



The University of  
**Nottingham**

UNITED KINGDOM • CHINA • MALAYSIA

---

Integrated modelling of electrical energy  
systems for the study of residential  
demand response strategies

---

Ana SANCHO TOMÁS

*Supervised by:*

Prof. Darren ROBINSON

Prof. Mark SUMNER

*A thesis submitted to the University of Nottingham  
for the degree of Doctor of Philosophy*

*in the*

Laboratory of Urban Complexity And Sustainability  
Faculty of Engineering

AUGUST 2017



*To Georgios Kazas.*



## Abstract

Building and urban energy simulation software aim to model the energy flows in buildings and urban communities in which most of them are located, providing tools that assist in the decision-making process to improve their initial and ongoing energy performance. To maintain their utility, they must continually develop in tandem with emerging technologies in the energy field. Demand Response (DR) strategies represent one such family of technology that has been identified as a key and affordable solution in the global transition towards clean energy generation and use, in particular at the residential scale.

This thesis contributes towards the development and application of a comprehensive building and urban energy simulation capability that parsimoniously represents occupants' energy using behaviours and responses to strategies to influence them. This platform intends to better unify the modelling of Demand Response strategies, by integrating the modelling of different energy systems through Multi Agent Simulation, considering stochastic processes taking place in electricity demand and supply. This is addressed by: (a) improving the fidelity of predictions of household electricity demand, using stochastic models, (b) demonstrating the potential of Demand Response strategies using Multi-Agent Simulation and machine learning techniques, (c) integrating a suitable model for the low voltage network to study and incorporate effects on the grid, (d) identifying how this platform should be extended to better represent human-to-device interactions; to test strategies designed to influence the scope and timing of occupants' energy using services.

Stochastic demand models provide the means to realistically simulate power demands, which are subject to naturally random human behaviour. In this work, the

power demand arising from small household appliances is identified as a stochastic variable, for which different candidate modelling methods are explored. Variants of two types of stochastic models have been tested, based on discrete time and continuous time stochastic processes. The alternative candidate models are compared and validated using Household Electricity Survey data, which is also used to test strategies, informed by advanced cluster analysis techniques, to simplify the form of these models.

The recommended small appliance model is integrated with a Multi Agent Simulation (MAS) platform, which is in turn extended and deployed to test DR strategies, such as load shifting and electric storage operation. In the search for optimal load-shifting strategies, machine learning algorithms, Q-learning in particular, are utilised. The application of this new developed tool, No-MASS/DR, is demonstrated through the study of strategies to maximise the locally generated renewable energy of a single household and a small community of buildings connected to a Low Voltage network.

Finally, an explicit model of the Low Voltage (LV) network has been developed and coupled with the DR framework. The model solves for power-flow analysis of a general low-voltage distribution network, using an electrical circuit-based approach, implemented as a novel recursive algorithm, that can efficiently calculate the voltages at different nodes of a complex branched network.

The work accomplished in this thesis contributes to the understanding of residential electricity management, by developing better unified modelling of Demand Response strategies, that require integrated modelling of energy systems, with a particular focus on the study of maximising locally generated renewable energy.

## Acknowledgements

This thesis is the result of the work of many people. I have been constantly supported and encouraged by my supervisors, Prof. Darren Robinson and Prof. Mark Sumner, who were always available for discussions, interested in my work and providing me with valuable guidance. My sincerest gratitude goes to my thesis committee, Prof. Mark Johnson and Prof. David Infield, for taking the time to read my thesis and make excellent comments. My special thanks go to Dr Jacob Chapman, for his patience and support, and for making work so enjoyable and fun.

I am very grateful to all my colleagues in Nottingham, and the time I have shared with you in the last three years. Thanks to all of those in B14, for all the good times inside but also outside the office. Thank you for our friendship and, of course, thank you for the coffees, pictionary, Wollaton Park, the Depot and so much more.

This PhD would not have been the same without the Ci-ENERGY family, with whom research became a unique and incredible journey. I will be forever grateful for all the time we shared together, which made me grow professionally and, more importantly, personally. I have learnt so much from each of you.

Dedico esta tesis a mi familia y amigos, gracias por vuestro afecto incansable. A mis amigas, que me ayudaron y ayudan a no sentirme lejos nunca. A mi madre y hermana, por ser las grandes mujeres que son. Con especial cariño dedico esta tesis a mi padre, que me inculcó el valor del esfuerzo y la perseverancia, pilares básicos para acabar este PhD.

A Pablo, por ser mi familia y mi hogar durante todo este tiempo.



# Contents

|   |             |
|---|-------------|
| <b>List of Figures</b>  | <b>xiii</b> |
| <b>List of Tables</b>   | <b>xvii</b> |
| <b>List of Code examples</b>                                  | <b>xix</b>  |
| <b>1 Introduction</b>   | <b>1</b>    |
| 1.1 Motivation . . . . .                                      | 1           |
| 1.2 Objectives . . . . .                                      | 3           |
| 1.3 Research approach . . . . .                               | 6           |
| 1.4 Contribution to knowledge . . . . .                       | 7           |
| 1.5 List of publications . . . . .                            | 8           |
| 1.6 Thesis structure . . . . .                                | 9           |
| <b>2 Stochastic Methods for Appliance Modelling</b>           | <b>11</b>   |
| 2.1 Introduction . . . . .                                    | 11          |
| 2.2 Literature Review . . . . .                               | 13          |
| 2.3 Proposed modelling tasks . . . . .                        | 17          |
| 2.4 Methods: modelling fractional energy . . . . .            | 19          |
| 2.4.1 Discrete-time Markov processes . . . . .                | 19          |
| 2.4.2 Matrix dimensioning: Density based clustering . . . . . | 21          |

|          |  |           |
|----------|--|-----------|
| 2.4.3    | Survival analysis . . . . .  | 23        |
| 2.4.4    | Monte Carlo simulation of fractional energy states . . . . .   | 25        |
| 2.5      | Validation methods . . . . .   | 30        |
| 2.5.1    | Cross validation . . . . .   | 30        |
| 2.5.2    | Time series analysis . . . . .   | 31        |
| 2.5.3    | Sensitivity and specificity analysis . . . . .   | 33        |
| 2.5.4    | Application of validation methods . . . . .  | 34        |
| 2.6      | Summary . . . . .  | 35        |
| <b>3</b> | <b>Small Appliance Modelling</b>   | <b>39</b> |
| 3.1      | Household Electricity Survey data set . . . . .  | 39        |
| 3.1.1    | Audio-visual category . . . . .  | 41        |
| 3.1.2    | Data preprocessing: outliers and maximum energy estimation   | 43        |
| 3.2      | Results and discussion . . . . .   | 45        |
| 3.2.1    | Application of Markov model . . . . .  | 46        |
| 3.2.2    | Application of Survival analysis . . . . .   | 48        |
| 3.2.3    | Approach selection . . . . .   | 49        |
| 3.2.4    | Application of Survival Multistate approach to other categories  | 57        |
| 3.2.5    | Global performance and application of the model . . . . .  | 61        |
| 3.3      | Conclusion . . . . .   | 64        |
| <b>4</b> | <b>Generalisation of multi-agent stochastic simulation architecture<br/>to support Demand Response</b> | <b>69</b> |
| 4.1      | Introduction . . . . .   | 70        |
| 4.2      | Formulation of DR software requirements . . . . .  | 74        |
| 4.3      | Literature Review . . . . .  | 75        |

---

|          |   |            |
|----------|---|------------|
| 4.3.1    | DR methods and techniques . . . . .   | 77         |
| 4.3.2    | Application into DR simulation software tools . . . . .                       | 81         |
| 4.3.3    | Proposed modelling framework . . . . .  | 83         |
| 4.4      | No-MASS background . . . . .  | 85         |
| 4.4.1    | No-MASS workflow . . . . .  | 87         |
| 4.4.2    | No-MASS for Demand Response . . . . .   | 88         |
| 4.5      | Methodology I: models and data in No-MASS/DR . . . . .                        | 89         |
| 4.5.1    | Electricity demand models (R1) . . . . .                                      | 89         |
| 4.5.2    | Supply and power flow (R2) . . . . .  | 91         |
| 4.6      | Methodology II: Software orchestration with<br>Agent Representation . . . . . | 93         |
| 4.6.1    | Agent representation . . . . .  | 93         |
| 4.6.2    | Type of agents . . . . .  | 94         |
| 4.7      | Methodology III: Agent Learning for<br>DR optimisation (R3) . . . . .         | 95         |
| 4.7.1    | Q-learning algorithm . . . . .  | 96         |
| 4.7.2    | DR mechanisms . . . . .   | 98         |
| I.       | Load shifting optimisation . . . . .  | 98         |
| II.      | Battery discharge operation . . . . .   | 100        |
| 4.7.3    | Self-consumption maximization . . . . .                                       | 102        |
| 4.8      | Summary . . . . .   | 102        |
| <b>5</b> | <b>Application of the extended No-MASS/DR framework</b>                       | <b>105</b> |
| 5.1      | Improving Self-Consumption . . . . .  | 105        |
| 5.2      | Case study I - maximise self-consumption for a single building . . . . .      | 108        |
| 5.2.1    | Scenario description . . . . .  | 108        |

---

|          |   |            |
|----------|---|------------|
| 5.2.2    | Reward functions . . . . .  | 109        |
|          | Appliance re-schedule . . . . .   | 109        |
|          | Battery discharge . . . . .   | 110        |
| 5.2.3    | Results and discussion . . . . .  | 112        |
| 5.3      | Case study II: community of buildings . . . . .                           | 116        |
| 5.3.1    | Scenario description . . . . .  | 116        |
| 5.3.2    | Implementation: multilevel coordination . . . . .                         | 117        |
| 5.3.3    | Reward functions . . . . .  | 118        |
|          | Appliance re-schedule . . . . .   | 118        |
|          | Battery discharge . . . . .   | 119        |
| 5.3.4    | Results and discussion: Improving community<br>self-consumption . . . . . | 122        |
| 5.4      | Conclusion . . . . .  | 127        |
| <b>6</b> | <b>Low Voltage Network modelling</b>                                      | <b>131</b> |
| 6.1      | Introduction . . . . .  | 131        |
| 6.2      | LV network modelling . . . . .  | 136        |
| 6.3      | Load-flow analysis . . . . .  | 137        |
|          | 6.3.1 Time granularity for load-flow simulation . . . . .                 | 141        |
| 6.4      | Load-flow analysis as a recursion algorithm . . . . .                     | 142        |
|          | 6.4.1 The model: pre-requisites and assumptions . . . . .                 | 142        |
|          | 6.4.2 The load flow algorithm . . . . .                                   | 143        |
|          | 6.4.3 Pseudo-code . . . . .   | 145        |
| 6.5      | Application to a case study . . . . .                                     | 147        |
| 6.6      | Conclusion . . . . .  | 152        |

---

|          |  |            |
|----------|--|------------|
| <b>7</b> | <b>Conclusions</b>                             | <b>155</b> |
| 7.1      | Introduction . . . . .                         | 155        |
| 7.2      | Achievements . . . . .                         | 155        |
| 7.2.1    | Objective I . . . . .                          | 155        |
| 7.2.2    | Objective II . . . . .                         | 156        |
| 7.2.3    | Objective III . . . . .                        | 158        |
| 7.3      | Further work and recommendations . . . . .     | 158        |
|          | Ideas to improve R1-R3 . . . . .               | 159        |
|          | Ideas to address R4 and R5 . . . . .           | 161        |
|          | Ideas to improve R6 . . . . .                  | 163        |
|          | <b>Appendices</b>                              | <b>165</b> |
| A        | Small Appliance Modelling parameters . . . . . | 165        |



# List of Figures

|             |  |    |
|-------------|--|----|
| Figure 1.1: | No-MASS and proposed extension to No-MASS/DR. . . . .  | 4  |
| Figure 1.2: | DR methodology proposed. . . . .   | 6  |
| Figure 1.3: | Thesis structure. . . . .  | 9  |
| Figure 2.1: | Exponential decay of survival times for eleven different frac-<br>tional energy states. Frequency is normalised for the distri-<br>bution. . . . . | 24 |
| Figure 2.2: | Details of data fitted to Weibull distribution using non-<br>linear least square method. . . . .   | 25 |
| Figure 2.3: | Simulation flowcharts. . . . .   | 28 |
| Figure 2.4: | Example of time series decomposition into trend, daily vari-<br>ation and remainder component. . . . .   | 32 |
| Figure 3.1: | Annual energy use of the types of appliances considered in<br>the modelling. . . . .   | 41 |
| Figure 3.2: | Distribution of audio-visual category data, for the values of<br>fractional energy over a day. . . . .   | 44 |
| Figure 3.3: | Fractional energy from data clusters. . . . .  | 47 |
| Figure 3.4: | Comparison of daily variation components. . . . .  | 50 |
| Figure 3.5: | Temporal dependency of RMSE, calculated for each 10-<br>minute time slot, for all states of each approach. . . . .                                 | 52 |

---

|   |     |
|---|-----|
| Figure 3.6: Daily profile residuals of fractional energy. . . . .   | 54  |
| Figure 3.7: Boxplot comparing observed and predicted total energy use<br>over the validation period. . . . .                | 56  |
| Figure 3.8: Summary of validation results . . . . .   | 57  |
| Figure 3.9: Effect of a dominant appliance (kettle) on the observed and<br>simulated data for the kitchen category. . . . . | 59  |
| Figure 3.10: Example of one-day simulation for low-load appliances. . . . .   | 62  |
| Figure 3.11: Example of one-day simulation for aggregates of appliances. . . . .  | 63  |
| Figure 3.12: Averaged daily total energy usage from the different cate-<br>gories. . . . .                                  | 64  |
| Figure 4.1: Trends in technology uptake. . . . .  | 71  |
| Figure 4.2: Demand Response categories. Source: [39]. . . . .   | 76  |
| Figure 4.3: Device-to-device and person-to-person interaction. . . . .  | 84  |
| Figure 4.4: No-MASS flowchart. . . . .  | 85  |
| Figure 4.5: Switch on probabilities of large appliances. . . . .  | 90  |
| Figure 4.6: Power flows. . . . .  | 93  |
| Figure 4.7: DR methodology. . . . .   | 96  |
| Figure 4.8: Appliance re-schedule diagram. . . . .  | 99  |
| Figure 4.9: Q-tables. . . . .   | 100 |
| Figure 5.1: Reward function uses price signals. . . . .   | 109 |
| Figure 5.2: Appliance learning: improvement of reward. . . . .  | 111 |
| Figure 5.3: 24h of simulation for a winter day. . . . .   | 114 |
| Figure 5.4: Import and export for the three scenarios, for summer and<br>winter. . . . .                                    | 115 |

---

|   |     |
|---|-----|
| Figure 5.5: Self consumption index for the scenarios, for summer and winter. . . . .  | 116 |
| Figure 5.6: Multilevel coordination. . . . .  | 119 |
| Figure 5.7: Reward values dependent on tariff. . . . .  | 121 |
| Figure 5.8: Imports and exports from/to the grid. . . . .   | 123 |
| Figure 5.9: Self-consumption index for the three scenarios, in winter and summer. . . . .   | 125 |
| Figure 5.10: Cost reduction when load shifting and battery are introduced. Values based on unit tariff signal proposed. . . . .   | 127 |
| Figure 6.1: Power system . . . . .  | 133 |
| Figure 6.2: Description of circuit-based methods. . . . .   | 140 |
| Figure 6.3: Power grid time-scales. . . . .   | 141 |
| Figure 6.4: Diagram of network, relating parent and child nodes. . . . .  | 144 |
| Figure 6.5: Meadows network layout. . . . .   | 147 |
| Figure 6.6: Network scheme. . . . .   | 148 |
| Figure 6.7: Voltage levels through a day in winter and a day in summer, for nodes at the end of the lines in each branch. . . . . | 151 |
| Figure 6.8: Imaginary modified topology for performance comparison. . . . .   | 152 |
| Figure 6.9: Distribution of simulation times to achieve convergence in one time-step, for both topologies. . . . .                | 152 |
| Figure 7.1: Linear behavioural model. . . . .   | 163 |



# List of Tables

|      |   |     |
|------|---|-----|
| 3.1  | Types of appliance ( $j$ ) available in the audio-visual category . . . . .   | 42  |
| 3.2  | Outlier filtering for appliances in the audio-visual category . . . . .   | 46  |
| 3.3  | Hierarchical levels of clustering considered for $\{t, f\}$ space partition,<br>with number of clusters and number of noise cells identified. . . . . | 48  |
| 3.4  | Summary of validation time series decomposition. . . . .  | 51  |
| 3.5  | Accuracy of model. . . . .  | 53  |
| 3.6  | RMSE values of the daily profile results. . . . .   | 53  |
| 3.7  | Residual error between observation and simulation, for mean and<br>median of the total energy use over the validation period. . . . .                 | 55  |
| 3.8  | Number and type of parameters needed for the different type of model.   | 56  |
| 3.9  | Summary of results on application of Survival Multistate approach<br>to other categories. . . . .   | 60  |
| 3.10 | Available appliances in house "103028". . . . .   | 60  |
| 4.1  | Examples of optimization problems and methods in DR. . . . .  | 78  |
| 5.1  | Average weekly import/export. . . . .   | 122 |
| 5.2  | Average self-consumption (%). . . . .   | 126 |
| 6.1  | Actors and roles in the transmission and distribution power system.   | 134 |
| 6.2  | Cable lengths on four branches. . . . .   | 149 |

|     |   |     |
|-----|---|-----|
| A.1 | Survival distribution parameters ( $\gamma$ : location, $k$ : shape, $\lambda$ : scale)<br>for four categories. . . . . | 166 |
| A.2 | Median fractional energy for transforming energy states into frac-<br>tional energy profiles. . . . .                   | 166 |
| A.3 | Hourly state probability for Audio-visual and Computing categories.   | 167 |
| A.4 | Hourly state probability for Kitchen and Other categories. . . . .  | 168 |
| A.5 | Sum of maximum power (W) for all the corresponding low-load<br>appliances in each household. . . . .                    | 169 |

# List of Algorithms

|   |  |     |
|---|--|-----|
| 1 | FORWARDSWEEP()   | 146 |
| 2 | BACKWARDSWEEP( $V_{j-1}$ )   | 146 |
| 3 | POWERFLOW  | 146 |
| 4 | SIMULATE SMALL APPLIANCE USAGE ( $[\lambda, k, \gamma]_{\mathbf{s}}, P(t, \mathbf{s})$ ) | 165 |



# Chapter 1

## Introduction

### 1.1 Motivation

The uptake of low carbon technologies is growing worldwide, as a consequence of political regulation to reduce carbon emissions. In the UK, the domestic sector is responsible for 20% of such emissions, and current trends indicate larger shares in the future, as a result of the electrification of heat and transport. Electrical devices and appliances used in households are thus becoming very important, and it is crucial to understand how and why they are being used. Deployment of Distributed Generation (DG) sources at domestic scale is also expanding, bringing new challenges to manage and control transmission and distribution of electricity. The energy generation infrastructure is evolving, and new and improved technology solutions are emerging, such as energy storage devices and Demand Response solutions. Generally, these solutions may result in capital investment benefits for utility companies, as they can defer the need for costly upgrades in the transmission and distribution systems.

Building and urban energy simulation software aim to model the energy flows

in buildings and urban communities in which most of them are located, providing tools that assist in the decision-making process to improve their initial and ongoing energy performance. It is therefore important that building and urban simulation software is able to address new technological opportunities, such as those mentioned above, using appropriate computing modelling techniques. This task is complicated by the fact that in real households, there is a large variability in occupants' energy using behaviours, referring to demand patterns, device interaction or consumption management, which significantly influence electricity and heating demands. These stochastic behavioural processes should therefore be represented in our computer simulations.

This thesis aims to address the above challenges; contributing towards the achievement of a comprehensive building and urban energy simulation capability that parsimoniously represents occupants' energy using behaviours and responses to strategies to influence them; demonstrating the application of this new platform, No-MASS/DR, through the study of strategies to maximise the locally generated renewable energy of a single household and a small community of buildings connected to a Low Voltage network.

## **Research context**

A Multi-Agent Stochastic Simulation platform called No-MASS has been developed in previous research (Figure 1.1), whose purpose is to augment existing building/urban simulation software to consider the stochastic actions of occupants. Thus far it has mostly focused on thermal energy effects by implementing models of occupants' locations and presence, activities and associated metabolic heat gains and their interactions with windows, lights and shading devices; accounting also for

negotiations between peers that influence these interactions. These models allow building simulation software to more accurately and realistically calculate heat flows and associated energy demands.

Electrical energy loads have also previously been studied through models related to appliance usage: both appliance ownership and appliance use, whether conditional or not on the associated activities. However, these models have only considered *large* devices, meaning those that individually have a significant energy use and are commonly owned (e.g. fridge, dishwasher, washing-machine, cooker, microwave and TV). There is no such model of small electrical appliances that may not be significant individually, but are when considered as like groups (aggregates of similar type).

An extension to No-MASS is proposed in this thesis to also consider Demand Response modelling, as depicted in Figure 1.1.

## 1.2 Objectives

This thesis contributes to knowledge by *a)* improving the fidelity of predictions of household electricity demand, *b)* demonstrating the potential of Demand Response strategies using Multi-Agent Simulation and machine learning techniques and *c)* integrating a suitable model for the low voltage network to study and incorporate effects on the grid. The objectives of this thesis are therefore to:

**Objective I.** Develop stochastic models of demand for the use of small appliances in homes (to complement existing models of relatively large appliances). This requires that a range of modelling strategies are compared selecting the most parsimonious.

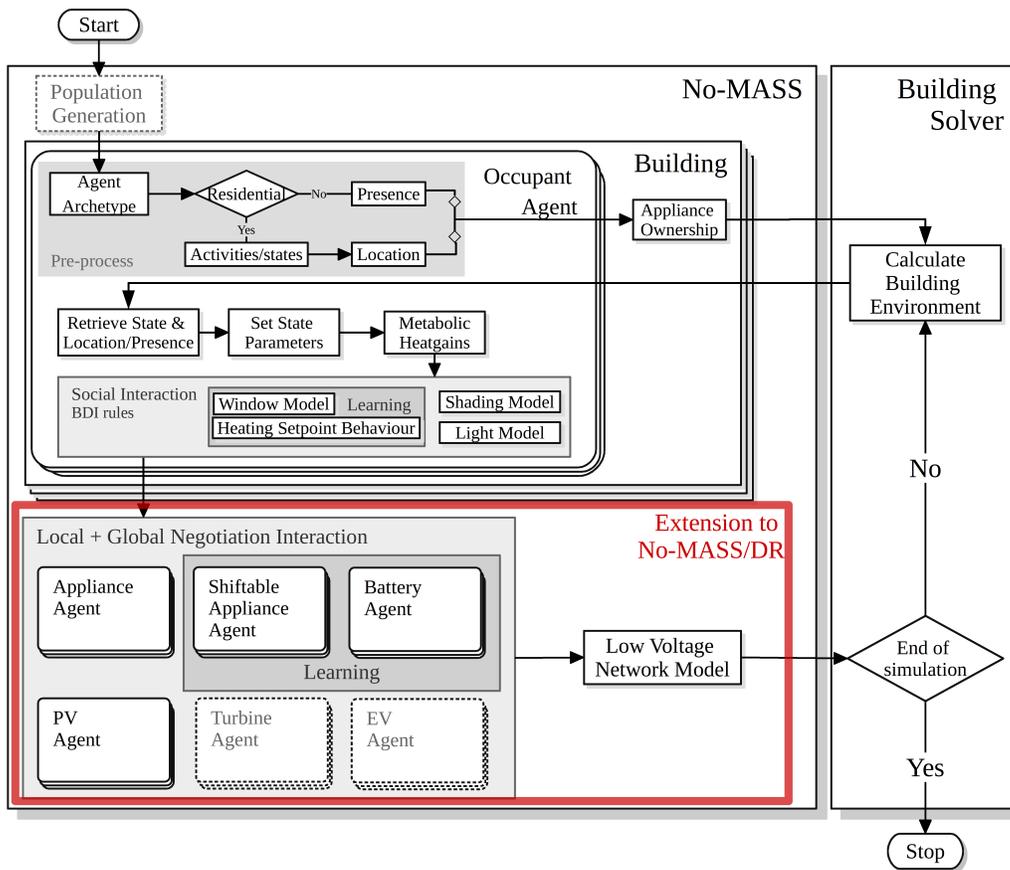


Figure 1.1: No-MASS and proposed extension to No-MASS/DR.

Key requirements here are to:

- a. Account for the time-dependency of electricity demand due to small appliances.
- b. Realistically describe demand in individual households, through appropriate aggregation, either modelling appliances individually or as a group.

**Objective II.** Demonstrate the potential for Multi-Agent Simulation and machine learning algorithms to evaluate DR strategies, and propose a feasible methodology to investigate the maximization of renewable self-consumption in residential communities. To achieve this it is necessary to:

1. Identify a set of requirements that a DR simulation platform should aim to satisfy, including the optimised use of local power sources, power loads and storage devices and its impact on the power grid, but also enabling interactions between devices and users.
2. Test the use of Multi Agent Simulation as a software architecture to address these requirements.
3. Test the use of agent learning as a method for simulating DR strategies (appliance re-schedule and battery operation) that support maximization of on-site renewable energy use.
4. Explore applications for a single house and for a community of buildings.

A further general goal of this work is to develop and couple it with building/urban energy simulation software using exclusively open-source software, to maximise accessibility by the international research community and thereby its impact.

**Objective III.** Build a model for load flow analysis of a residential low voltage distribution network, which satisfies the following requirements:

1. The source code of the model should be available for its future implementation with building/urban energy simulation software.
2. Load-flow analysis is efficiently performed for a typical radial distribution network, including branched layouts.
3. The nodes of the network may represent either electricity use from single appliances or aggregated power for the household.
4. Models (or input data) for microgeneration, electrical storage devices or Electric Vehicles (EV) can be easily added to the network model.

### 1.3 Research approach

Objective I is achieved by developing data-driven models, which are based on stochastic methods. More specifically, discrete (Markov) time processes and continuous (survival) time processes. For this thesis, the Household Electricity Survey dataset [1] was analysed for parameter extraction. The methodology presented in this work has been thoroughly validated using 10-fold cross validation on the same dataset.

Objective II has been implemented as a Multi-Agent Stochastic Simulation, where occupants and electrical devices are represented as agents. Each agent has a range of properties and the ability to negotiate and interact with other agents, in order to achieve individual and common goals. Some device agents are conferred with learning intelligence to achieve such goals, via machine learning algorithms (Q-learning in particular). Two DR mechanisms are implemented: load shifting and battery discharge, for which the learning agents use information about the system (such as the electricity cost signal, total power demands or renewable energy available). The strategies studied in this thesis are focused on improving electricity use from renewable local sources.

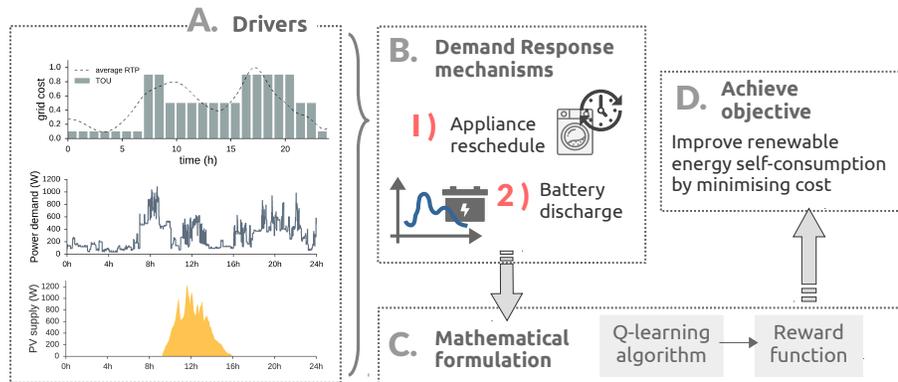


Figure 1.2: DR methodology proposed.

Objective III uses a forward/backward sweep method to solve the power-flow. Computationally speaking, it has been implemented as a recursive algorithm, which solves branched network layouts very efficiently. In its early stages, the method was validated against equivalent MATLAB Simulink models, producing the same results, with improvements in solver speed. MATLAB Simulink is a well-established software, extensively used for power-flow simulation. However, it requires an expensive license, failing in our open-source requirement as stated above.

## 1.4 Contribution to knowledge

The primary contribution of this thesis is the delivery of a simulation framework that, when used as stand-alone software, is able to handle simultaneously *i)* stochastic power demands, *ii)* device-to-device interactions for energy balancing and DR of electrical equipment and *iii)* load flow analysis, complementing existing functionality to handle occupants' behaviour and associated human-to-human interaction. Moreover, it has been coupled with building/urban energy simulation software and it is readily extensible to emulate device-to-human interactions.

More specifically, this thesis presents an effective methodology to model groups of small appliances, even when they are of a different type. It provides a mathematical formulation to reduce Markov's transition matrix' dimensionality, by using a state-of-the-art density clustering algorithm, with potential applications beyond the energy modelling community.

It also puts into practice a novel implementation to perform load-flow analysis with a forward/backward sweep method using Object Oriented programming and recursion, which is suitable to be integrated into a Demand Response simulation framework.

## 1.5 List of publications

- A. Sancho Tomás, et al. On the energy demands of small appliances in homes. *Energy Procedia* 78 (2015): 3384-3390.
- A Sancho Tomás, M. Sumner and D. Robinson. A generalised model of electrical energy demand from small household appliances. *Energy and Buildings* 135 (2017): 350-366.
- A. Sancho Tomás, et al. Extending No-MASS: Multi-Agent Stochastic Simulation for Demand Response of residential appliances. *Building Simulation 2017*, August 2017.
- A. Sancho Tomás, et al. On the energy demands of small appliances in homes. *IEEE Transactions on Smart Grid* In preparation.

## Foreword

The work presented in Chapters 4 and 5 is the result of a close collaboration with Dr J. Chapman. As part of our goal to produce robust software, this collaboration benefited from Chapman's advanced computing and programming skills, being responsible for developing C++ code to implement the models and algorithms used and debugging tasks. The author's contribution to those chapters is mainly based on the development, testing and application of the fundamental algorithms. Including also literature review, identification of software requirements, testing and using the software platform, finding appropriate parameters and reward functions, producing results, carrying out their analysis and identifying where and how improvements to the software are required.

## 1.6 Thesis structure

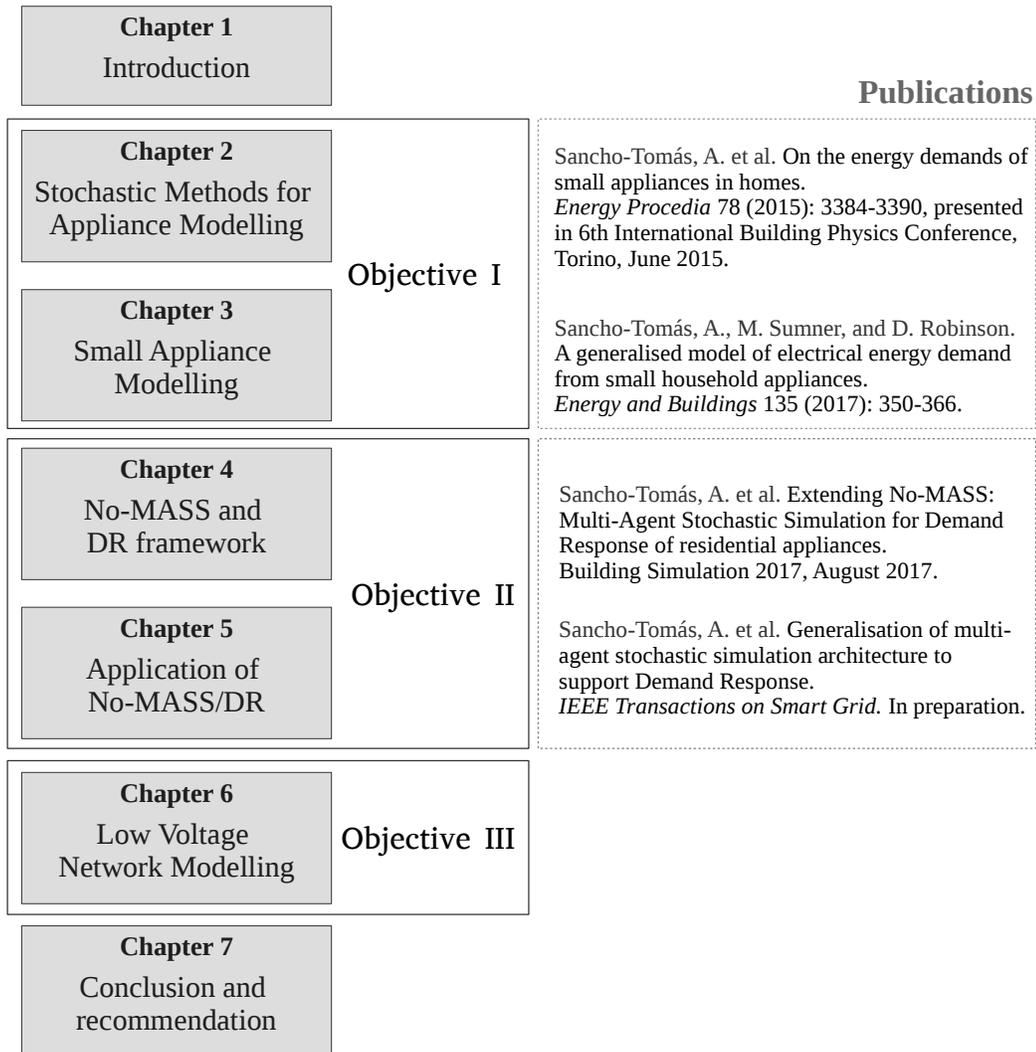


Figure 1.3: Thesis structure.



# Chapter 2

## Stochastic Methods for Appliance Modelling

Stochastic processes describe systems that are subject to random features mathematically. When modelling them, we do not attempt to model the occurrence of events (with respect to time) but the probability of occurrence of such events. Stochastic processes are thus represented with random variables. In our work, we identify the power demand coming from small appliances to be a stochastic variable. Based on that, we explore different methods to model the electricity usage of appliances in homes. In this chapter, the mathematical methods and techniques used are explained and developed in detail. In particular, two types of models have been tested: *a)* a discrete time discrete state process: Markov chain, whose transition matrix dimension has been reduced using cluster analysis; and *b)* a continuous-time and discrete state process: survival analysis.

### 2.1 Introduction

In the UK approximately 20% of energy use in households is due to electrical appliances [2], and this proportion is higher in better insulated homes. Residential electrical appliance use has direct implications for local Low Voltage (LV) networks, the loads on them and their integrity; and indirect implications for thermal energy

demands, since electrical energy is ultimately dissipated as heat, most of which is emitted within the building envelope. It is therefore important to be able to reliably predict electrical appliance use, in particular the magnitude and temporal variation of the energy use and power demand profiles arising from the aggregation of individual appliances, to support design and regulation of LV networks serving communities of buildings and of building's thermal systems.

But this is a complicated task, for the ownership and use of different types of appliance varies significantly from house to house, and between users. Addressing this diversity requires that we have an appropriate basis for allocating appliances to households depending on their composition and socio-economic characteristics and for predicting their subsequent use. This in turn implies the use of stochastic simulation and bottom-up approaches that may also facilitate the future testing of Demand Side Management (DSM) strategies.

So far, bottom-up approaches have focused on the modelling of high-load appliances: those that are commonly owned and which contribute significantly to total annual electricity use. Examples include cold (fridge and freezer), wet (washing machine and dishwasher) and cooking appliances. For example, the model of Jaboob [3] predicts when the appliances are switched on, the duration for which they will remain on and their fluctuating power demands whilst on. But in our everyday lives we also use myriad low-load appliances. Their individual share of energy use may be small, in some cases even negligible, but it is significant when considering them as a group (or groups).

As stated in section 1.2, one of the objectives of this thesis is to find a parsimonious strategy for modelling low-load appliances. In this we distinguish between four categories of appliance: audio-visual, computing, kitchen and other appli-

ances, which collectively account for those that are not represented by current device specific models.

The remainder of this chapter is organised as follows: section 2.2 introduces former work in the field; the modelling tasks proposed are outlined in section 2.3; in section 2.4 the mathematical methods employed in the modelling are described; finally the validation methods used are presented in section 2.5.

## 2.2 Literature Review

*Bottom-up* approaches describe the dynamics of a system by explicitly modelling the behaviour of the individual parts of that system. For the case of energy use, they consider the individual modelling of every end-use, or aggregates of them, in order to obtain aggregate profiles. These approaches are particularly promising given their potential for a) improving predictions of energy use of individual buildings or neighbourhoods when integrated with building energy simulation, b) sizing decentralised generation and storage devices, and c) testing Demand Side Management (DSM) strategies and rules for load management. Moreover, bottom-up approaches have the potential to explicitly include the effects of household composition and individuals' behavioural diversity.

Regarding appliance modelling, bottom-up approaches can be configured at different aggregation levels: from a pure microsimulation where each device is explicitly modelled, to strategies that consider aggregations of device for typologies of them.

Detailed microsimulation approaches are considered in probabilistic empirical models (as defined in [4]), which tend to model appliances one-by-one. Collected data, information on dwelling and household (occupants) characteristics, techni-

cal properties of appliances and aggregated values of energy use are combined in such approaches, and probabilistic methods are applied to generate results with profile diversity. Stokes' model [5] generates profiles at three aggregation levels: 30-minute-resolution average household, 30-minute-resolution specific household (with occupancy considerations) and 1-minute resolution specific household (including information relating to appliances' cycles). It considers 14 appliances plus miscellaneous, although only 9 different input monitored power cycles are taken into account, resulting in a limitation on the diversity of profiles generated, which in turn leads to poor results estimating the energy demand in the validation of the models. Paatero and Lund [6] introduce a social random factor (supposed to capture the social variety of the demand) that improves the diversity of patterns obtained; however, only yearly consumption data for the 16 end-uses is used for the generation of the models, together with other aggregate statistics, restricting the resolution to hourly time steps. In general, these approaches do not describe in terms of model parameters the dynamic behaviour of appliances, but they generate empirical profiles of power demand as a function of time.

Relatively more aggregated methods are models based on time-use-survey (TUS) datasets. In TUS datasets, the respondents fill in diaries of their activities during the day usually for one week periods, such as cooking, sleeping, travelling to work, etc. This data provides a powerful input to bottom-up models, since it encapsulates highly detailed information describing occupants' activities, that can be related to the use of appliances. To this end, Capasso [7] presents a first strategy linking occupants' activities with appliance use, using TUS data. The model produces 15-minute profiles of electricity use, considering aggregations of appliance that correspond to just four type of activities: cooking, housework, leisure and

hygiene; each associated with a blend of large and small appliances, which are allocated by considering the average range of appliances present in the simulated household. The relation between performing an activity and using an appliance is described with a single coefficient  $\alpha$  (defined as a human resources). Tanimoto [8] combines TUS with statistical data of ownership of appliances and its peak and stand-by powers. 31 activities are considered in this case, so that the level of aggregation is low, but this is contrasted by a small dataset size (58 households over 2 days).

In a similar vein, Widén [9] uses Swedish TUS data to model electricity use by assigning appliances to related activities (9 different categories in this case) and imposing five standard end-use profiles based on the type of their demand profile: demand disconnected from activity, power demand constant during activity, power demand constant after activity (with and without addition of temporal constraint) and fluctuating power demand (only applied to lighting). This approach is further developed in [10], where inhomogeneous Markov chains generate sequences of domestic activities that have an impact on power demand (5 minute and 1 hour granularity), including dependencies with the number of occupants performing these activities. A yet finer temporal resolution of 1 minute is achieved in the work developed by Richardson et al.[11]. Based on 7 different activities, a load curve for the appliances is created using the probability of switching on an appliance when an activity is being performed, and applying a fixed power conversion scheme. Using a calibration procedure based on the total time of use of an appliance, they obtain annual energy predictions. Although this tuning ensures a good overall match in annual energy demand, this does not imply the absence of compensating errors in the modelling of different appliance typologies, or that the

dynamic characteristics of appliance use are well represented.

Although activity modelling is a promising method to obtain accurate energy demand profiles, this activity-appliance pairing approach does not facilitate the modelling of the range of appliances, because the activities that are recorded in time use surveys is insufficiently detailed, limiting the applicability of this approach to the modelling of either relatively high-load appliances or aggregations of small and large appliances for which there is weak empirical evidence. There has been no rigorous validation of those bottom-up modelling strategies to date, whether these are based on TUS data or not, to demonstrate their ability to faithfully capture energy use/power demand dynamics. These methods also have no rigorous basis for modelling the dependency of appliance ownership and related use characteristics as a function of household socio-demographic composition.

In partial response to these shortcomings, Jaboob [3] assigns (exclusively large) appliances to households as a function of their socio-demographic characteristics. The activities of the members of these households are then predicted, from which the conditional likelihood that related appliances will be switch on is modelled, as is the corresponding duration that they will remain on and their time-varying mean power demands whilst on. Thus, this modelling chain rigorously resolves for dynamic variations in mean power demand, in contrast to static power conversion schemes. Moreover, it presents the possibility of being used together with explicit models of low-load appliances, in order to obtain accurate values of the total electricity use of a house.

To this end and informed by these past endeavours, our task is to develop a parsimonious strategy for the use of relatively low-load appliances, in complement to Jaboob's model of high-load appliances.

## 2.3 Proposed modelling tasks

In this work we are interested in modelling the energy and power demands of low-load appliances to support building, systems and network design. In order to contribute to accurate predictions of residential energy use, we need to address the diversity in dwelling characteristics and human behaviours. Thus, we identify the following modelling tasks:

- I Low-load appliances are categorised into four groups: audio-visual, computing, small kitchen and other (miscellaneous housework, garden and personal care appliances). This classification keeps the possibility of linking the modelling of low-load appliances with occupants' activity modelling, which in turn allows for considering socio-demographic factors. Low-load appliance allocation is performed using a random sampling of the total rated power (sum of all devices in the category) of aggregates of appliances, based on the available data set.
- II Model the characteristic use of these appliances in individual households. To this end, we utilise the fractional energy use  $f(t)$ : the ratio of the actual to the maximum energy  $E_{max\ k,i}$  that an appliance  $i$  belonging to a category  $k$  can use, determined by its rated power. Modelling  $f(t)$ , we can distinguish between:
  - Switching on/off events.
  - Fluctuating demands whilst the appliances are in use.

The modelling of fractional power can be applied to modelling single devices ( $f_j(t)$ , where  $j$  is a specific type of appliance) or aggregates of them ( $f_k(t)$ )

where  $k$  refers to a category of appliances).

Two considerations need to be taken into account in carrying out these tasks. Firstly, stochastic methods are required, as we are interested in describing the underlying randomness in households' appliance use and investment decisions. These methods rely on the definition of coefficients that represent the system as a probability distribution, which can be dependent on different variables such as time of the day, number of occupants, weather, etc. Secondly, using the normalized fractional energy of individual appliances  $f_j(t)$  instead of absolute energy allows us to evaluate load profiles from different appliances of a similar type, but that do not necessarily have the same magnitude. In this way, appliances can be classified into groups and modelled as a category ( $f_k(t)$ ).

Candidate techniques that have been used to good effect in the modelling of occupants' behaviours include Bernoulli processes (activities [3]), discrete-time random or Markov processes (presence [12], blinds [13], windows [14]) and continuous-time random processes (blinds [13], windows [14]): the latter being a hybrid between discrete and continuous time random process models.

Furthermore, it has been previously shown [3, 9, 11] that stochastic methods are successful in describing energy demands and the information listed in Task II. In the work here presented, two of these statistical approaches have been exploited:

- *Discrete-time Markov processes* can model the probability of transitions occurring between energy states  $s(t)$ , with or without time dependency. Energy states are the result of discretising the range of fractional energy values. This discretization process can be more efficiently achieved if complemented with clustering techniques.
- *Survival analysis* can model the switching-on/off of appliances, as well as the

duration an appliance remains in different energy states.

In the methodology presented here we have tested a range of strategies in order to find the most parsimonious approach. In this, we have ensured that the number of subjective decisions needed for modelling have been minimised, so that the methodology can be appropriately applied independently of the data set employed to estimate the models' coefficients.

## 2.4 Methods: modelling fractional energy

### 2.4.1 Discrete-time Markov processes

A Markov process is a stochastic process that fulfils the Markov property, by which a future state depends on the most recent state, and not on any prior history [15]. A stochastic process  $X(t)$  is therefore a Markov process if for every  $n$  and  $t_1 < t_2 < \dots < t_n$ :

$$P[X(t_n) = x_n | X(t_{n-1}) = x_{n-1}, \dots, X(t_1) = x_1] = P[X(t_n) = x_n | X(t_{n-1}) = x_{n-1}]. \quad (2.1)$$

Markov chains describe the process of making transitions between a present state  $i$  to a future state  $j$ , according to a probability distribution, described by a state transition probability matrix (or Markov matrix) as follows:

$$P_{ij} = \begin{pmatrix} p_{11}(t) & p_{12}(t) & \dots & p_{1m}(t) \\ p_{21}(t) & p_{22}(t) & \dots & p_{2m}(t) \\ \vdots & \vdots & \ddots & \vdots \\ p_{m1}(t) & p_{m2}(t) & \dots & p_{mm}(t) \end{pmatrix}, \quad (2.2)$$

where

$$p_{ij}(t) = \frac{n_{ij}(t)}{n_i(t)} = \frac{n_{ij}(t)}{\sum_j n_{ij}(t)} \quad (2.3)$$

is the probability that a transition from  $i$  to  $j$  takes place, given by the ratio of transitions that occur to state  $j$  from  $i$  to the total number of transitions occurring from  $i$ .

The dimensions of a Markov matrix  $m \times m$  are given by the number of states  $m$  defined in the system. At the same time, the coefficients in the matrix may or may not depend on time. In the first instance, a time-homogeneous Markov process is considered, where the system can be described using a single matrix. We then consider a time-inhomogeneous Markov chain, in which the number of matrices  $r$  is given by the number of time slots considered to have different transition probabilities. For instance, if it is assumed that the probabilities are different for each hour of a day, then  $r = 24$  (considering a single-day). This means that the probability distribution is given by a matrix of dimension  $r \times m \times m$ .

Appropriate dimensioning of the Markov matrices is not a trivial task: if  $m$  and  $r$  are set too low or even equal to 1, the dynamics or temporal variation of the system may not be suitably described by the model. On the other hand, if  $m$  and  $r$  are set too high, there is a risk of performing redundant calculations, adding unnecessary computing complexity, as well as a risk of overfitting the model.

In our case, fractional energy  $f(t)$  is a continuous variable with values between 0 and 1, that is discretized into  $m$  energy states  $s$ . The time variable  $t$  is discrete, and it takes values every 10 minutes, but it can also be divided into  $r$  temporal states. In this sense, the subdivision chosen of the two-dimensional space  $\{t, f\}$  generated by the time of the day and the fractional energy of a category of appliances will set the values of  $m$  and  $r$ , that determine the dimension of the transition

matrix. Consequently, estimating an adequate and efficient subdivision of  $\{t, f\}$  is key in the formulation of a parsimonious model. In our search for an objective methodology, clustering techniques were identified as good candidates to evaluate the partitioning of this space.

### 2.4.2 Matrix dimensioning: Density based clustering

Cluster analysis techniques provide a powerful and systematic mechanism for identifying groups or common features of a database  $D$  of  $n$  objects. There are a large number of clustering algorithms, two of the main being hierarchical and partitional [16] algorithms. The former decomposes  $D$  into a nested hierarchy of clusters, represented by a dendrogram, i.e. a tree diagram that splits the database into subsets of smaller size, until each object belongs to one subset. The process can be agglomerative or divisive, depending on whether the structure is made from the leaves towards the root or from the root to the leaves. The latter creates a single-level partition of  $D$  into  $k$  clusters based on similarity and distance measures. The parameter  $k$  is required as an input, even though it is not generally known *a priori*.

A third type of clustering method is density-based clustering algorithms, which apply local cluster criteria [17] in order to classify  $D$ . They identify regions of high density that are separated from other clusters by regions of a low density of points, which can be classified as noise. Each object of the database is evaluated in terms of density in the neighbourhood, which has to exceed some threshold. Density-based clustering algorithms present some advantages over other types of clustering:

- (i) They are suitable for large data sets.
- (ii) Clusters may have irregular shapes.

- (iii) Although distance metrics are employed, clusters are identified based on density estimations of areas of the data set. The advantage of this is the identification of points that do not belong to any cluster, allowing for the treatment of unstructured (noise) points.

DBSCAN is a typical density-based clustering algorithm that was developed in 1996 [18]. The core idea behind DBSCAN is that for each object in a cluster, the neighbourhood of radius  $\epsilon$  has to be populated with a minimum number of points *MinPts*.  $\epsilon$  and *MinPts* are the two only parameters required.

However, the cluster structure of a real data set cannot usually be identified with a single global density parameter, but rather by clusters of different density, as well as their intrinsic structure. The OPTICS algorithm [17] is a generalization of DBSCAN. Instead of a clustering division, OPTICS outputs an ordering of the database relative to its density-based clustering structure, containing information for every density level up to a "generating distance"  $\epsilon_0$ , that allows for analysis of the grouping structure (hierarchy). A graphical interpretation of the ordering is available through a reachability plot [17], where clusters are identified as "dents" in the plot. The authors of this algorithm provide a method for automatically determining the cluster hierarchy using the information extracted from the reachability plot. However, a simpler alternative method for automatic extraction of the clusters is described in [19], in which the most significant clusters are simultaneously selected from different density levels. Interestingly the authors also show that reachability plots are equivalent to the dendrograms of single-link clustering methods.

### 2.4.3 Survival analysis

Survival analysis [20] models the waiting time until a given event occurs, also referred to as *survival time*. Let  $T$  be a non-negative continuous random variable representing the survival time until an on-appliance is switched off (or an off-appliance is switched on); with probability density function (p.d.f.)  $g(t)$ .  $g(t)$  can follow multiple distributions depending on the problem studied, and commonly  $g(t)$  is identified with an exponential decay. In those cases, Weibull distributions can be used to model survival time, and then:

$$g(t) = \begin{cases} \frac{k}{\lambda} \left(\frac{t-\gamma}{\lambda}\right)^{k-1} e^{-\left(\frac{t-\gamma}{\lambda}\right)^k} & t > \gamma \\ 0 & t < \gamma \end{cases} \quad (2.4)$$

where  $k > 0$ ,  $\lambda > 0$  and  $\gamma > 0$  are the shape, scale and location parameters of the Weibull distribution [20]. Thus, the cumulative distribution function (c.d.f.)  $G(t) = P\{T < t\}$  gives the probability of the event to have occurred by duration  $t$ . The *survival function*  $S(t) = 1 - G(t) = P\{T \geq t\}$  is then defined as the complement of the c.d.f, and describes the probability to remain in a given state before  $t$ :

$$S(t) = e^{-\left(\frac{t-\gamma}{\lambda}\right)^k}. \quad (2.5)$$

By inverting equation (2.5), it is possible to obtain directly the duration for which an appliance will continue (survive) in a specific energy state  $s$  as:

$$t_s = \gamma + \lambda [-\ln(w)]^{1/k}, \quad (2.6)$$

given a number  $w \in [0, 1)$  drawn randomly from a uniform distribution.

Fitting survival times to Weibull distributions has been successfully deployed in the past to model the times that appliances are in a particular energy state (on, off, stand-by or other). An illustrative example of the shape of the survival times in our data is presented in Figure 2.1, for a group of fractional energy states.

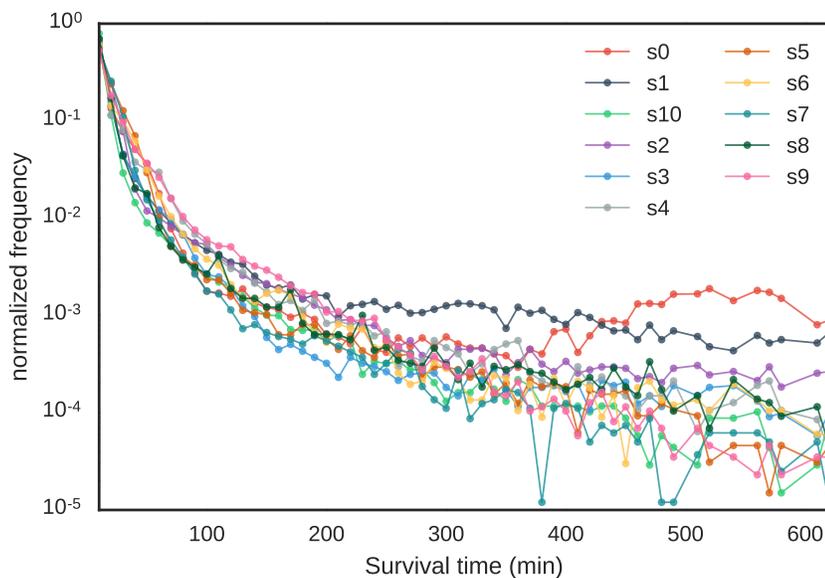


Figure 2.1: Exponential decay of survival times for eleven different fractional energy states. Frequency is normalised for the distribution.

Occurrences of each event and durations are first extracted from the data, and used to fit Weibull distributions, obtaining scale, shape and location parameters  $\lambda$ ,  $k$  and  $\gamma$ . The fitting process makes use of the `scipy` optimization package in Python, and fits the data to a Weibull function using a non-linear least squares method. Interpolation between points was necessary to fit the curves more effectively, given that the resolution of the original data (10 minutes) was not high enough. Two examples of the fitting process are given in Figure 2.2.

These distributions are then used to calculate survival times in a simulation

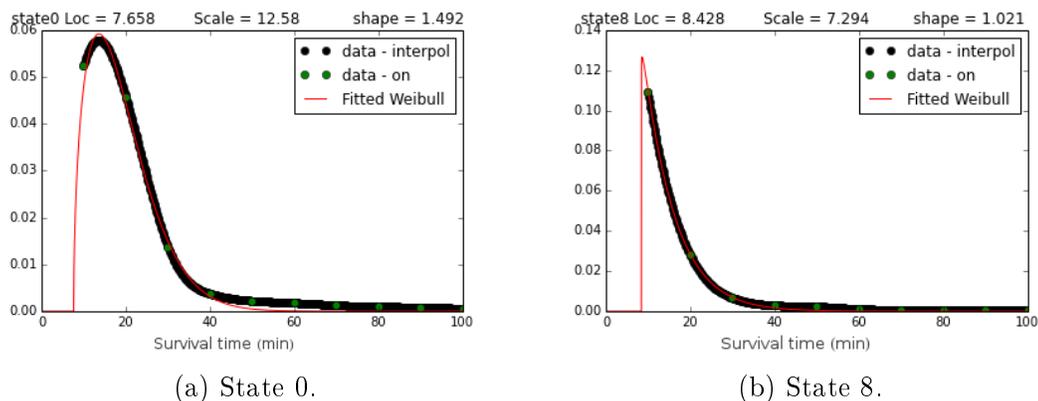


Figure 2.2: Details of data fitted to Weibull distribution using non-linear least square method.

using equation (2.6).

#### 2.4.4 Monte Carlo simulation of fractional energy states

Monte Carlo methods may be defined as *the representation of a mathematical system by a sampling procedure which satisfies the same probability laws* [15]. They provide a method to artificially represent a stochastic process by a sampling procedure, which will be determined by the particular underlying probability distribution of the given process.

For the specific problem posed here, the probability structure is given by either the parameters of the Markov chain or the survival analysis. In both cases, the purposes of the Monte Carlo simulation is to produce a time series of energy states  $s_k(t)$  for a given category of appliances. The process to obtain this sequence depends on which of the methods is being used.

### Simulation of energy states using a Markov chain

The sequence of energy states  $s_k(t) = \{s_{t=0}, s_{t=1}, \dots, s_{t=n}\}$  is simulated for a category of appliances  $k$ , employing an inverse function method, based on the probabilities given by the matrix  $P_{ij}$ . A given state  $s(t = 0)$  is assumed to be at the start of the simulation. From that state, the transition matrix  $P_{ij}$  gives the probabilities to make a transition to the next state. A random number is drawn from a continuous uniform distribution over the interval  $[0, 1)$  and the corresponding interval in the c.d.f. is selected as the next state. This process is repeated for each time step until the end of the simulation (see Figure 2.3 (a)).

This is the basic operation of all the Markov models employed in this thesis. However, as it was explained in Section 2.4.2, clustering techniques were used to produce more effective dimensioning of the  $P_{ij}$  matrix, resulting in 5 different Markov model implementations that will be deployed and explained in the next Chapter.

### Simulation of energy states using survival analysis

At the start of the simulation, state  $s(t = 0)$  is assumed. A random number  $w$  is drawn from a continuous uniform distribution over the interval  $[0, 1)$  and entered in equation (2.6) to obtain the survival time  $t_{s(t=0)}$  until a change of state occurs, covering a number  $n_{s(t=0)}$  time steps. When this time is over, the next state is calculated. In this case, transitions are not modelled. Instead, they were extracted using a sampling following the distribution of hourly probabilities  $P_s(t)$  of finding each of the states. Once at the next state  $s(t = t_{s(t=0)})$ , the process starts again: survival time is computed, until a change of state occurs. Thus, times and transitions between states are successively calculated for the simulation period.

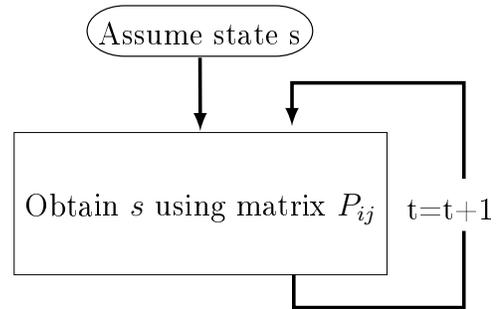
In this thesis, two survival models have been implemented, as outlined in figure 2.3, both following the operation described above, with different number of defined fractional energy states:

- **On/off survival model.** Two energy states,  $s_{OFF}$  for values of fractional energy  $f(t) = 0$  and  $s_{ON}$  for  $f(t) \in (0, 1]$ , corresponding to *on/off* states. Thus, switching on/off events are explicitly modelled.
- **Multistate survival model.** Eleven energy states, following an arbitrary division of 10 equidistant fractional energy states, plus the *off* state:  $s_i$  for  $f(t) = \{0; 0-0.1; 0.1-0.2; \dots; 0.9-1\}$ , respectively. Such a division of  $f(t)$  allows us to test the added value of refined characterisation of energy states. As it was said above, transitions between the different states are not modelled, but they are calculated using an inverse function method with the hourly likelihood of finding each of the states.

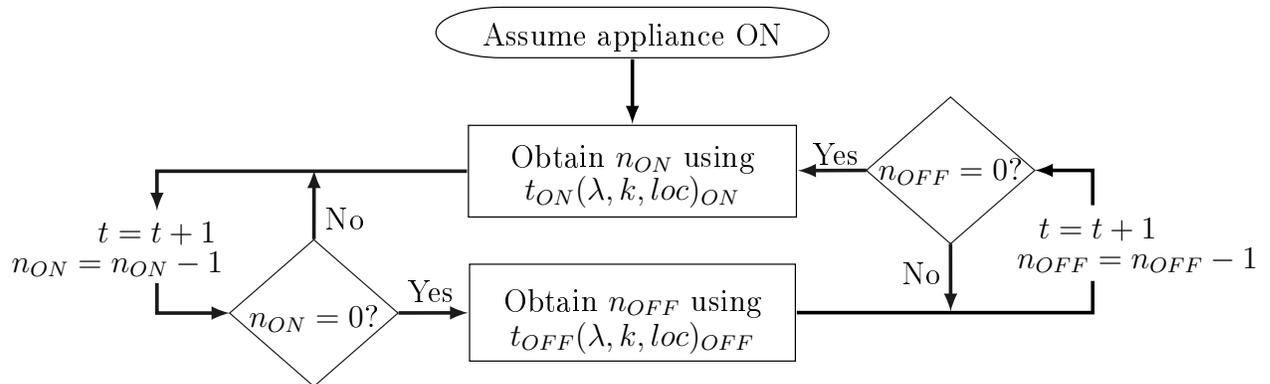
Both approaches, Markov and survival, present strengths and weaknesses. The main advantage of Markov chain models is that it explicitly models transitions between states, while the survival approach is not able to do so. On the other hand, Markov models are subject to the Markov property, by which the transition at the current state is independent of the previously visited states. In that sense, modelling the survival times is a more realistic approach. Another key advantage of the survival approach is that it does not require calculations for the all time steps while the devices remain in the same state, saving computation time.

### From energy states to an energy profile

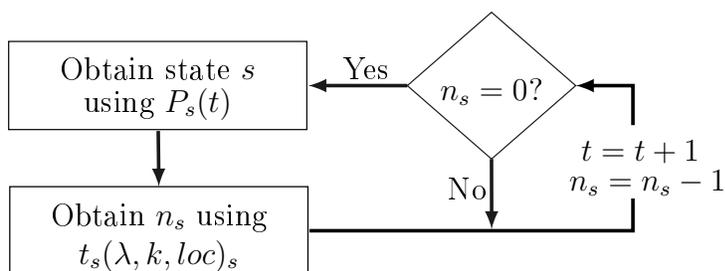
For category  $k$  the sequence of energy states  $s_k(t)$  now needs to be transformed back into a fractional energy profile  $f_{sim k}(t)$ . Thus, each energy-temporal state is



(a) Markov models.



(b) On/off survival model.



(c) Multistate survival model.

Figure 2.3: Simulation flowcharts.

multiplied by its corresponding mean or median fractional energy  $\tilde{F}_k(s)$ , depending on the strategy employed, leading to a simulated fractional energy profile

$$f_{sim k}(t) = s_k(t) \cdot \tilde{F}_k(s). \quad (2.7)$$

One final step transforms these profiles from fractional to actual energy values:

$$E_{sim k}(t) = f_{sim k}(t) \cdot \tilde{E}_{max k}, \quad (2.8)$$

where  $\tilde{E}_{max k}$  is a statistical measure of the maximum energy (or power  $\tilde{P}$  as required) for all instances of appliance in the category  $k$ . At this stage, depending on the measure selected for  $\tilde{E}_{max k}$ , the approach can be applied to simulate the use of individual devices belonging to a category  $k$ , or the use of the category as a whole. When considering the first case, assignment of  $\tilde{E}_{max k}$  is performed using the mean value:

$$\bar{E}_{max k} = \frac{1}{N_k} \sum_{i=1}^{N_k} E_{max k,i}, \quad (2.9)$$

where  $N_k$  denotes the number of instances  $i$  in category  $k$ . The second option consists of adding maximum energy values off all appliances in a category for each house, and perform a random sampling of them (values given in  $W$  in Table A.5). In any case, the estimation or selection of  $\tilde{E}_{max k}$  (or  $\bar{E}_{max k}$ ), becomes critical to calculating accurate aggregate energy profiles. If there is not such data available (as it was in our case), is essential to produce robust estimation. If the data is available, then the assignment of  $\tilde{E}_{max k}$  is trivial. In what follows then, we focus on testing the underlying hypothesis in our modelling strategies rather than in the fidelity of predictions of aggregate energy profiles that require a random assignment

process.

## 2.5 Validation methods

### 2.5.1 Cross validation

In statistical modelling, cross-validation processes are used to assess how effectively the results will generalize to a different data set [21]. Cross-validation computes the average error obtained from evaluation measures of different partitions of the data set. There are several methods for cross-validation, such as random sub-sampling, leave-one-out cross validation and  $K$ -fold cross validation. In our work we favour  $K$ -fold cross validation in which the data set is partitioned into  $K$  sub-samples. A single sub-sample is used as the validation set and the remaining  $(K-1)$  sub-samples are used as the training set. This process is then iteratively repeated  $K$  times (*folds*), until each partition has been used once as a validation set. A mean performance error can then be computed as the average error:

$$e = \frac{1}{K} \sum_{i=1}^K e_i, \quad (2.10)$$

where  $e_i$  represents some error between prediction  $\hat{y}_i$  and observation  $y_i$ .  $K$ -fold cross validation is a computationally expensive method, but produces an accurate estimation of the goodness of fit. The data set was split into 10 consecutive sub-samples of the time series. Therefore, each partition corresponds to a different time of the year, which could potentially cause seasonality issues on the analysis. In our study, we did not find evidence of such effects for the appliances we were considering. Cross-validation is usually more reliable when the partitions are per-

formed using random sampling. However, random sampling would break the time serial dependence of the data.

### 2.5.2 Time series analysis

Selecting an adequate strategy for the modelling of fractional energy requires a comparison of performance between simulation and observation data sets during the validation period. Time series analysis provides a powerful method to compare and understand internal structure on both temporal profiles, extracting meaningful statistical information. The objective is to describe the validation time series with a set of parameters that should be replicated by the simulation time series. In particular, it is possible to decompose the fractional energy profile into trend, seasonal and irregular (or remainder) component, allowing for evaluations of each of the components at a different level. Figure 2.4 shows an example of a decomposed time series.

The following information is used from the decomposition exercise:

- *Trend component.* There is no strong evidence supporting that there is trend in the observed data, or that this component changes over the year. Therefore, a constant value over the whole year has been assumed. This component is thus used as an average of the fractional energy over the simulation period.
- *Seasonal component.* A daily variation (or seasonal component) is expected in the use of appliances. The models are expected to reproduce this variation correctly, and this can be studied using the cross-correlation function [22] between two signals  $(X_t, Y_t)$ , which is defined as

$$\rho_{XY}(\tau) = \frac{1}{N-1} \frac{\sum_{t=1}^N (X_t - \mu_X)(Y_{t+\tau} - \mu_Y)}{\sigma_X \sigma_Y}, \quad (2.11)$$

where  $\mu_k, \sigma_k$  are the mean and standard deviation of process  $k = X, Y$ , respectively, and  $\tau$  is the lag or time delay between both. Equation 2.11 provides an insight into the relationship and dependence between observed and simulated periodic components. Based on that, we examine:

- Pearson’s coefficient, as an index of the linear correlation at  $\tau = 0$  (considering both signals to be synchronised); ideally this should be equal to 1.
  - Time delay of maximum correlation, in order to determine whether the signals are in phase with each other.
- *Irregular component.* After extracting the trend and seasonal components, a residual fluctuating variation remains.

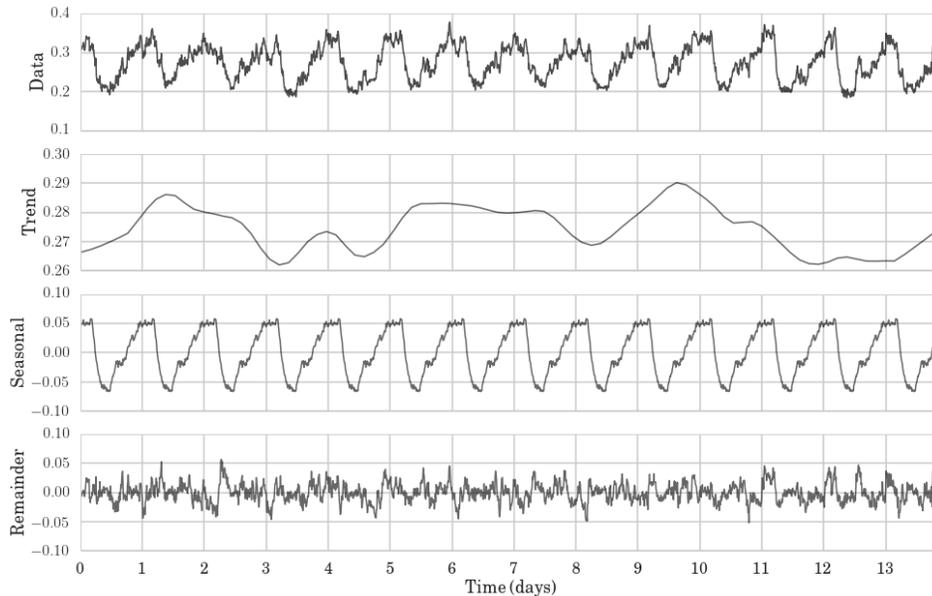


Figure 2.4: Example of time series decomposition into trend, daily variation and remainder component.

### 2.5.3 Sensitivity and specificity analysis

Sensitivity and specificity analysis represents a strong indicator of the model's absolute aggregate performance: its ability to correctly reproduce the time dependent properties of the process being simulated. Sensitivity or true positive rate (TPR) is defined as the proportion of matching cases between simulated and observed values, i.e.:

$$TPR = \frac{TP}{TP + FN}, \quad (2.12)$$

where  $T, F, P, N$  represent True, False, Positive and Negative and TP is the total number of truly predicted positive outcomes (true positives). Specificity or true negative rate (TNR) is defined as [23]:

$$TNR = \frac{TN}{TN + FP}. \quad (2.13)$$

In an ideal case, one would have  $TPR = 1$  and  $TNR = 1$  (or  $FPR = 1 - TNR = 0$ ). Comparison of these indicators can be plotted in receiver-operating characteristic (ROC) space. This analysis can be complemented with the model accuracy

$$ACC = \frac{TP + TN}{P + N}, \quad (2.14)$$

giving an indication on the overall performance of the model. In this thesis, sensitivity and specificity analysis relates to the modelling of multiple fractional energy states. For multi-state systems this is a particularly exigent evaluation technique.

## 2.5.4 Application of validation methods

In this work, 10-fold cross validation is performed for every approach suggested. For each iteration, several error measures at three different levels have been taken into account:

1. At the first level, we are interested in evaluating the quality of the fractional energy signal produced by our simulations. *Time series decomposition* has been performed (section 2.5.2) to extract the following comparative measures:
  - Relative error of the average energy usage over the period of simulation (using the trend component).
  - Pearson's coefficient and time delay of maximum correlation of the daily variation.
2. We are also interested in evaluating the accuracy of the averaged daily profile of fractional energy, as well as the models' effectiveness in predicting energy states. For this we use:
  - *Simulated energy states*. Sensitivity and specificity analysis is directly applied to the simulated energy states, producing ACC values and an ROC plot.
  - *Absolute state prediction*. The probability of predicting each of the states during the validation period is calculated and compared for observation and simulation. Discrepancies between both magnitudes are represented with RMSE.
  - *Temporal probability of state prediction*. The probability distribution for each state over time provides insight into the temporal variation

of each state, allowing us to identify situations when some states are over or under-predicted, and even at which periods during the day. Discrepancies again are calculated with RMSE.

- *Fractional energy daytime profile.* Once the energy states have been converted into fractional energy values, it is possible to evaluate the results for a typical day over the validation period (averaged over all the days for which fractional energy values are available for each time step). Residuals and RMSE are calculated to describe performance.

3. Finally, the selected methodology should perform well in calculating *total energy use*. Each simulated instance is converted to an energy profile using maximum energy values of the appliances present. The total energy use over the validation period is then obtained and compared for the relevant category of appliance.

The validation data set is a subset of the data that corresponds to 10% of the available total time range. This subset does not contain a unique time series of values, but a number  $N_k$  equal to the number of instances in the category  $k$ . The simulation of energy states was performed  $N_k$  times over the validation period, in order to perform the sensitivity and specificity analysis for the energy states. For the other evaluation measures, averaged values for all instances were considered for both observation and simulation.

## 2.6 Summary

Stochastic methods provide a simple and powerful technique to model dynamic electricity demand. In this chapter, we have introduced two types of them: discrete-

time and continuous time stochastic processes, which will be deployed and tested on real data on the following chapter.

A review of the relevant literature was given in Section 2.2. Stochastic methods, specially Markov process had been already successfully used in the past. However, the majority of the work was focused on determining the use of dominant appliances and electrical devices in homes, most of the times deriving its use from related activities. Although such task is important, in this research we go a step further and apply the aforementioned methods to categories of low-load appliances. They are less commonly owned and used, but its power demands are still relevant. We believe that having access to more detailed models of electricity consumption is valuable for building simulation.

Also, most of the current work does not consider power fluctuations when the device is being used. Considering fractional power as the stochastic variable, allows us to model variations in the total energy use. Section 2.4 described in detail the methodology of these processes. Markov processes were refined using a state-of-the-art clustering method that, to the best of the authors' knowledge, has never been applied before for reduction of the Markov matrix dimensioning.

A series of validation methods to evaluate the goodness of fit of the models are explained in 2.5. They pretend to cover different measures of the results. First, their temporal dependency using time series analysis. Second, their average daily trends, with sensitivity and specificity analysis, comparison of absolute state prediction, comparison of temporal probability of state prediction and comparison of the fractional energy daytime profile. Third, the total energy use is also evaluated.

In this chapter, a series of mathematical methods have been presented, necessary to carry out Objectives I.1 and I.2 of this thesis. The next chapter is a direct

continuation of this one, to demonstrate these objectives, showing the application of the methods here described to a real dataset: the Household Electricity Survey dataset.



## Chapter 3

# Small Appliance Modelling

The mathematical methods that we deploy in the modelling and validation of alternative appliance modelling strategies were described in the previous chapter. In this chapter, we first present the data used in our modelling: the Household Electricity Survey; and justify the choice of our four categories in which we model aggregates of appliances. Second, seven different variations of stochastic model are discussed and evaluated. From those seven, one is found to outperform the others. Deployment and validation of this strategy is presented at the end of the chapter.

### 3.1 Household Electricity Survey data set

The Household Electricity Survey [1] is an extensive monitoring survey of 250 households in the UK, carried out during 2010 and 2011. Apart from detailed socio-demographic information, it contains data describing the appliances present in every monitored household and their temporal electrical energy use during 1 or 2 months, with records every 2 minutes. Of the 250 households, 26 were additionally monitored for a whole year, with a 10 minute resolution. Since the one-month data was not measured during the same month for all households, only the data recorded for the 26 houses during a whole year was utilised in the analysis presented here,

in order to avoid possible seasonal effects on the use of appliances.

The relevant low-load appliances found in the dataset were classified into four categories, following the types of activity that relate to their use:

- audio-visual (excluding TVs, that are considered as high-load appliances given their extensive use),
- computing,
- small kitchen appliances (excluding cookers, microwaves and ovens),
- other small appliances.

In this thesis, subscript  $k$  denotes a category, and its values correspond to the four categories:  $k = \{1: \text{audio-visual}; 2: \text{computing}; 3: \text{kitchen}; 4: \text{other}\}$ .

Figure 3.1 shows the types of device available in the data set and their contribution to annual energy use, with categories depicted in different colours. The height of the bars represents the mean value of annual energy use of the corresponding type of appliance, whereas the width is proportional to the number of instances observed in the 26 households for the given device. Thus, the area of the bar indicates the total energy use of that appliance throughout the stock of houses surveyed.

One shortfall encountered in the data set is that there is no information describing the rated power of the appliances, posing a challenge to the accurate estimation of  $E_{maxk,i}$ . Consequences derived from this and the solution proposed are discussed in section 3.1.2.

The procedure adopted in our work was to test a range of strategies to model one appliance category, the audio-visual category, in order to identify the most parsimonious approach, and then to deploy this to other categories of appliance.

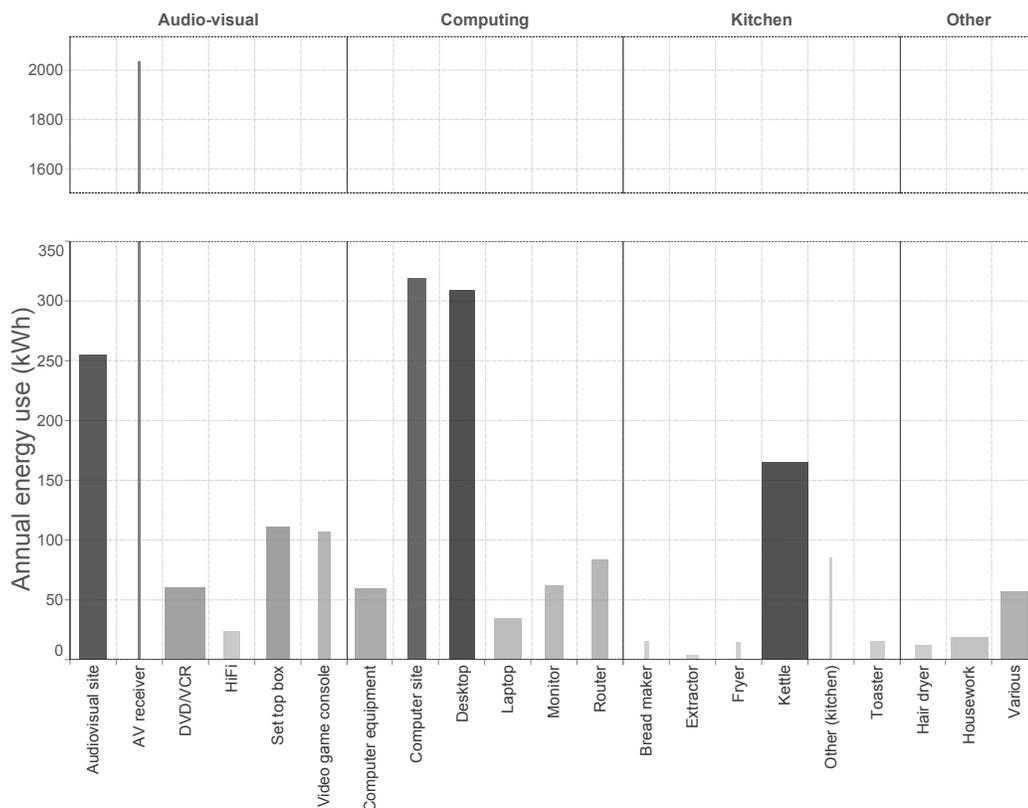


Figure 3.1: Annual energy use of the types of appliances considered in the modelling, divided in four categories: audio-visual, computing, kitchen and other. The height of the bars corresponds to the mean annual energy use, while the width is proportional to the number of instances recorded in the 26 houses. Combining this information, darker bars identify the dominant types of appliance for the category.

### 3.1.1 Audio-visual category

In this section the nature of the data used to test the modelling techniques is presented. With respect to the notation in Section 2.4.4 of the previous Chapter, we introduce here the subscript  $j$  referring to the type of appliance in a category  $k$ . Subscript  $i$  denotes now the instances of an appliance of type  $j$ . Ultimately, all appliances in a category  $k$  are modelled using the same parameters (independently of  $j$ ), but in this and the following sections it is necessary to distinguish between

| Type of appliance ( $j$ ) | Number instances ( $N_j^{k=1}$ ) |
|---------------------------|----------------------------------|
| AV receiver               | 1                                |
| Audio-visual site         | 33                               |
| DVD/VCR                   | 28                               |
| HiFi                      | 14                               |
| Set top box               | 17                               |
| Video-game console        | 9                                |

Table 3.1: Types of appliance available in the audio-visual category ( $k = 1$ ), and the number of instances (devices recorded in the data set through all the houses).

types, for the sake of the explanation. Table 3.1 displays the total number of instances of each subcategory of appliance considered in the audio-visual category,  $N_j^{k=1}$ , present in the 26 houses, leading to a total of 102 instances for the category. Our first step was to extract fractional energy values from the electricity use records, as

$$f_{k,j,i}(t) = \frac{E_{k,j,i}(t)}{E_{max\ k,j,i}}, \quad (3.1)$$

where  $k$  is the category,  $j$  the type of appliance and  $i$  the instance (one device in a specific house).

Given an estimate of  $E_{max\ k,j,i}$  this transformation outputs a normalised profile for each appliance in each house with values in the interval  $f_{k,j,i}(t) \in [0, 1]$ , that can be combined now with other instances or other types of device, allowing the category to be modelled. It is also possible to explore how these profiles vary during the period of a day, in order to identify patterns or dominant behaviours. Interesting characteristics of the data set include that:

- i. The data set contains over 4.2 million data points.
- ii. 33.9% of the data are zero values ( $f_{k,j,i}(t) = 0$ ), suggesting that appliances

are off for around a third of the time.

- iii. The off-state exhibits temporal dependency, reaching maximum values in the early hours of the morning (39%) when most people are sleeping, and a minimum (29%) between 20h and 22h, when most people are present and awake.
- iv. 15% of the entries have fractional energy lower than 0.1, which likely corresponds to a stand-by state, a common feature of audio-visual devices.
- v. As with [iii.](#), a concentration of stand-by states is found during the early hours of the day, whereas appliances are most often used at maximum power during the late hours of the evening.

A preliminary visualization of the two-dimensional space created by the time period of a day and the fractional energy values  $\{t, f\}$ , is depicted in figure [3.2](#). Dark areas represent denser regions of the data set, showing common values recorded during certain times of the day. Values of  $f_{k,j,i}(t) = 0$  were excluded to help with the interpretation.

### 3.1.2 Data preprocessing: outliers and maximum energy estimation

As previously mentioned, our data set does not include appliance name plate (power) ratings. The fractional energy modelling approach, however, is dependent on the values of  $E_{max\ k,j,i}$  and requires this input at two specific stages. Firstly in using equation [\(3.1\)](#) to extract fractional energy profiles for each instance. Secondly after the simulations have been performed, to compute an energy profile from a simulated fractional energy time-series for the category.

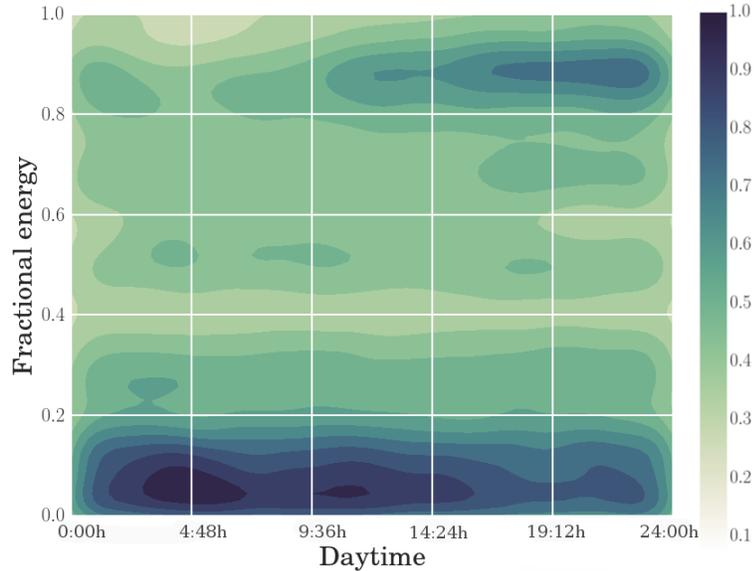


Figure 3.2: Distribution of audio-visual category data, for the values of fractional energy over a day (normalised values). A sample of 50000 entries is plotted. Since the data set contains a large amount of values with  $f_{k,j,i}(t) = 0$ , these data were excluded to facilitate comparison between other values.

In order to estimate  $E_{max k,j,i}$  from the data, the maximum energy record for each profile was used. The existence of discrepant entries for the same type of appliance suggested that a data cleaning process was necessary. In many cases, excessively high spikes were found, which corresponded to erroneous entries. Although in some cases this could be due to the fact that each data point represents the energy corresponding to the mean power drawn by an appliance over a period of ten minutes. Since this may fluctuate between 0 and the nameplate rating it could be that the selected value of  $E_{max k,j,i}$  results from an appliance that has been working at higher power during a shorter period of time (*e.g.* a kettle that never takes 10 minutes to boil). This problem was overcome by obtaining maximum energy values from the 2-minute data (also subjected to a cleaning pre-process),

with the purpose of selecting consistent entries. However, the data needed a cleaning process (filtering) to reject false spikes (outliers) which could lead to erroneous selection of  $E_{max\ k,j,i}$ .

A Seasonal Hybrid Extreme Studentized Deviate test (S-H-ESD) [24] was employed to detect anomalies. S-H-ESD is based on the generalized ESD algorithm to detect one or more outliers in a univariate data set that follows an approximately normal distribution, and is applicable to time-series data. Its main feature is that it is able to predict both local and global anomalies, taking into account long-term trends on the temporal profile to minimize the number of false positives. In other words, the conditions to detect an outlier vary depending on local temporal windows. When no trend is identified, the algorithm works as an ordinary outlier filter. The algorithm is part of the *AnomalyDetection* package in R [25].

This outlier filter is applied to all the time series, corresponding to each instance  $i$ , of each type of appliance  $j$  of each category  $k$ , in order to obtain a better estimation of their individual maximum energy  $E_{max\ k,j,i}$ , detecting anomalous spikes in the measured data. Individual maximum energy values are then averaged to calculate maximum energy for the type of appliance  $\bar{E}_{max\ k,j}$ . Table 3.2 presents these values before and after applying the filter, for the audio-visual category,  $k = 1$ . For other categories, the value of  $\bar{E}_{max\ k,j}$  after the exclusion of outliers can be reduced up to a 300% with respect to the value before applying the filter.

## 3.2 Results and discussion

In this section we first explain the application of the techniques presented in 2.4.1 and 2.4.3, respectively, to the data set introduced in section 3.1. Then, simulation results are described and evaluated for each of the strategies tested to model

| <b>Type of appliance</b> | $\bar{E}_{j,before}^{k=1}$ (Wh) | $\bar{E}_{j,after}^{k=1}$ (Wh) |
|--------------------------|---------------------------------|--------------------------------|
| AV receiver              | 86.4                            | 86.4                           |
| Audio-visual site        | 42.4                            | 34.9                           |
| DVD/VCR                  | 8.2                             | 5.41                           |
| HiFi                     | 33.1                            | 29.4                           |
| Set top box              | 5.2                             | 5.2                            |
| Video-game console       | 19.9                            | 16.7                           |

Table 3.2: Outlier filtering for appliances in the audio-visual category ( $k = 1$ ): mean maximum energy for the type of appliance, before ( $\bar{E}_{j,before}^{k=1}$ ) and after ( $\bar{E}_{j,after}^{k=1}$ ) applying outlier filters.

fractional energy use of audio-visual appliances, justifying the selection of one of them. Finally, the selected strategy is applied to the other appliance categories, and a final evaluation of the model is given.

### 3.2.1 Application of Markov model

As a first approach, the  $\{t, f\}$  space was arbitrarily divided with  $m = 11$  (11 energy states: one for the off-state plus ten of 0.1 fractional energy width) and  $r = 24$  (one temporal state per hour), leading to 264 subdivisions.

Clustering techniques were then applied to the audio-visual appliances data set. Excluding entries when the appliances are switched off (i.e.  $f_{k,j,i}(t) = 0.0$ ), there are over  $2.8 \cdot 10^6$  data points (from a total of over  $4 \cdot 10^6$ ), which is still large given the computational expense of the clustering algorithms used. In order to overcome this problem, a random sampling process [26] was carried out, selecting 50,000 points that roughly represent 2% of the total size of the data set.

Implementations of the DBSCAN and OPTICS algorithms were tested, corroborating that the unique global density parameter of DBSCAN was not effective at finding a satisfactory partition of the data set into clusters; since we are interested in finding clusters of different density.

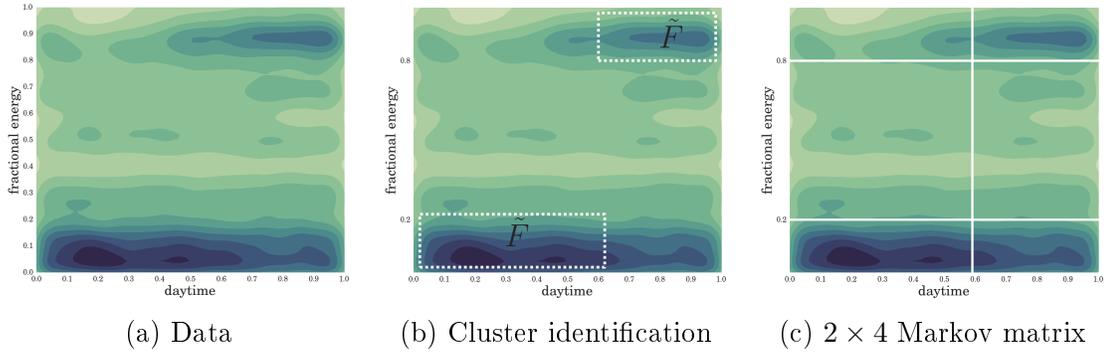


Figure 3.3: From the data: (a) clusters are identified by the algorithm allowing rectangles ranging from the 1st to 99th percentiles to be extracted (b); the rest of the space will be divided into noise cells. From that, a grid is defined, (c) whose partitions will be the dimension of the matrix (taking into account the off state  $f_{k,j,i}(t) = 0$ ), associating with each cell the median value of fractional energy  $\tilde{F}$  of the data points it contains.

**Subdivision of  $\{t, f\}$  space** The objective of applying a density-based clustering algorithm (OPTICS) is to produce an efficient subdivision of the two-dimensional space  $\{t, f\}$ , as described in section 2.4.1. As summarised graphically in 3.3, the process works as follows:

1. Find parameters that produce a good clustering structure.
2. Adjust the clusters found to fit a cell of rectangular shape. In order to avoid overlapping of cells, data points between the 1st and 99th percentile are selected for each cell. Points identified by OPTICS as noise are grouped into noise cells that will fill the empty space not covered by the clusters.
3. Define the grid established by the edges of the rectangles.
4. Associate with each cell a fractional energy value  $\tilde{F}$  corresponding to the median value of the points it contains.

OPTICS requires two parameters to produce the ordering of the points: first, the generating distance  $\epsilon_0$ , referring to the largest distance considered for clustering

(clusters will be able to be extracted for all  $\epsilon_i$  such that  $0 < \epsilon_i < \epsilon_0$ ); second, the minimum number of points that will define a cluster *MinPts*. However, the algorithm used to automatically extract the clusters from the ordering of the points and their reachability distance makes use of a further 7 parameters [19], upon which the clustering structure obtained will vary. For this work, the OPTICS algorithm was implemented in Python<sup>1</sup>.

After a systematic search for a well performing solution, a set of successful parameters was identified<sup>2</sup>. These values lead to a hierarchical solution, with four incremental nested partitions, from two clusters at the top level of the hierarchy, to eleven at the leaves. The four different levels (summarized in Table 3.3) lead to four different divisions of the  $\{t, f\}$  space and four Markov matrices with different dimensions. The relative performance of these different structures is evaluated in the following sections.

| Name          | Hierarchy level | Clusters | Noise cells | $\{t, f\}$ dimension |
|---------------|-----------------|----------|-------------|----------------------|
| OPTICS - 5x15 | IV              | 11       | 9           | 5x12                 |
| OPTICS - 4x14 | III             | 10       | 5           | 4x12                 |
| OPTICS - 3x11 | II              | 7        | 4           | 3x11                 |
| OPTICS - 1x3  | I               | 2        | None        | 1x3                  |

Table 3.3: Hierarchical levels of clustering considered for  $\{t, f\}$  space partition, with number of clusters and number of noise cells identified.

### 3.2.2 Application of Survival analysis

The two alternatives considered for the survival models are a simple two-state (on-off) model and a multistate model with 0.1 divisions in  $f(t)$ , so that there are

<sup>1</sup>Aided by script provided in <https://github.com/amyxzhang/OPTICS-Automatic-Clustering.git>

<sup>2</sup>Parameters found following description in [19] are:  $\epsilon = 0.08$ ;  $MinPts = 50$ ;  $minClustSizeRatio = 0.03$ ;  $minMaximaRatio = 0.001$ ;  $significantMin = 0.003$ ;  $checkRatio = 0.8$ ;  $maximaRatio = 0.87$ ;  $rejectionRatio = 0.8$  and  $similarityThreshold = 0.6$ .

eleven states in total (*cf.* 2.4.3). This multistate model encapsulates temporal variations since the transitions to the following state are computed based on the temporal probability of finding each of the states. Weibull parameters (shape, scale and location) introduced in equation (2.4) are estimated from the data points for each of the energy states. Once obtained, the simulation runs as depicted in figure 2.3.

### 3.2.3 Approach selection

The fractional energy use of the audio-visual category of appliances was modelled using a range of strategies. The goodness of fit of the models is evaluated from different points of view, following the description in section 2.5.4.

#### Fractional energy time series

Decomposition of the time series over the validation period allows for the extraction of statistical information from the structure of the observed and simulated data and to compare their different components: trend, daily variation and remainder (see section 2.5.2).

**Trend** The trend component has been assumed to be constant over the whole year of observed data, so there is no need to fit a function. It gives an estimation of the average value of the fractional energy over the simulation period, given that the daily variation has been removed. Values are shown in table 3.4.

**Daily variation** The daily variation components are shown in figure 3.4, plotted for several days. There are two models, OPTICS-1x3 and Survival, for which an inadequate handling of dynamics is clearly apparent. For the other cases, those

with larger numbers of temporal states produce an understandably more accurate profile of the daily variation (11x24-SHESD, OPTICS-5x14, with 24 and 5 temporal states, respectively). Also, the Survival Multistate model represents surprisingly well the daily variation, considering that the temporal dependency is included only in the transitions between states, but not in their duration.

Table 3.4 complements those results with numerical values for Pearson's coefficient and temporal lag at maximum correlation. Again, the best value for Pearson's coefficient and time lag is achieved using the models with a larger number of temporal states 11x24-SHESD and Survival Multistate, followed by OPTICS-4x12 and OPTICS-5x14.

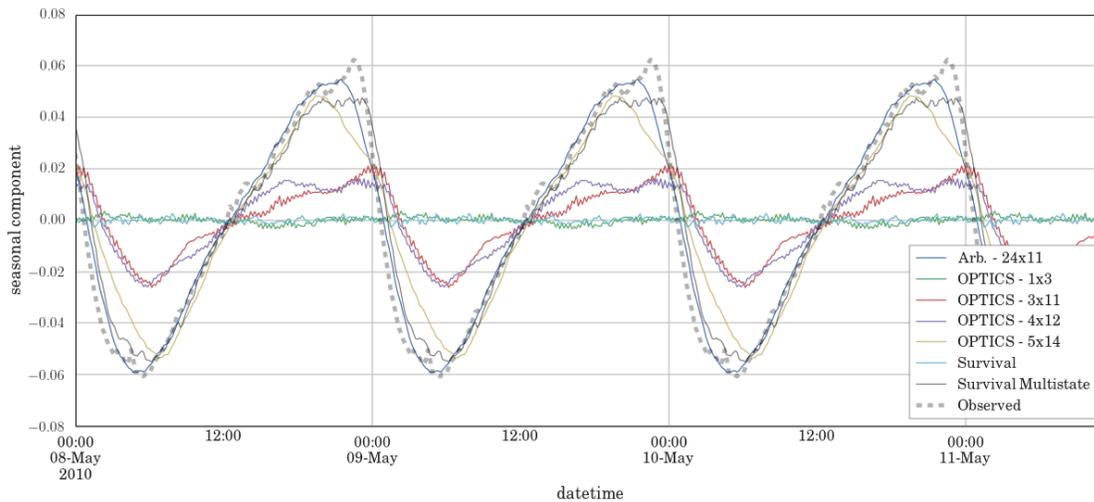


Figure 3.4: Comparison of daily variation components.

### Evaluation of average daily profile

In the previous section the signal simulated over the whole period was compared; but we are also concerned with how well the averaged daily profile is represented, in terms of the predictive power of simulated energy states and the consequent fractional energy profile.

|                     | Rel. error trend | Pearson's coeff. | lag     |
|---------------------|------------------|------------------|---------|
| Arb - 24x11         | 2.35%            | 0.989            | 0.0     |
| OPTICS - 5x14       | 5.08%            | 0.9501           | -0.0278 |
| OPTICS - 4x12       | 4.91%            | 0.960            | -0.0486 |
| OPTICS - 3x11       | 5.40%            | 0.854            | -0.0625 |
| OPTICS - 1x3        | 3.83%            | -0.446           | -0.424  |
| Survival            | 0.831%           | -0.0896          | 0.236   |
| Survival multistate | 5.61%            | 0.983            | -0.0208 |

Table 3.4: Summary of validation time series decomposition. From left to right: error on the average value of the trend; Pearson's correlation coefficient and time delay (lag) of daily variation. "Arb" refers to the arbitrary subdivision of the data into 24 time states and 11 fractional states.

**Fractional energy states prediction** Figure 3.5 shows the dependency of the RMSE (calculated for every 10-minute timeslot) with time for the probability of finding the system in each of the defined energy states. The *on/off* Survival approach gives RMSE values an order of magnitude larger than for the Markov models, indicating a poor overall estimation of the two states considered. Since there are only two states defined, their probabilities of being simulated are complementary,  $P_{s=0}(t) = 1 - P_{s=1}(t)$ ; therefore, a poor estimation of  $P_{s=0}(t)$  implies a poor estimation of  $P_{s=1}(t)$ .

Furthermore, the shape of the curves for Survival and OPTICS-1x3 models implies that the temporal dependency of the system is not well encapsulated. The former exhibits an increase in error during the late hours, suggesting a worse prediction of the on-state; while the RMSE in the latter increases both in the evening and during the night, revealing an under performance for both the off state and the maximum energy state.

Temporal dependency is well represented in the Survival Multistate model, although the overall error in energy state prediction is higher than with the Markov

approaches. This could suggest that the specific energy states are better represented when clustered energy values have been considered.

The total RMSE for temporal and average daily state predictions are presented in table 3.6. In both cases, OPTICS-5x14 outperforms the other strategies, suggesting that the larger the number of energy states (14 in this case), the more accurate the probability prediction.

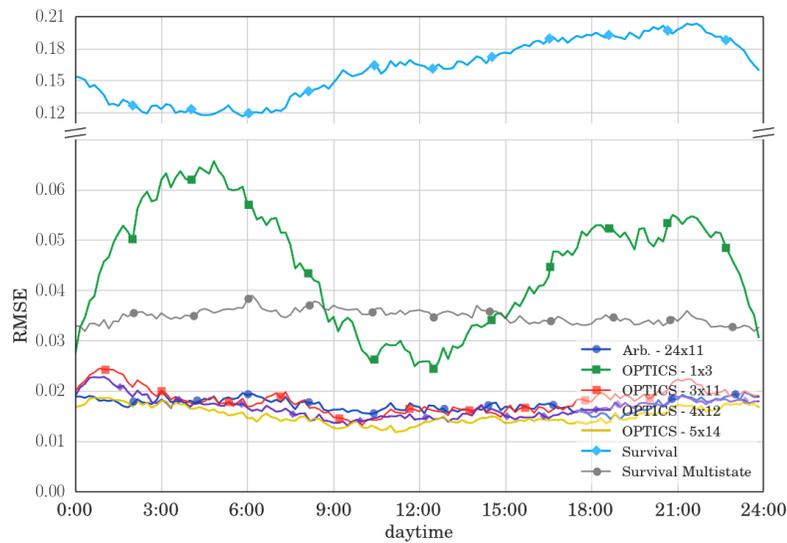


Figure 3.5: Temporal dependency of RMSE, calculated for each 10-minute time slot, for all states of each approach.

As noted in section 2.5 the accuracy of the modelling of states can also be evaluated using ROC parameters, as shown in table 3.5; although this is a particularly onerous test when applied to multi-state systems, so that TPR is not expected to be high. Once again the OPTICS 5x14 and Survival Multistate models outperform their counterparts.

|                            | <b>TPR</b> | <b>TNR</b> | <b>ACC</b> |
|----------------------------|------------|------------|------------|
| <b>Arb - 24x11</b>         | 0.173      | 0.917      | 0.483      |
| <b>OPTICS - 5x14</b>       | 0.160      | 0.935      | 0.487      |
| <b>OPTICS - 4x12</b>       | 0.172      | 0.925      | 0.485      |
| <b>OPTICS - 3x11</b>       | 0.178      | 0.918      | 0.483      |
| <b>OPTICS - 1x3</b>        | 0.339      | 0.669      | 0.456      |
| <b>Survival</b>            | 0.499      | 0.499      | 0.450      |
| <b>Survival multistate</b> | 0.189      | 0.919      | 0.484      |

Table 3.5: Accuracy of model.

|                     | <b>Fractional energy <math>f(t)</math></b> | <b>Absolute state prediction</b> | <b>Temporal state prediction</b> |
|---------------------|--|----------------------------------|----------------------------------|
| Arb - 24x11         | $2.056 \cdot 10^{-2}$                      | $2.00 \cdot 10^{-2}$             | $1.76 \cdot 10^{-2}$             |
| OPTICS - 5x14       | $2.32 \cdot 10^{-2}$                       | $1.86 \cdot 10^{-2}$             | $1.56 \cdot 10^{-2}$             |
| OPTICS - 4x12       | $2.54 \cdot 10^{-2}$                       | $2.06 \cdot 10^{-2}$             | $1.73 \cdot 10^{-2}$             |
| OPTICS - 3x11       | $2.70 \cdot 10^{-2}$                       | $2.23 \cdot 10^{-2}$             | $1.89 \cdot 10^{-2}$             |
| OPTICS - 1x3        | $4.66 \cdot 10^{-2}$                       | $1.84 \cdot 10^{-2}$             | $4.85 \cdot 10^{-2}$             |
| Survival            | $4.22 \cdot 10^{-2}$                       | $1.13 \cdot 10^{-1}$             | $1.63 \cdot 10^{-1}$             |
| Survival multistate | $2.53 \cdot 10^{-2}$                       | $3.55 \cdot 10^{-2}$             | $3.50 \cdot 10^{-2}$             |

Table 3.6: RMSE values of the daily profile results, in terms of the fractional energy profile, absolute state prediction (without temporal dependency), and temporal state prediction.

**Fractional energy averaged daily profile** The residuals in fractional energy for an average day tend to increase towards the boundaries of the day (Figure 3.6), where users are more active in switching devices and regulating them. Nevertheless, the results suggest that even with temporally crude models, dynamics are well encapsulated (with the exception of OPTICS-1x3 and Survival); particularly in the case of the model with the largest number of temporal states, Arb.-24x11, as reflected in table 3.6.

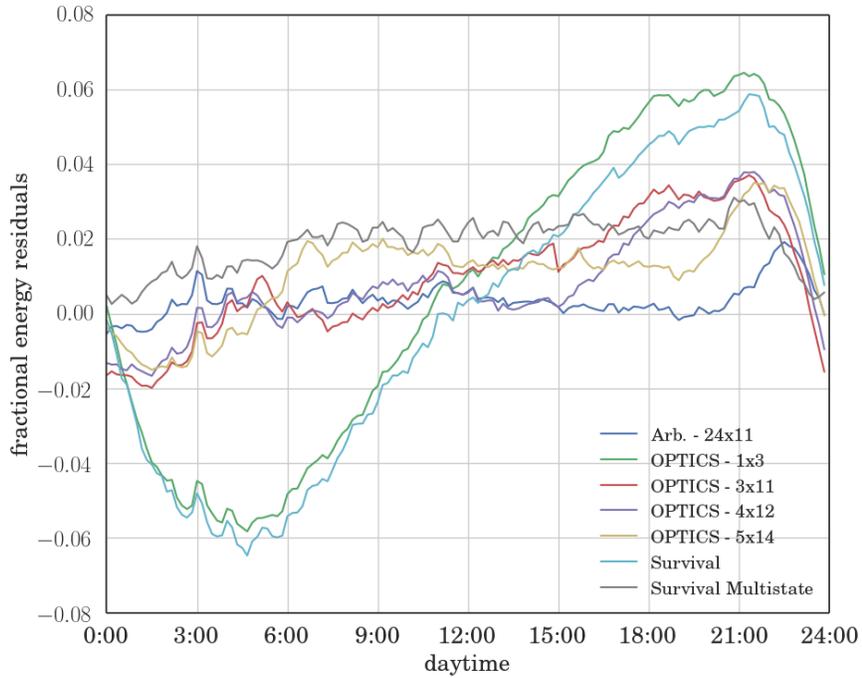


Figure 3.6: Daily profile residuals of fractional energy.

### Total energy prediction

For the total energy prediction over the validation period, a sampling process of maximum energy values is performed following the procedure described in section 2.4.4.

In order to compare the results, a box plot is presented in figure 3.7, and the residual error in energy use prediction is presented in table 3.7. Whilst the median residual error is in all cases relatively low, the simulated values are consistently positively skewed, overestimating the upper quartile in total energy use. This is caused by a loss of information during the modelling process. Errors compound from the modelling of fractional states, through the assignment of maximum energy values to the subsequent prediction of energy use for the relevant appliance category. Also, appliances in the same category may still have different behaviours, which this modelling technique cannot distinguish. The mean energy overestimation could

|                     | <b>Mean Residual<br/>(kWh)</b> | <b>Median Residual<br/>(kWh)</b> |
|---------------------|--------------------------------|----------------------------------|
| Arb - 24x11         | -13.5                          | -1.83                            |
| OPTICS - 5x14       | -13.5                          | -1.76                            |
| OPTICS - 4x12       | -12.7                          | -1.43                            |
| OPTICS - 3x11       | -12.5                          | -1.48                            |
| OPTICS - 1x3        | -11.8                          | -1.68                            |
| Survival            | +1.13                          | +1.80                            |
| Survival multistate | -11.6                          | -1.44                            |

Table 3.7: Residual error between observation and simulation, for mean and median of the total energy use over the validation period.

come from the fact that some appliances may have high power ratings, but be used in seldom occasions; on the contrary, those devices actively use may have low power ratings. Our approach is not able to capture that. Thus, even though each task in our modelling process faithfully reproduces reality, errors inevitably arise when using models estimated from aggregate data of the four typologies of appliance to the prediction of specific device behaviours; errors that will reduce in magnitude as the size of the stock of appliances simulated increases. This is reasonable considering that our goal is to estimate communities of buildings and the appliances contained within them.

### Summary

The complexity of the different approaches can also be used for comparison, based on the type and number of parameters that the models need. They are summarised in table 3.8. For the Markov based approaches, the parameters needed are those that build the Markov matrix, and are dependent on its dimension. Additionally, the clustering process requires 9 extra parameter values, which are not easily extracted, as the clustering algorithm requires a trial and error process which is

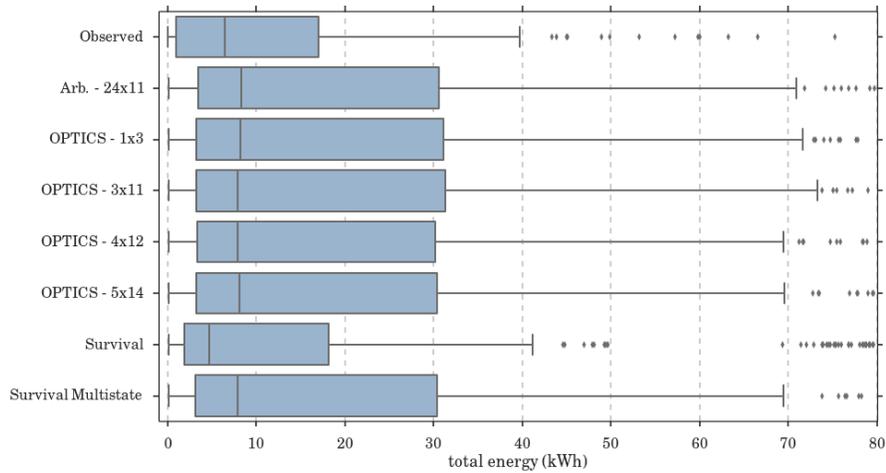


Figure 3.7: Boxplot comparing observed and predicted total energy use over the validation period.

|                     | <b>Number of<br/>parameters</b> | <b>Type</b>                     |
|---------------------|---------------------------------|---------------------------------|
| Arb - 24x11         | $24 \times 11 \times 11$        | Markov matrix                   |
| OPTICS - 5x14       | $5 \times 14 \times 11 + 9$     | Markov matrix and clustering    |
| OPTICS - 4x12       | $4 \times 12 \times 11 + 9$     | Markov matrix and clustering    |
| OPTICS - 3x11       | $3 \times 11 \times 11 + 9$     | Markov matrix and clustering    |
| OPTICS - 1x3        | $1 \times 3 \times 11 + 9$      | Markov matrix and clustering    |
| Survival            | $3 \times 2$                    | Weibull parameters              |
| Survival multistate | $3 \times 11 + 24 \times 11$    | Weibull and states' probability |

Table 3.8: Number and type of parameters needed for the different type of model.

complicated and time-consuming. The parameters needed in the survival models are those that describe the Weibull distribution, plus the hourly probability distribution of each state to occur (trivial to obtain and which can be simplified using less time slots.)

To inform our selection of the most parsimonious modelling strategy the relative performance of each of the models tested is qualitatively summarised in figure 3.8, using a color coded diagram. From this it is apparent that the Survival Multistate strategy outperforms its counterparts: its predictive power is comparable to that of

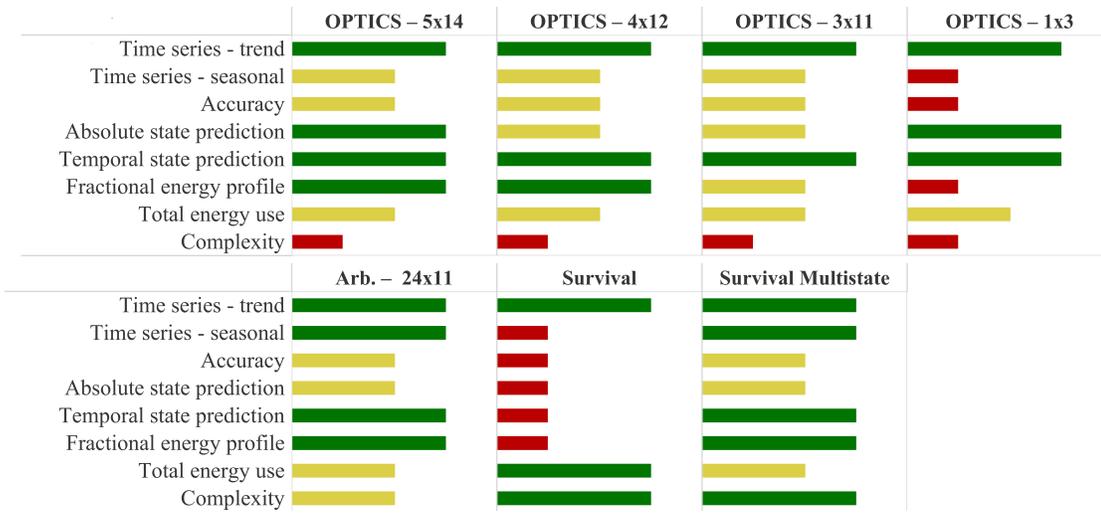


Figure 3.8: Summary of validation results between the approaches tested, qualitatively represented as *good* (long green bar), *average* (yellow medium bar) and *poor* (short red bar).

the more refined Markov models, but is considerably simpler in formulation, both in the estimation of its coefficients and in subsequent implementation. It performs well in the time series analysis, temporal state prediction and fractional energy profile, acceptably well in absolute state prediction, accuracy and total energy use. For these reasons, the Survival Multistate approach has been deployed to model the other categories.

### 3.2.4 Application of Survival Multistate approach to other categories

In this section results for the simulation of the energy use of computing devices ( $k = 2$ ), small kitchen appliances ( $k = 3$ ) and a category of other appliances ( $k = 3$ ) is presented, following the Survival Multistate approach.

### Discussion on modelling a diversity of appliances

Modelling categories of appliances prevents from the analysis of different behaviours from specific devices, which are in a large range of total time of use (between commonly-used and seldom-used appliances) and peak demand values (low-rated and high-rated appliances). At the extremes of this range two types of behaviours have been identified: dominant appliances (commonly-used and high-rated) and infrequent appliances (very rarely used over the course of a year, independently of their rated power). In both cases, these behaviours are undetected by our modelling approach, with corresponding implications for predictive accuracy.

In the case of the kitchen category, preliminary results as described in 3.1 led to the elimination of the kettle as part of the category. As a high-rated appliance that is commonly owned and used, its behaviour is dominant, misleading the extraction of parameters of the model. Figure 3.9 shows the results for the survival multistate model applied to the kitchen category with and without the kettle. In this particular case, the total energy use was underestimated by the model, due to its inability to discriminate between the power use pattern of this particular appliance and the other small kitchen appliances<sup>3</sup> Once removed, the result shows a very good fit with the observed data.

---

<sup>3</sup>Note that the small appliance modelling strategy is predicated on the modelling of appliances with potentially complex dynamic behaviours, expressed in variations in fractional power demand whilst in use. The kettle is a considerably simpler case, having a constant power demand but for a short duration depending on the temperature and quantity of water to be boiled. A detailed usage model of the kettle is presented in [27].

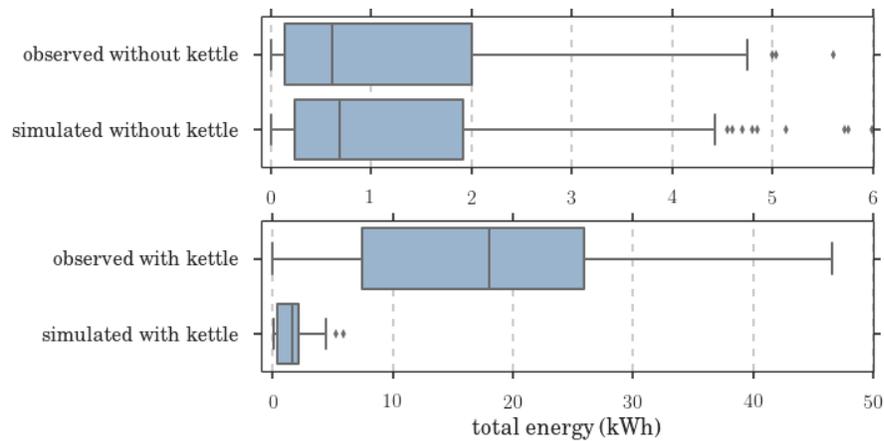


Figure 3.9: Effect of a dominant appliance (kettle) on the observed and simulated data for the kitchen category.

The category of *other* appliances, on the other hand, is biased by the effect of infrequent appliances, which were monitored in the survey but are very rarely used: several being used only for less than 1% of the total recorded time. Consequently, the total energy use predicted was overestimated.

### Performance of the model

The performance of the model has been evaluated in a similar fashion to that for the audio-visual category, and is summarized in table 3.9. In general, the model performs comparably to that of the modelling of audio-visual appliances. The results are remarkably good for the case of the kitchen devices, once the kettle was removed, proving that the strategy is very powerful for modelling relatively homogeneous type of appliances. Larger errors in energy prediction are found for the other two cases, related again to the diversity of behaviours present on the dataset, as explained in 3.2.3. Notwithstanding this, the average fractional energy use is well predicted, as are the states (table 3.9).

| Error measure               |                           | Computing            | Kitchen              | Other                |
|-----------------------------|---------------------------|----------------------|----------------------|----------------------|
| <b>Time series analysis</b> | Relative error trend      | 15.7%                | 0.429%               | 8,49%                |
|                             | Pearson's coefficient     | 0.967                | 0.946                | 0.927                |
|                             | Lag                       | -0.014               | -0.0005              | -0.014               |
| <b>ROC curve</b>            | TPR                       | 0.205                | 0.842                | 0.783                |
|                             | TNR                       | 0.920                | 0.984                | 0.978                |
|                             | ACC                       | 0.485                | 0.517                | 0.514                |
| <b>Daily profile (RMSE)</b> | Fractional energy ( $f$ ) | $4.44 \cdot 10^{-2}$ | $2.23 \cdot 10^{-3}$ | $1.27 \cdot 10^{-2}$ |
|                             | Absolute state prediction | $5.48 \cdot 10^{-2}$ | $6.16 \cdot 10^{-3}$ | $1.77 \cdot 10^{-2}$ |
|                             | Temporal state prediction | $4.08 \cdot 10^{-2}$ | $1.06 \cdot 10^{-2}$ | $1.70 \cdot 10^{-2}$ |
| <b>Total Energy (kWh)</b>   | Mean Residual             | -26.4                | -0.318               | 17.7                 |
|                             | Median Residual           | -19.0                | -0.160               | -20.6                |

Table 3.9: Summary of results on application of Survival Multistate approach to other categories for: time series analysis, sensitivity and specificity, RMSE of daily profile and total energy.

|                     | Appliance type     | Rated Power (W) |
|---------------------|--------------------|-----------------|
| <b>Audio-visual</b> | Set top box        | 30              |
|                     | Audio-visual site  | 43.2            |
|                     | DVD/VCR            | 12.6            |
| <b>Kitchen</b>      | Bread maker        | 97.2            |
|                     | Toaster            | 720             |
|                     | Extractor          | 16.8            |
| <b>Computing</b>    | Laptop             | 48.6            |
|                     | Computer equipment | 3.6             |
|                     | Desktop            | 108             |
|                     | Monitor            | 30              |
|                     | Router             | 7.2             |
| <b>Other</b>        | Housework          | 1248            |
|                     | Various            | 54.6            |

Table 3.10: Available appliances in house "103028" and their rated values, extracted from the data set.

### 3.2.5 Global performance and application of the model

The application of the model is shown in this section in two ways: the first involves a single day simulation for a specific household (labelled in the dataset as "103028"), presented in figure 3.10. It contains 13 different low-load appliances, which are described in table 3.10. Figure 3.10 displays the output of the model, for the four categories, when using the listed appliances. As expected, the model predicts usages of different duration and it is able to capture the spikes in the profiles. But, as expected, the model does not resolve for the specific characteristics of the individual appliances, and it does not represent different behaviours between them, as the parameters inside the model have been defined to describe aggregates of appliances.

A more suitable use of the model is presented in Figure 3.11, where the one-day simulation has been carried out for the categories as wholes, aggregating devices in Table 3.10. In this case, simulated and observed series are plotted for comparison. For the audio-visual and computing categories the simulated profile shows a more dynamic behaviour than the observed one, with more predicted on/off switches. The observed data, on the other hand, is showing one or more devices that remain switched on but with low variations during the day. On the contrary, the kitchen and other category predict well the sporadic use of these devices.

The third includes the averaged daily energy usage arising from all the devices in the different categories of appliances over the year, when aggregated to a community of 20 households (figure 3.12). This situation is much more representative of the intended usage of the model than for the modelling of individual appliances in a single household. In this case, the total energy use for each category (adding up all the available devices in each household) is averaged in order

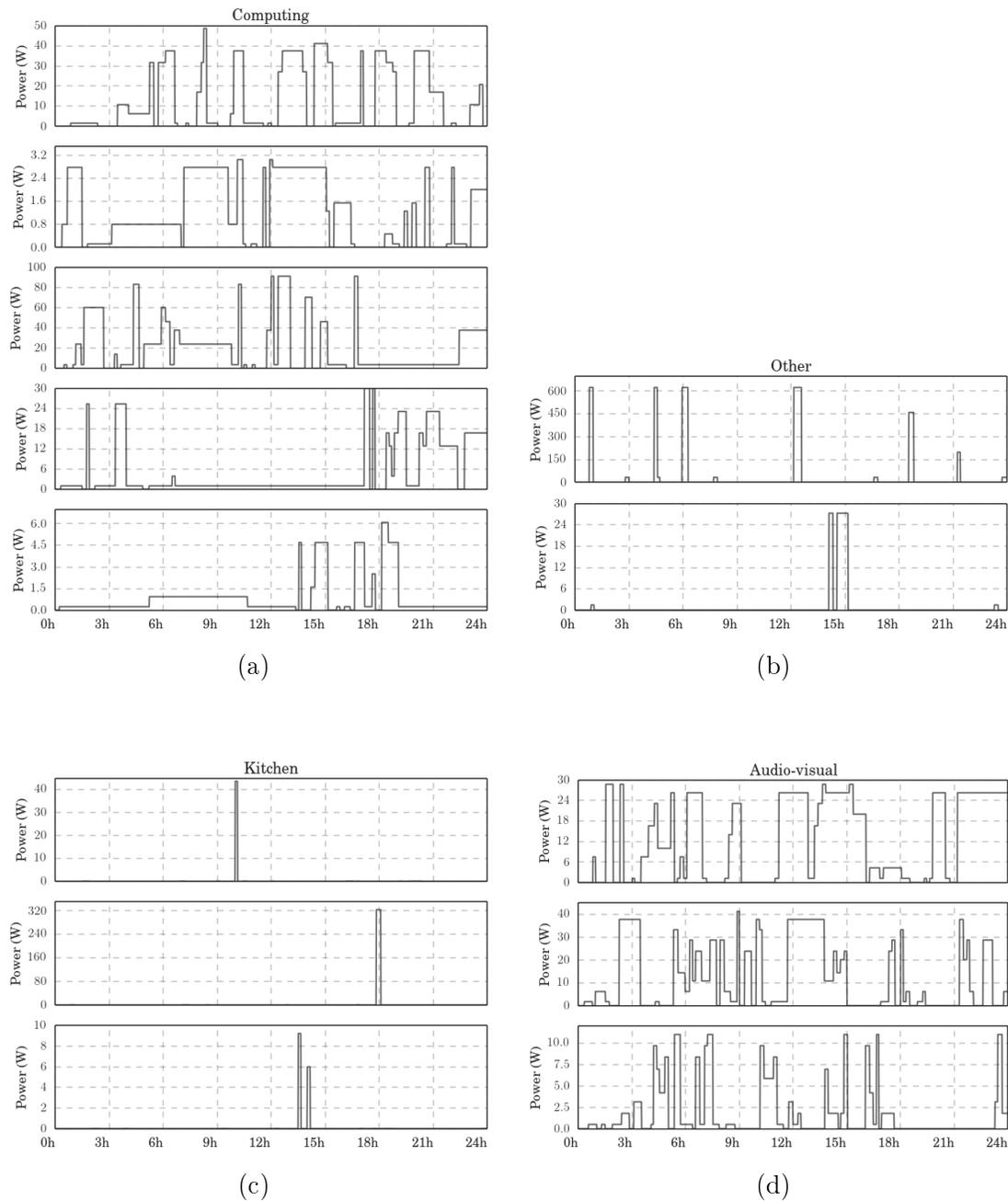


Figure 3.10: Example of one-day simulation for individual low-load appliances, by category: (a) computing, (b) other, (c) kitchen, (d) audio-visual. They correspond from upper to lower as indicated in table 3.10.

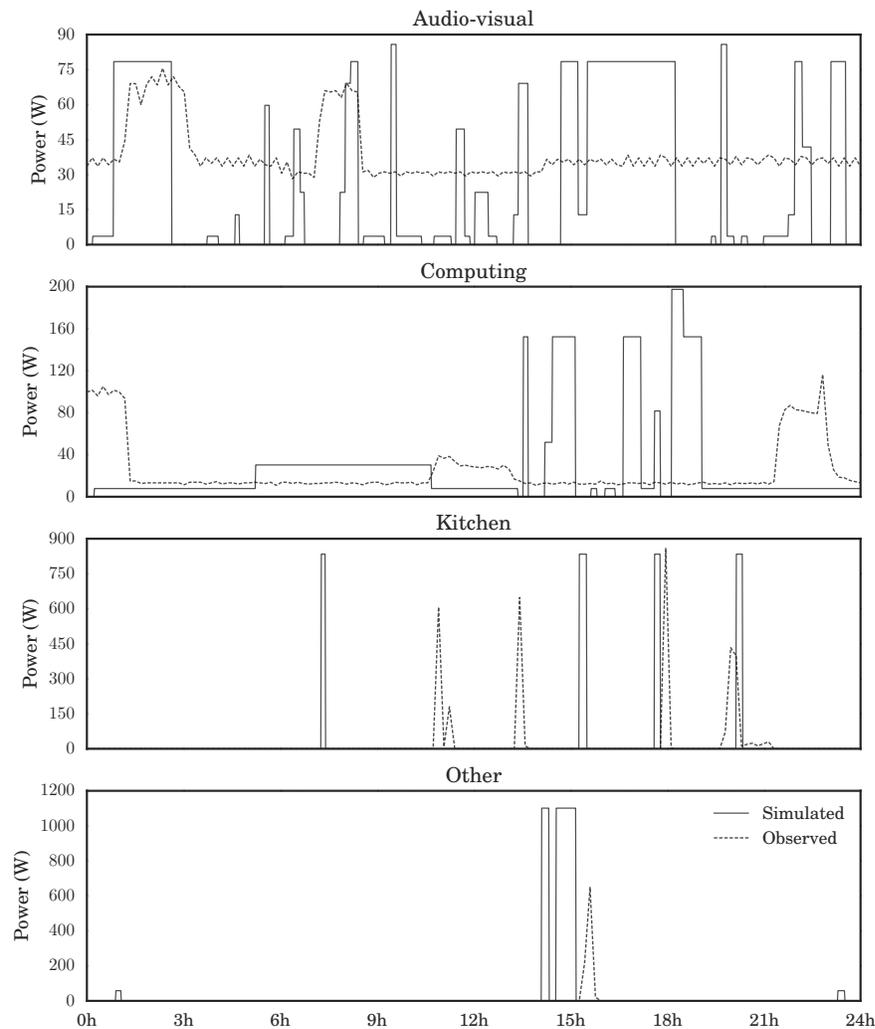


Figure 3.11: Example of one-day simulation for aggregates of appliances.

to create a typical day profile for the community. Then observed and simulated data are compared. Simulated audio-visual appliances describe temporal variability, although its dependency is not as strong as in reality. Again, the choice of category maximum energy values for appliances with different behaviours impacts the total energy use predicted by the model. This effect is more clear in the high values of energy use at night hours, and lower values than the observed during the evening peak. Something similar occurs for the computing appliances, although in

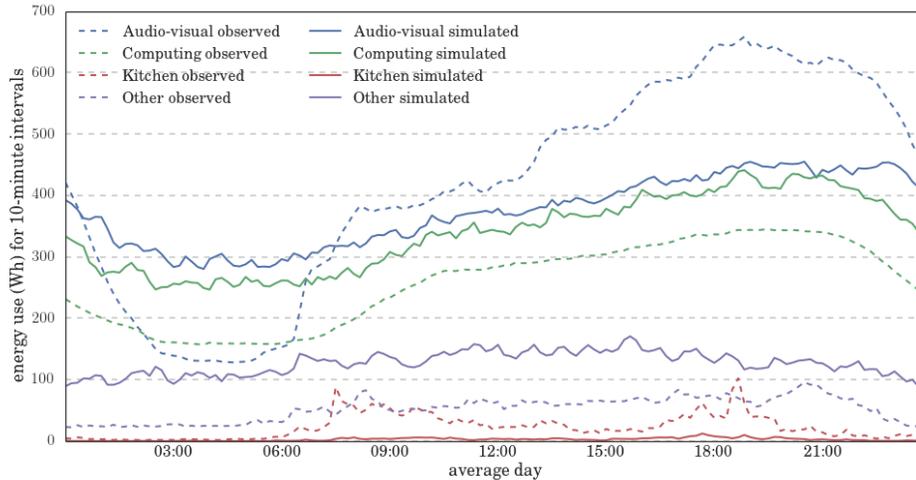


Figure 3.12: Averaged daily total energy usage from the different categories.

this case, its energy use is consistently overestimated, but still captures variations during the day. The use of kitchen devices is underestimated, and the opposite happens with the other category; this could be related to the amount and type of devices available for this particular group of households, given that their use is reduced. One way of improving these consistent over/underestimations would be to make a correction after the simulation to the total energy use or to the maximum energy value employed. To conclude, although we are modelling exclusively aggregates of appliance and in a generalised way, realistic magnitudes for the electrical energy use with respect to time can still be obtained. The parameters used in both cases are those detailed in Appendix A.

### 3.3 Conclusion

As the integrity of the envelope of both new and existing houses improves, so the proportion of energy that is used by household electrical appliances, which

are becoming increasingly ubiquitous, is likely to increase. It is important then that modellers have at their disposal reliable models of appliance energy use, if they are to accurately predict the thermal performance and energy use of future homes. Furthermore, there is increasing interest in the concept of smart grids, to better regulate the distributed supply, storage and demand of electrical energy. This places increasing onus on the ability to predict the dynamic behaviour of household electrical appliances. Whilst good progress has recently been made in the modelling of relatively large appliances: those whose prevalence and cumulative energy use supports the estimation of device-specific models. Poor progress has been made in the modelling of relatively small appliances: those whose cumulative energy use is individually small, but significant when considered as aggregates by typology. To this end we have tested a range of strategies for the modelling of small appliance categories; first predicting discrete states in fractional energy demand, then converting these into absolute energy demand, given an estimate of the corresponding maximum power demand.

In this we deploy (as described in Chapter 2) both discrete (Markov) and continuous (survival) time random processes; for the former also utilising cluster analysis to effectively partition the state transition probability space.

From this process of model developments, application and validation we draw the following conclusions:

- Modelling appliances by their typologies presents many advantages: it provides a straightforward solution for modelling the range of types of appliances, it reduces the amount of input data needed to estimate the model and the risk of overfitting, and it avoids the time-consuming process of modelling appliances individually, simplifying dynamic energy simulation.

- The model predicting time varying fractional power demands is surprisingly robust, given that it is modelling aggregates. In particular we find that:
  - Finer discretisation of temporal states improved predictive power, but these improvements are modest beyond 5 temporal states.
  - Appropriate estimation of the number of fractional energy states is not as influential as the number of temporal states.
- Clustering techniques have been effectively deployed to objectively search for a parsimonious form of model: minimising the size and partitioning of state transition probability matrices. The methods presented can be used for many other areas of research.
- Based on three types of evaluation measure (time series analysis, model accuracy and aggregated energy use), the survival multistate approach, in which survival times are estimated for selected bins of fractional energy demand, clearly outperforms its Markov process counterparts.
- However, analysing categories can compromise the fidelity of predictions of aggregate energy use, particularly if modelling small numbers of households. In our case, a successful strategy consisted of allocating maximum energy values with a random assignment process. The survival multistate approach has been effectively deployed to model low-load appliances in four categories: audio-visual, computing kitchen and other. The profiles output by the model have been satisfactorily compared to those of a community of households.

As previously noted, this work forms part of a larger programme of research to reliably predict appliance energy demand using bottom-up techniques for communities of households, and to test strategies for the management of these appliance

demands to improve community energy autonomy. The proposed Low Voltage network model and the testing and evaluation of these Demand Side Management strategies are reported in the following chapters.



## Chapter 4

# Generalisation of multi-agent stochastic simulation architecture to support Demand Response

There exist a multitude of different mathematical approaches to define the Demand Response problem, and maybe more ways to computationally implement the optimal operation of devices and resources. Having reviewed the pros and contras of candidate strategies, we have opted for a Multi Agent Simulation (MAS) approach, for the following reasons. First, the landscape of microgrids and community energy concepts is intrinsically represented as a system of multiple actors (*agents*) which are technologically ready to employ varying degrees of intelligence to pursue an objective (*make individual decisions*), and enabled communication with other actors in the network (*agent interaction*). Secondly, the architecture of No-MASS lends itself to the ready extension from modelling occupants as agents to devices as agents, with potential interactions between the two typologies. In this chapter, we present ideas for testing DR, such as load shifting and battery operation, in a MAS framework. Our general methodology, including Q-learning for system optimization, is explained here.

## 4.1 Introduction

The power system at the distribution level is evolving towards decentralised monitoring, supply and control. Small-scale renewable energy generation and storage devices are increasingly incorporated into the network, which in turn needs intelligent management and information communication systems to fulfill local power demands. Such an evolved distribution grid is part of a broader concept, the *Smart Grid*. There is no unified definition of Smart Grid in the literature<sup>1</sup>. Common concepts among the multiple definitions cover: bidirectional communication between supply and demand, use of monitoring technology, asset utilisation optimization, integration of renewable power sources and intelligent management of all the actors involved in the transmission and distribution systems. Thus, the Smart Grid is bringing new challenges at technical, regulatory and policy levels.

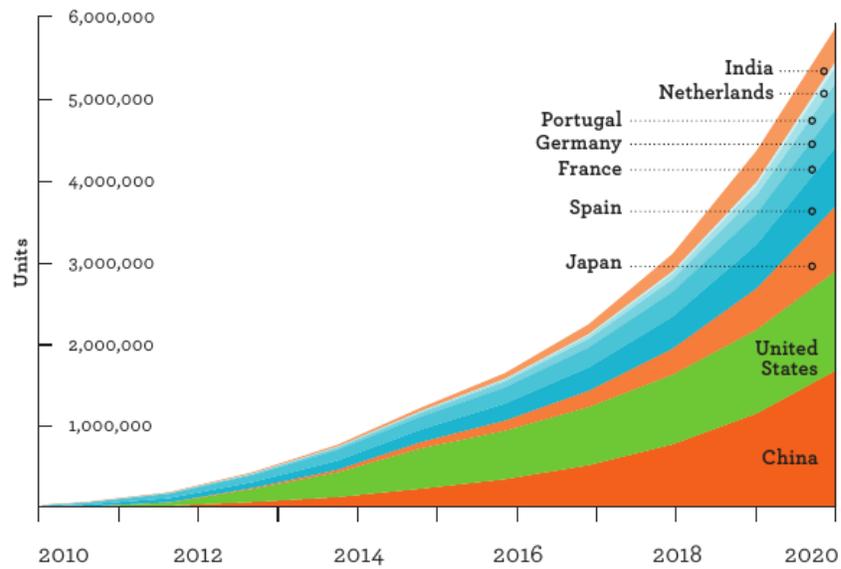
As a consequence of political regulation to reduce carbon emissions, such as the 2°C scenario (2DS)<sup>2</sup> or the 2008 UK Climate Change Act (80% reduction), the uptake of Low Carbon Technologies (LCT) is growing. Power sector decarbonisation, leading to the electrification of heat and transport through greater penetration of heat pumps and electric vehicles among other technologies, is predicted to have a great impact on power demands, which are expected to significantly increase in future years, despite of efficiency improvements. In 2015 over 1 million electric cars<sup>3</sup> were on the roads [30], with this figure expected to increase in the following years (figure 4.1a). The IEA Photovoltaic Energy Roadmap [31] envisions 4600GW of installed capacity by 2050 (figure 4.1b), with PV system prices reduced to a

---

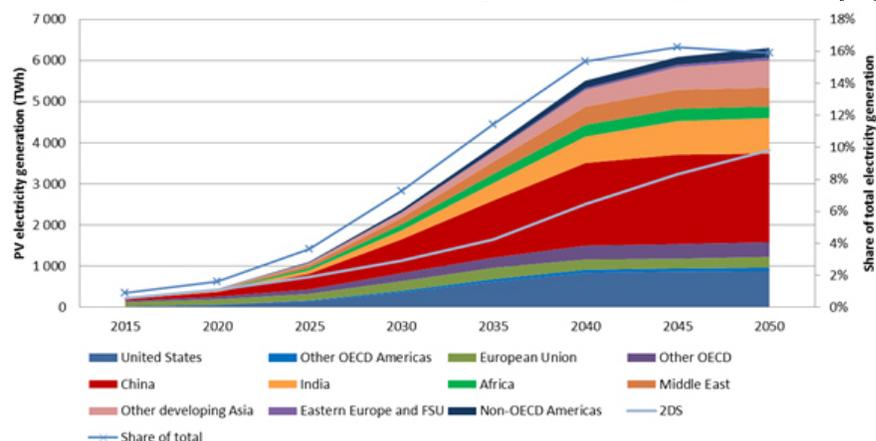
<sup>1</sup>A compendium of definitions from different institutions can be found in [28].

<sup>2</sup>The 2DS targets the energy-related CO<sub>2</sub> emissions to reduce by more than half in 2050 (compared to 2009) and to fall thereafter [29].

<sup>3</sup>Including battery electric vehicles, plug-in hybrid electric, and fuel-cell electric vehicles (BEVs, PHEVs and FCEVs).



(a) EV sales targets forecast worldwide. Note: A 20% annual growth rate is assumed for countries without specific sales target. Source: [33]



(b) Regional production of PV electricity envisioned. Source: [31]

Figure 4.1

third since 2008. The 2DS vision estimates 310GW of additional grid-connected electricity storage needed in Europe, China, India and United States (80GW approximately were available in 2011) [32]; but a more intense deployment would be achieved with a cost reduction breakthrough in the technologies.

In this landscape, Demand Side Management (DSM) and Demand Response

(DR) measures have been identified as key components of the proposed infrastructure [34, 35]. This is because they represent a core and affordable technology to better allocate resources, without the need of costly upgrades to the current power distribution system. It is important to clarify the difference between the two. DSM is understood as “the planning, implementation and monitoring of any activity designed to encourage consumers to modify patterns of energy usage, including the timing and level of electricity demand” [36]. DR refers to strategies and technologies that modify consumption patterns in response to signals (e.g. supply conditions or electricity price). In this sense, DSM is a broader concept that includes Demand Response and Energy Efficiency (EE).

Electricity flexibility in industry and the high volume/high capacity customer segment has been important in balancing and matching demand and supply for a long time, playing a key role in system reliability measures (dispatchable DR). Due to the capacity increase of Distributed Generation (DG) of Renewable Energy Sources (RES), demand flexibility is increasing its value in lower-volume/lower-capacity end-users.

At this level, the application of DSM and DR measures is not only trying to propose technical solutions, but it requires close interaction and involvement of stakeholders. At domestic level, previous studies [37] have consistently shown the relevance of behavioural interventions in the overall effectiveness of DSM programmes. However, the underlying determinants for energy-related actions is not clear yet, and there is a lack in consistency between different DSM programmes in the assessment of their effectiveness. A set of scales that can be used for evaluation of behavioural interventions is presented in [38], as part of the IEA DSM Task 24

collaboration<sup>4</sup>. It proposes a range of metrics, with the aim of standardize programme effectiveness evaluation<sup>5</sup>. The need for such standardization implies that experts cannot compare clearly the effectiveness of different programmes. Consequently, existing programmes could not be designed to target the engagement of specific sectors of the population, reducing the possibilities for achieving their maximum potential. There is great scope for simulation software to help bridge this gap, as it allows to test different strategies and scenarios using computer simulation.

In summary, DR is expected to have a major impact on domestic energy efficiency and energy-related behaviours in the coming years, but there are still so many open questions: what control strategies work best?, are tariff structures the right incentive for behavioural change engagement?, do they have different impact on different type of consumers? We are interested in support the design of successful DR programmes using computing simulation to model and optimise different strategies and scenarios. To carry out such task, we have identified a set of requirements that we acknowledge a DR simulation platform should satisfy; they are articulated in Section 4.2. Background information and major developments in the area are described in Section 4.3. Section 4.4 introduces the software tool developed and used: No-MASS. The methodology for DR simulation is detailed in sections 4.5 to 4.7. Finally the chapter concludes in 4.8.

---

<sup>4</sup><http://www.ieadsm.org/task/task-24-phase-1/> and <http://www.ieadsm.org/task/task-24-phase-2/>

<sup>5</sup>Based on norms (e.g. motivation to engage or energy literacy), practices (behaviours and intentions), material culture (appliance ownership), context (e.g. physical properties of dwelling or demographics) and user experience (ease of use, engagement, trust and satisfaction).

## 4.2 Formulation of DR software requirements

In our research, we have identified the need of simulation software that is capable of modelling and optimising Demand Response (DR) strategies. To this end, we can articulate some key requirements R that a DR simulation platform should aim to satisfy. It should be capable of:

R1. Simulating demands for four appliance archetypes:

- i)* switched on, regulated by and switched off by the user (e.g. cooker),
- ii)* switched on by the user and off when a programme is complete (e.g. washing machine),
- iii)* switched on and off according to some programme or schedule (e.g. hot water system),
- iv)* continuous cycling (e.g. refrigerator).

The user-interaction should be stochastic.

R2. Drawing power to satisfy demands from:

- i)* local generation capacity,
- ii)* local storage devices,
- iii)* the local microgrid and/or the national grid;

similarly of diverting locally generated power to either local demand and storage devices, or to the local/national grid.

R3. Deciding, to satisfy some objective function, where power should be drawn from/diverted to; rescheduling demands (or the provision of energy related

services), given some pre-defined constraints related to service delivery (e.g. a washing machine may be activated after having been enabled, but must complete the wash by a predefined time).

- R4. Presenting information to the user and emulating the users' decision making rationale regarding the rescheduling of user-controlled devices.
- R5. Accounting for diversity in the extent to which users' are willing to relinquish control and to actively engage in behavioural change.
- R6. Facilitating the above for communities consisting of buildings with numerous demand devices and potentially numerous supply and storage devices which can communicate within and between buildings to achieve individual homeowners' requirements; potentially also those of the local low voltage network to which they are connected (e.g. in terms of network stability and safety).

### 4.3 Literature Review

Distinction between DSM and DR was explained in Section 4.1, and it is depicted in Figure 4.2. A broad classification of DR programmes distinguish between event based (dispatchable) and non-event based (non-dispatchable) programmes. Dispatchable DR concerns responses to emergency reliability events and/or peak load reduction events. They are usually triggered by the appropriate system operator (after previous agreement with end-users). In most of the programmes in this category, the decision to activate a certain DR service relies on contingency events, and users are obligated to respond. In some other cases (such as demand-bidding), the users follow dispatch instructions from a third party in response to some form

of pricing signal. Alternatively, non-dispatchable methods allow the end-user to choose to activate DR, following some form of time-sensitive pricing.

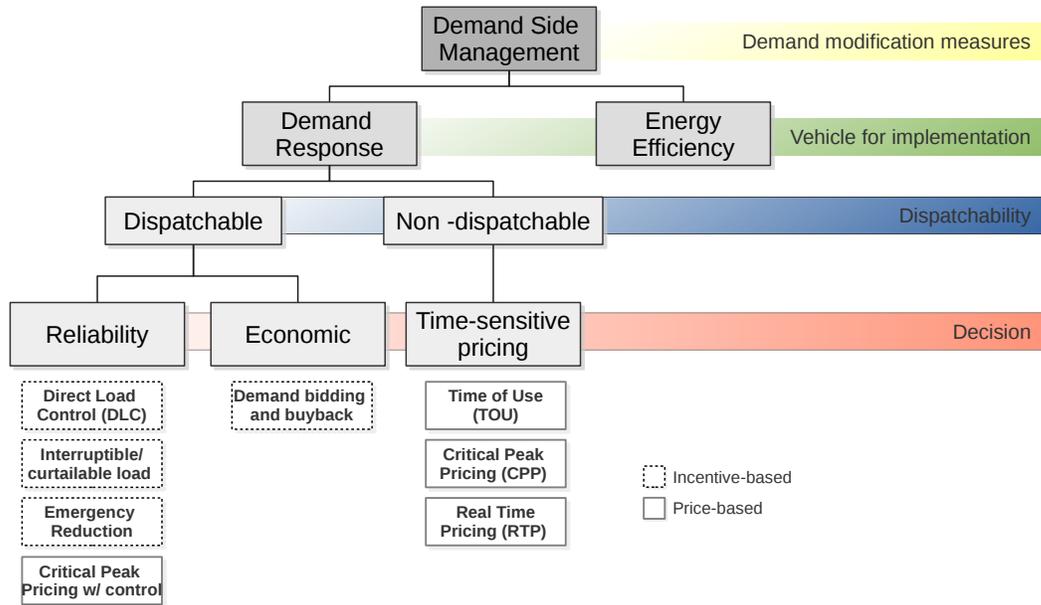


Figure 4.2: Demand Response categories. Source: [39].

Solutions for both dispatchable or non-dispatchable DR can be classified into price-based and incentive-based schemes [40, 41]. Price-based programmes concern the modification of the electricity profile in response to price variations. Different tariff structures and regulation contribute to load-shifting in peak times. Examples of those are Time Of Use<sup>6</sup> (TOU) tariffs, Critical Peak Price<sup>7</sup> (CPP) and Real Time Pricing<sup>8</sup> (RTP). Incentive-based programmes, on the other hand, are related to fixed or time-varying financial incentive plans. They have been commonly implemented at industrial end-user level for a long time now, however, they are less common in domestic trials. Some examples include direct load control<sup>9</sup>, interrupt-

<sup>6</sup>Demand rates for peak and off-peak hours. In the UK, it corresponds to *Economy7* tariff.

<sup>7</sup>Uses real-time prices for cases of extreme peak.

<sup>8</sup>Considers (half-)hourly prices in the same day or day ahead.

<sup>9</sup>The DR operator can remotely control user equipment.

ible/curtailable services<sup>10</sup>, demand bidding<sup>11</sup> or emergency programmes<sup>12</sup>.

### 4.3.1 DR methods and techniques

The problem of identifying optimal DR strategies can be formulated as an optimization problem, in which the objective function  $F(\mathbf{x})$  must be optimised under certain constraints. The formulation of the function  $F(\mathbf{x})$  can take different forms, such as the daily cost of the energy bill or the fraction of renewable energy use on a given period of time [40], constituting the DR objective. Depending on the case, the optimization will involve maximizing or minimizing  $F(\mathbf{x})$ . For complex cases, in which  $\mathbf{x}$  represents a set of parameters and functions, this optimization problem can be difficult to solve analytically. In this scenario, there are various algorithmic formalisms that enable the parameter space to be explored in the search for a global optimum.

The following section include more detailed information of the most commonly used methods for DR algorithms.

#### Optimization methods employed in DR

The mathematical formulation for the optimization algorithms will depend on the type of objective function  $F(\mathbf{x})$ , the parameters and functions considered in  $\mathbf{x}$ , and the specific DR mechanisms considered (such as load shifting or thermostat regulation). Consequently, there are various techniques available for DR optimization. They range from traditional calculus methods to heuristic optimization [40]. Some specific and widespread optimization techniques for DR are listed in Table 4.1,

---

<sup>10</sup>Curtailment options are integrated into the contracts with the users, who can be penalised if they fail to reduce demand during critical periods.

<sup>11</sup>Large customers bid to curtail demand, based on wholesale electricity market prices.

<sup>12</sup>Incentives are offered during reserve shortfalls.

with considerations and examples for each one. Each of the methods presented in the table is linked to a specific problem, and formulated under certain considerations. It is not the purpose of this table to rank the different techniques, but to provide an overview of those available, and how they are adapted to a given DR problem. If  $F(\mathbf{x})$  is linear, linear programming can be a good option; when  $F(\mathbf{x})$  is a convex function, convex optimization is available. Linearity and convexity, however, are restrictive properties, difficult to formulate in complex scenarios. Other approaches consider time-varying parameters, for which dynamic programming, stochastic programming and Markov decision methods provide better options.

| METHOD                      | CONSIDERATIONS   | EXAMPLES         |
|-----------------------------|--|------------------|
| Linear programming          | Including Integer/Mixed Linear/Non-linear programming. Challenging for complex scenarios.  | DemSi [42]       |
| Convex optimization         | Basic approach to DR. Challenging for complex scenarios.   | [43]             |
| Game theory                 | Useful to study interactive decision-making processes, with the potential to address interactions between different actors in the network. Restricted to rather simple constrained problems. | [44]             |
| Particle Swarm Optimization | Bio-inspired heuristic approach, to tackle challenging optimization problems. Other options consider neural networks, genetic algorithms.  | [45]             |
| Dynamic programming         | Basic approach to deal with time-varying parameters, such as power generation or price.  | [46]             |
| Stochastic programming      | Addresses time-varying parameters whose probability distributions are known. It is a special type of dynamic programming.  | [47]             |
| Markov Decision Problem     | Specially addresses time-varying parameters. Usually solved with a combination of dynamic programming and reinforcement learning.  | [48], [49], [50] |

Table 4.1: Examples of optimization problems and methods in DR.

**Centralised and decentralised DR** With respect to the implementation of such algorithms, two groups arise: centralised or decentralised programmes [40, 41], and they relate to which component(s) in the system is (are) allowed to react to stimuli. In the former, decisions on load re-scheduling or power dispatch are made by a central controller, informed by the operation of each individual element in the network. Optimization methods used in these cases range from traditional calculus methods [43] to heuristic optimization [51]. However, these approaches become challenging when large and complex networks are considered. In the case of linear programming or convex optimization problems, they attempt to provide with a definite optimal solution; when a large number of customers or devices are considered for DR, such problems become increasingly complex and heuristic or other methods capable of finding near-optimal solutions are necessary. When applied in real life situations, centralised systems sometimes may raise consumer privacy protection concerns, as they require a central authority to collect data and information for decision making.

On the other hand, decentralised DR schemes incorporate the ability of distributed decision-making, assuming a certain degree of intelligence of the devices involved (such as smart meters and appliances, power electronics and so on), which ensures direct communication between the elements in the network. End-users can directly access indicators of the state of the grid, or of other points of the network, and react based on them. Certain software architectures, such as object-oriented and agent-based programming, are naturally configured to deal with decentralised algorithms. And for such architectures, some optimization algorithms are more suitable; for instance, game theory methods (adequate for relatively constrained problems, not highly complex typologies), or Markov decision problems are easily

adaptable to distributed approaches (see Table 4.1.).

**DR and Multi-Agent Simulation** Multi-agent systems are inherently a successful way of designing distributed DR programmes, with the option of employing agent learning for optimization. Multi-agent simulation (MAS) systems have been proposed in different power engineering applications [52], from marketplace simulations to operation and control methods of the power system. They have been exploited in different ways such as monitoring and diagnostics, distributed control [53], fault protection and modelling and simulation. Particular interest is posed in multi-agent systems being used for Demand Response [51, 54, 55]. Being inherently decentralised, they present a highly flexible and extensible modelling approach for simulating Device-to-Device (D2D) communication between supply, storage and demand devices: systems are defined in terms of agents which fulfil individual and collective goals, being able to compete, collaborate, negotiate and learn behaviours.

Multi-agent systems can be found extensively in the literature to tackle DR modelling [46, 48, 49, 56]. The implementation of machine learning techniques allow for agent learning, meaning that different agents can learn to optimise their behaviours. Reinforcement learning is a popular choice for agent learning, and in particular Q-learning has been proved effective for DR applications, from explicit Vehicle-to-Grid control [48], to microgrid coordination [56]. A useful mathematical description of the Q-learning algorithm in a device-centric approach is presented in [46], and applied to a single household. Finally, the Q-learning algorithm is taken a step forward in [49], adding another "learning layer" to design a "distributed-W-learning" algorithm, in which agents learn how their local actions affect their neighbours' performance. Although the formulation is promising for microgrid

studies, it has only been put into practise for the case of maximising energy use of electric vehicles from wind-generated electricity.

### 4.3.2 Application into DR simulation software tools

The previous section reveals that there are multiple modelling strategies and algorithms available for DR simulation. Although many interesting candidate models and ideas are found in those studies, usually they are focused on developing certain aspects of DR, or testing specific algorithms. They do not have the potential to address all the requirements articulated in section 4.2. However, we are interested in developing flexible, scalable and extensible software for DR simulation. In that sense, this section provides an insight on those studies aiming to be DR simulation tools, which in our view, should target to address the range of these requirements in a comprehensive way.

DRSim [55] is a simulator for DR, which mainly focuses on R1, R4 and R5 requirements. They use an agent-oriented approach, defining house, human and appliance agents, which makes DRSim a highly modular software, capable of incorporating new models easily, and to confer different behaviours to its agents. Demand in appliance agents (R1) is inferred using conditional probabilities, combining information of the correlated distribution of type of occupant and house, rooms in the house, set of appliances in each room, activity taken place and the use of an appliance during that activity. More interestingly, user behaviour is incorporated into human agents, defining a responsivity parameter (how agent reacts on receiving a signal) to perform a DR action. Both responsivity and action are modelled using three parameters: price, perception and communication sensitivities. This way, users' decision making (R4) and its engagement (R5) are sophisticatedly

parametrized. Two types of DR mechanisms are allowed to the human agents: reduce load and shift load; but “the specific algorithm (R3) used to create a revised schedule is left unspecified as many candidates can be used.” On the downside, there is no attempt on this study to consider supply agents (R2), or to involve network operation in any way (R6).

Another DR simulator called DemSi [42], “aims at providing a flexible tool to analyse DR actions and schemes” by developing “a software application that can be used by DNO’s [Distributed Network Operator] and consumers to optimize their resource management.” DemSi focuses on distributed generation issues (R2), covering a broad range of renewable energy technologies (photovoltaic, wind, co-generation, fuel cells, small hydro and biomass), and its effects on the electrical network and energy market elasticity (R6). To do so, they use PSCAD<sup>13</sup> software for network simulation. Their objective is to simulate a variety of DR methods that minimize the costs of a generation shortage situation: minimizing the total cost that the DNO and the suppliers have to pay for non-supplied loads, which is implemented as a mixed-integer linear model in (commercial) software MATLAB and GAMS<sup>14</sup> (R3). The case of an isolated microgrid with only available electricity from DG gives some hints on how to simulate DR for network faults. It presents some interesting ideas for DR from a DNO perspective, however, energy use is considered with typical profiles, making this simulator incapable (and inextensible) of dealing with R1, R4 and R5 in its present form.

SMASH (SiMulated Adaptable Smart Home) [57] is claimed to be a demand side focused simulation platform, “specifically built to provide insights on the effect of different approaches to consumers, in terms of discomfort and decreased elec-

---

<sup>13</sup><https://hvdc.ca/pscad/>

<sup>14</sup><https://www.gams.com/>

tricity costs.” It mainly addresses R4, proposing an interesting DR programme: consumer-centric load control (CLC). It is an incentive-based programme similar to Direct Load Control (DLC), but with maintained freedom and comfort to consumers. The authors formulate a power reduction request sent by the DSO, and the home management system performs a load control action based on policy reasoning and extended finite-state machines<sup>15</sup>. The user will receive a reward based on the energy reduction, and a penalty fee in case of rejection. The objective is to minimise times of reduced comfort through the modification of space heating (R3). However, the other requirements are poorly fulfilled: only space heating and electric water heating (shower) are included (R1); it does not consider local generation or storage so far (R2); the interaction between end-users and the home management system is unclear (R5) and there is not intention for a multi-building simulation.

### 4.3.3 Proposed modelling framework

Although good progress has been made in the development of DR software, with interesting initiatives available in the literature, there remains no comprehensive framework capable of addressing R1 to R6 requirements.

The remainder of this chapter describes the extension to an existing Multi-Agent Stochastic Simulation platform (No-MASS), with the aim of comprehensively address the shortfall in DR simulation capability. We present a tool capable of dealing (completely or partially) with R1, R2, R3 and R6, and with potential to incorporate R4 and R5 in the future.

No-MASS was initially developed to model the presence, activities and related

---

<sup>15</sup>Decisions are made based on a set of possible states and a set of transitions or outcomes from each state.

behaviours of synthetic occupants (agents) of buildings that are co-simulated with EnergyPlus using the Functional Mockup Interface standard [58]. In this way it is straightforward to model occupants, interactions between them, (person-to-person: P2P) and their impacts on the energy performance and indoor comfort of the buildings they occupy, in contrast with most DR tools, which were designed to address Device-to-Device (D2D) interactions (see Figure 4.3).

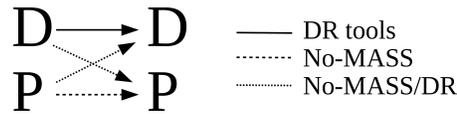


Figure 4.3: Device-to-device and person-to-person interaction.

Interestingly, the architecture of No-MASS is readily extensible to also consider D2D communications. Thus creating a platform that has the capability to simulate:

- Occupant-agents' behaviours and interactions between them (P2P, the prior No-MASS).
- Device-agents' behaviours and interactions between them (D2D).
- Interactions between occupants and devices (D2P and P2D).

In this thesis we focus on developing the second one of these capabilities. Next, the fundamentals of No-MASS are described in more detail, its workflow and the models that have been incorporated. The algorithms implemented to support load re-scheduling and battery operation in a multi-agent representation are then explained. In the next chapter we illustrate the application of this new prototypical platform to estimate the effects of appliance load shifting and electrical storage with the objective of maximising the use of locally generated renewable energy of *a*) a domestic building and *b*) a group of buildings.

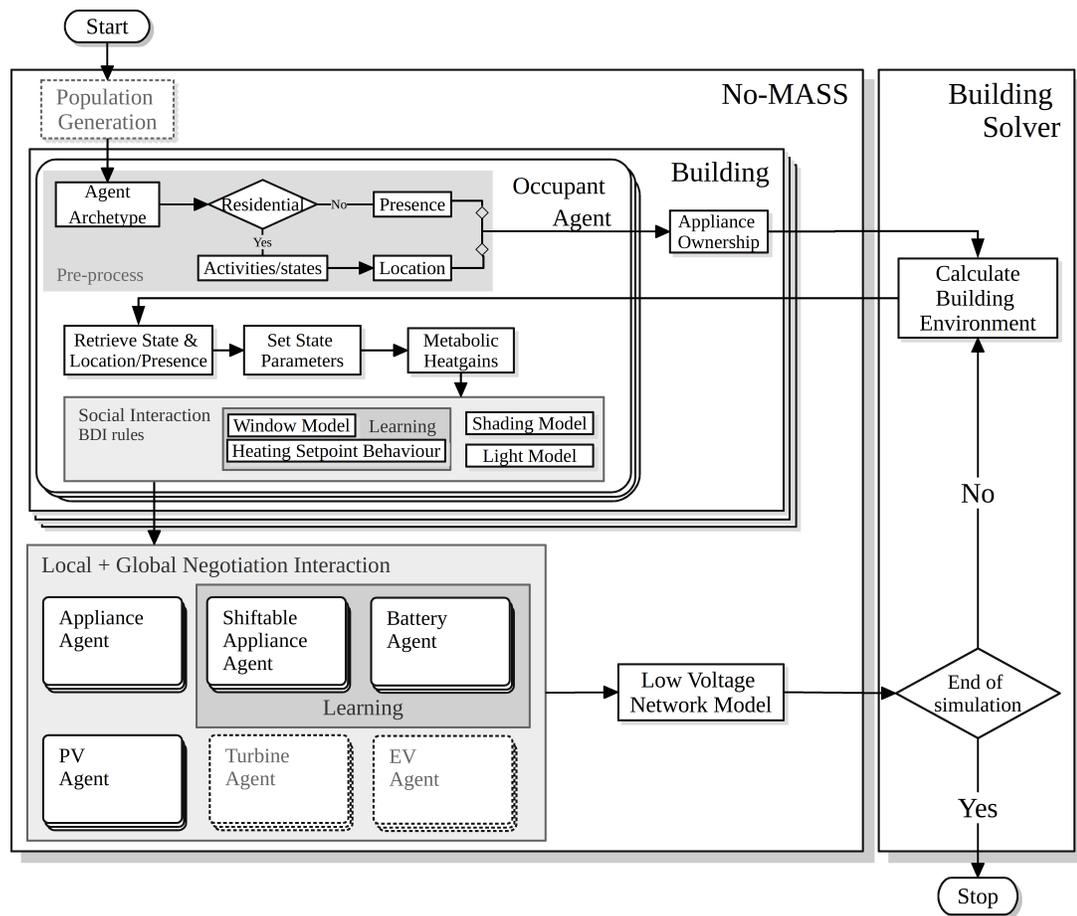


Figure 4.4: No-MASS flowchart.

## 4.4 No-MASS background

As noted earlier, No-MASS (Nottingham Multi-Agent Stochastic Simulation) is a platform that was originally developed to model the presence, activities and related behaviours of occupants of buildings, or groups of buildings, and the energy and indoor comfort consequences of these behaviours. To this end No-MASS employs four complementary modelling strategies:

1. **Data driven stochastic models** predicting occupants' presence [59] and associated metabolic heat gains, their activities whilst present [60] from which

location is also inferred, and their use of lights [61], windows [62] and shading devices [63].

2. The processing of votes for multiply-occupied spaces to emulate agents' **negotiation mechanisms** (e.g. agents' influence-weighted votes to open [1] or close [0] a window; the outcome receiving the greatest number of votes, weighted by influence, being effected).
3. Use of a **Belief-Desire-Intention** (BDI) framework to emulate agents' behaviours that are simple in character (e.g. closing curtains when it is dark, the closing of a window whilst bathing) but for which data is scarce.
4. **Agent learning** mechanisms to emulate agents' behaviours that are complex in character (e.g. choice of heating set-points) and for which data is also scarce.

The architecture of No-MASS (and its coupling with EnergyPlus), the data-driven (Strategy 1) models and their application to the simulation of both domestic and non-domestic buildings are described in [64]. The extension of No-MASS to model agents' negotiations and the data scarce modelling of both simple and more complex behaviours (Strategies 2-4) are explained in [65].

To facilitate the extension of No-MASS to handle DR simulation and optimisation (and indeed the more complete modelling of the impacts of occupants' behaviours), our Strategy 1 models have been complemented with models of occupants' ownership and use of large [60] and small [66] electrical appliances. These models have themselves been complemented with models and data of electrical storage (an electric battery) and supply or conversion (a photovoltaic panel) devices; and of a low-voltage network model.

#### 4.4.1 No-MASS workflow

Figure 4.4 illustrates the newly extended No-MASS architecture. First (upper left) a population generator creates a household of a size and demographic composition that is suited to the house being simulated [67]. Next the conventional No-MASS modelling tasks are executed.

Household member agents are assigned archetypal properties (room associated with activities, clothing and activity characteristics which randomly assign unique behaviour probabilistic models). Activities are then computed for the length of the simulation in a pre-process (for activities are not dependent upon environmental inputs); likewise electrical appliances are assigned to the household.

A loop then commences in which indoor/outdoor environment parameters are parsed from EnergyPlus to No-MASS for the present time step. From the pre-processed activities, agents' location, activity and clothing level are set, from which metabolic heat gains are calculated.

A series of models predicting interactions with windows, shading devices, heating systems and lights are then called. For heating interactions, agent learning is employed to determine the transient setpoints that minimise a cost function that combines heating costs and discomfort costs. In the case of multiply-occupied spaces, social interactions are considered at this stage through the vote processor to determine the negotiated outcome. Finally, BDI rules are used for straightforward interactions, for which data is scarce.

The workflow proceeds with the prediction of demands for small and large appliance, supply from energy conversion systems (only PV being enabled at this stage) and storage (only from batteries at present). Appliances are also shifted at this stage, again using agent learning, to maximise a cost reward function (of

which more later). A LV network model has been developed (see Chapter 6) and integrated with No-MASS. The user can choose to enable it for simulation.

The calculation then proceeds to the next time-step, exiting this loop at the end of the simulation period.

The corresponding physical models (for D2D modelling) are described below.

#### 4.4.2 No-MASS for Demand Response

The extension of No-MASS will, once complete (for this thesis describes a partial proof of concept), address the range of requirements outlined earlier. In its present form No-MASS/DR addresses requirements R1 (demand models), R2 (various power flows) and R3 (DR optimisation), further explained in this section. Progress on R6 is explained in Chapter 6. The strategies for achieving this are as follows:

- Develop mechanisms for D2D communications between energy conversion, storage and demand devices represented using the models described below.
- Implement strategies for load-shifting and optimal charge/discharge of a battery using agent learning algorithms. This will reflect whether or not multi-agent simulation and machine learning are effective approaches to test different DR strategies on different socio-demographic groups<sup>16</sup>.
- Extract or quantify differences when simulating residential electrical self-consumption / autonomy (or other DR objectives) for different socio-demographic groups.

---

<sup>16</sup>Socio-demographic considerations are not included in the work reported.

Thus, No-MASS could potentially (given available data) be used to evaluate scenarios involving human interaction and behaviour change (R4 and R5) and to support tariff design, using for instance BDI rules and/or agent-learning. For instance, we could assess to what extent active engagement of users increases electrical energy autonomy, or to what extent price signal impacts on behaviour change.

The following sections, explore in more detail the methods and algorithms used.

## 4.5 Methodology I: models and data in No-MASS/DR

There are a broad range of available technologies for Distributed Energy Resources in residential areas. In regard to supply technologies there is: solar PV, micro-wind turbines, CHP plants (that can be disaggregated into fuel cells and fuel combustion), small-hydro power or biomass plants. For small-scale electrical storage there are different battery technologies available (chemical), supercapacitors (electrical), flywheels (mechanical); other electrical storage options are available for larger systems, such as water pumping or compressed air storage. Additionally, electric vehicles can be considered as storage systems.

Within this landscape, for our proof-of-concept, we are selecting solar PV for power generation and electric batteries for storage systems.

### 4.5.1 Electricity demand models (R1)

The demand forecast models developed for No-MASS are based on stochastic methods. Devices have been classified in large appliances (high-load and commonly owned) and small appliances (range of low-load appliances), and follow two different modelling methodologies.

## Large Appliances

Large appliances include the cooker, TV, microwave, washing machine and dishwasher. They have been modelled as a three-step process [60]. First, the probability of switching on is predicted using a time-dependent Bernoulli process. Second, the duration for which devices will remain on is predicted using survival analysis. Finally, transitions between categories of fractional power demand (the fraction of maximum possible) are predicted as a Markov process, at 10 minute resolution. Probabilities of switching on are depicted in Figure 4.5.

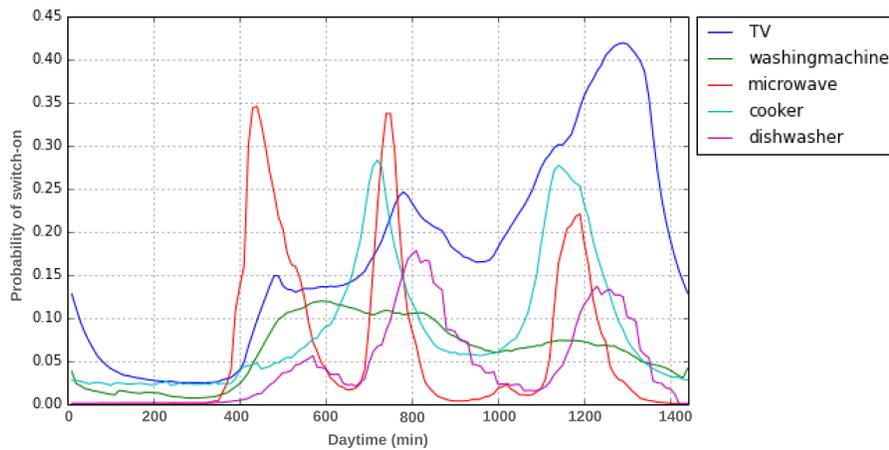


Figure 4.5: Switch on probabilities of large appliances.

## Small Appliances

Small appliances are modelled as aggregations of appliances, following four categories: small appliances in the kitchen, audio-visual appliances, computing appliances and other appliances [66], extensively explained in Chapters 2 and 3. A multi-state survival model is used: eleven fractional energy (ratio to maximum energy) states are defined  $f = \{0; 0-0.1; 0.1-0.2; \dots; 0.9-1\}$ . The survival time that the appliances remain in one of those states is calculated using a Weibull dis-

tribution (see section 2.4.3). A simplified flowchart of this process is represented in Figure 2.3(c).

## Heating

Electrical demand for heating can be obtained thanks to the coupling of No-MASS with Energy Plus.

Water heating is not considered for these simulations.

## 4.5.2 Supply and power flow (R2)

### Photovoltaic data

For the current purpose of demonstrating the proof of concept of our proposed modelling approach, we are using measured performance data that characterise PV systems' output in-lieu of a predictive model. In the future a suite of predictive energy conversion system models will be integrated.

### Electric Battery Storage System

Electric storage technologies provide valuable services in most energy systems. Small-scale systems are becoming (nearly) cost-competitive in certain situations, such as remote or off-grid scenarios [32]. Common applications for the use of small-scale electric storage are arbitrage (response times  $>1\text{h}$ ), demand shifting and peak reduction, variable supply resource integration (response times  $<15\text{min}$ ), frequency regulation (response times  $\approx 1\text{min}$ ) or voltage support (response times:  $0.001\text{s}$  to  $1\text{s}$ ).

In No-MASS/DR, an approximated model has been implemented for the charge and discharge of the battery, which considers the State Of Charge (*SOC*) of the

battery independent of its open circuit voltage  $V_{OC}$ . Conversion efficiency is fixed to  $\eta = 0.98$ , leading to constant conversion losses of 2% with respect to  $SOC$ . Given a nominal capacity of the battery  $Q_n$ , a maximum charge/discharge rate of  $P_n$  and a current capacity  $Q(t)$ , the  $SOC(t)$  is defined as

$$SOC(t) = \frac{Q(t)}{Q_n}. \quad (4.1)$$

The variation in  $SOC$  in each time step  $\Delta t$  (usually  $\Delta t = 1min$ ), is determined by the variation in capacity, from  $Q(t)$  to  $Q(t + \Delta t)$ . For the discharge of the battery, it is given by

$$Q(t + \Delta t) = Q(t) - \eta P(t)\Delta t, \quad (4.2)$$

where  $P(t) \in [0, P_n]$  corresponds to the requested power by the appliances. For the charge of the battery,

$$Q(t + \Delta t) = Q(t) + \eta P(t)\Delta t, \quad (4.3)$$

where  $P(t)$  is the available power generated locally. So far, a fairly primitive (dis-)charging strategy is defined, which only considers (dis-)charge from the PV panel (Figure 4.6). We plan to refine this strategy to also consider (dis-)charging from the grid (dashed lines in Figure 4.6) in the future.

### Low Voltage network model

The operation and modelling features of the Low Voltage network implemented is fully described in Chapter 6. A forward/backward sweep solver has been used for power-flow analysis. It has been implemented as a recursive algorithm. In the case of a distribution network, recursion constitutes a method to efficiently solve

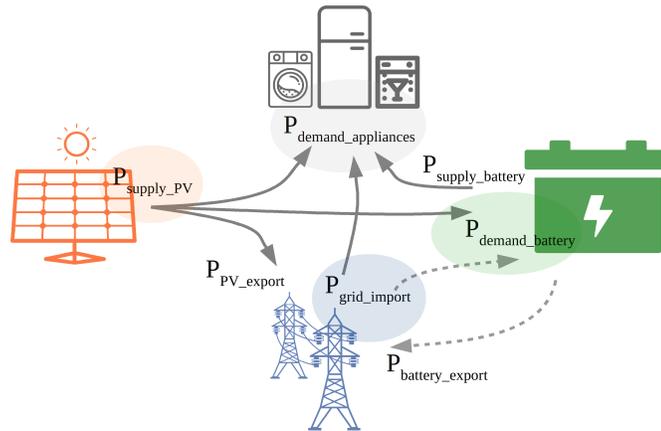


Figure 4.6: Power flows.

branched networks. During the forward sweep, the currents and voltages at each node depend on the currents and voltages at the child nodes connected to them (one if linear network, several if branched); whereas during the backward sweep, currents and voltages depend on the parent node.

## 4.6 Methodology II: Software orchestration with Agent Representation

### 4.6.1 Agent representation

In the initial methodology of No-MASS, household occupants were identified as software agents. In extending No-MASS to No-MASS/DR to handle D2D interactions, electrical appliances and energy systems are represented as an agent sub-class. Each agent has an ID, a peak power (either for demand or supply) and a priority of service. The priority of service is used for demand agents, and

it allows to rank all energy requests based on their priority, and it is activated when restrictions in supply are considered. This would be useful, for example, for simulations of an islanded network, to account for situations when not all services can be delivered. Also, the operation starting time and profile can be calculated (type of prediction depends on the type of device), given that information of the associated model is entered in the agent definition.

During each time step, each device communicates the amount of energy (power integrated over the duration of the time step) to be requested or delivered. First, energy from our PV panel is allocated to appliances based on their priority. Any shortage is provided (perfectly) by the upscale grid. When an electrical storage device is unavailable, any surplus is exported to the grid. When it is available, the battery can be charged using excess PV energy, and can be later discharged to run electric appliances when price conditions favour this strategy.

### **4.6.2 Type of agents**

Every actor in No-MASS/DR is configured as an agent. Thus, there are seven types of regular agents: LargeAppliance, SmallAppliance, PV, Battery, Grid, CSV and FMI agents. FMI agents are those linked to another software through Functional Mock-up Interface (FMI) for co-simulation; in the case of No-MASS this software is EnergyPlus. CSV agents offer the possibility to enter .csv data files as input, instead of specific models, giving flexibility to add demand or supply devices for which models are not available yet.

Additionally, two of these agents have been conferred with learning capabilities: LargeApplianceLearning and BatteryLearning. The specific details of the mathematical formulation for their operation is presented in the following section.

## 4.7 Methodology III: Agent Learning for DR optimisation (R3)

In its present form, No-MASS/DR combines the following components (in relation to Figure 4.7):

- A. Price-based strategies are considered: a cost signal based on indicative prices for electricity import is used as a **driver** stimulus to modify demand patterns. The specific tariff structure can be defined and modified for each simulation scenario. The current demand/supply status influence as well the response.
- B. Two **DR mechanisms** are implemented: load shifting (through appliance reschedule) and optimal electrical battery operation. A reinforcement learning algorithm is incorporated into learning agents: *a)* LargeApplianceLearning or *shiftable* appliance agents, that can be regulated autonomously: washing machines and dishwashers, and *b)* BatteryLearning agents, that learn optimised discharge operation.
- C. The selected **DR mathematical formulation** consists of a Q-learning algorithm for the optimization of the objective function.
- D. In this thesis, we have focused on the improvement of renewable energy self-consumption using cost minimization as our **objective function**. This cost is defined inside the reward function for each case, and consist of a combination of electricity tariff price, available renewable energy and intensity of the energy demand. We aim at finding a formulation for the reward function such that minimization of the objective measure leads to an indirect increase of self-consumption.

All these four elements are explained in more detail in the following sections.

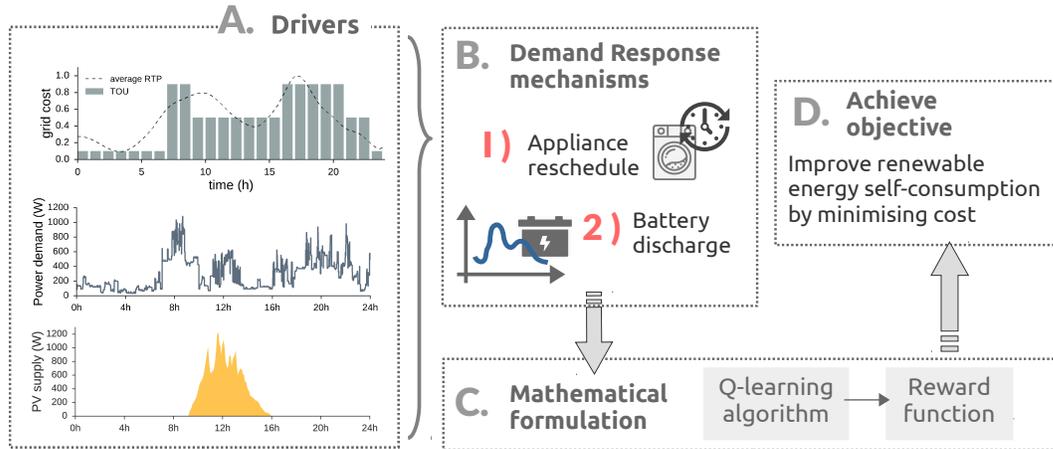


Figure 4.7: DR methodology.

### 4.7.1 Q-learning algorithm

There are multiple algorithms available for Reinforcement Learning (RL). Q-learning algorithms [68] are one of the most widely used RL methods, with its major advantage being their great simplicity. They require a minimal amount of computation and, on its basic formulation, they can be expressed by single equations [69], and easily implemented in computer programs.

Q-learning algorithms [68] allow agents to learn a response from a reward, to an action. This allows agents to develop an understanding of their preferences over time. An agent learns the best action in a given state by trying every action in a state and updating the expected reward with the actual reward for that action. This is particularly useful, compared to other machine learning methods, for modelling appliance shifting, as the appliances can test different strategies for maximising their reward where there is a clear link between an action and a driving stimulus. For example, does a chosen action (turn on at a later time) reduce peak

power demand over a time-period. This is a quick and effective methodology that would be difficult to model through rules due to the complexities involved, especially when considering multiple shifting appliances. Each appliance would need its own set of rules to ensure they would not turn on at the same time, whereas Q-learning allows them to learn their own preferences, considering other appliance demands.

In Q-learning algorithms, an agent chooses an action at a given state based on a Q-quantity, which is a weighted reward based on the expected highest long term reward [68]. The Q-quantity is defined for each state-action combination, creating a Q-table. The values in the Q-table are updated each time an agent selects an action. Let  $s_t$  express the agent's state at time step  $t$ , and  $a_t$  a chosen action. Using this information, the Q-value for the corresponding combination of  $(s_t, a_t)$  is updated:

$$Q_t(s_t, a_t) = Q_t(s_t, a_t) + \alpha [\mathcal{R}_t + \gamma * \max(Q(s_{t+1}, a)) - Q_t(s_t, a_t)], \quad (4.4)$$

where  $\mathcal{R}_t$  is the reward observed for the current state and the action taken. For each combination of  $(s_t, a_t)$ ,  $\mathcal{R}_t$  can be a single value or it can be a function depending also on other variables.  $\alpha \in (0, 1]$  is the learning rate and  $\gamma \in (0, 1]$  is the discount factor. The term  $\max(Q(s_{t+1}, a))$  is an estimate of the optimal future value; thus, the discount factor specifies how soon the agent cares about the reward: near terms goals when  $\gamma \sim 0$  (myopic agent), otherwise long term rewards when  $\gamma \sim 1$ . In summary, at time-step  $t$ , when the agent at state  $s_t$  is taking action  $a_t$ , the Q-value is updated using: its former value  $Q_t(s_t, a_t)$ , the reward  $\mathcal{R}_t$  and an estimate of the optimal future value  $\max(Q(s_{t+1}, a))$ . The effectiveness of the learning process is highly dependent on the selected parameters  $\alpha$ ,  $\gamma$  and  $\epsilon$ .

Once the Q-table is updated, the next state is set as the current state, and the agent will select a new action. The selection of an action is not completely deterministic (using the Q-table), but uses an epsilon-greedy approach: the best action is selected with  $1 - \epsilon$  probability, and a random action is selected otherwise. For instance, if  $\epsilon = 0.1$ , the (currently) best action will be adopted 90% of the time. This randomness is introduced so that the agent explores more in order to discover the best action over the whole period of time.

### **Reward function**

The reward function  $\mathcal{R}_t$  can consider different variables of interest, such as cost, power demands or voltage stability, allowing to use the same methodology to explore a range of objectives. An advantage of Q-learning over other black-box machine learning methods such as neural networks, is that the definition of the reward function allows to tune the algorithm in a explicit way, using knowledge of the system. Convergence of the learning process is achieved when the change in values of the Q-table no longer affects the results.

The reward function is different for each learning agent type. Thus, it needs to be formulated for each case study.

## **4.7.2 DR mechanisms**

### **I. Load shifting optimisation**

The core idea for appliance-reschedule is depicted in Figure 4.8. For each time step, the switch-on model is run. When a switch-on event is predicted, the profile of use and the new starting time are calculated. The temporal window within which appliances may be shifted has a maximum extent of 24h, but this may be

reduced to meet pre-defined delivery time constraints.

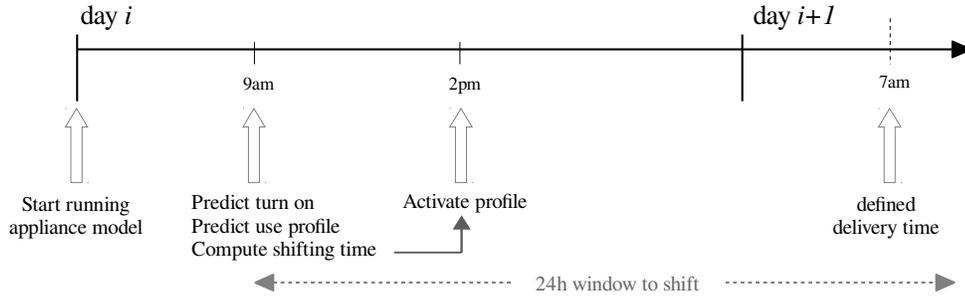


Figure 4.8: Appliance re-schedule diagram. In this example, state is  $s_t = 9(h)$ , and action  $a_t = 14(h)$ .

**Q-learning for load shifting** For a LargeApplianceLearning agent, we map the state (hour of the day) to an action (future hour of day to initiate the appliance). This creates a mapping of 24 hours to 24 hours, making the Q-table space of 576 combinations in size. An example of a profile shift might involve calculating whether at state  $s_t$  the appliance is required to be turned on (e.g. the dishwasher is loaded and ready). If so the appliance demand profile should be calculated using the large appliance model. However this should not yet be initiated. Instead the Q-table should be used to retrieve a time (action) at which the appliance programme should be initiated (see Figure 4.8).

A graphical example of a  $24 \times 24$  Q-table for appliance reschedule is presented in Figure 4.9a, in which each pixel is false coloured according to the q-value of an action from its state  $s_j$  to  $s_{j+1}$ . Shifting to peak times (7-9h and 16-21h) has a smaller reward (darker pixels) than shifting to sunshine hours, or off-peak (23-7h) times. Scope of reduce dimensionality should be explored, either visually or using cluster analysis, as it could allow a faster learning period.

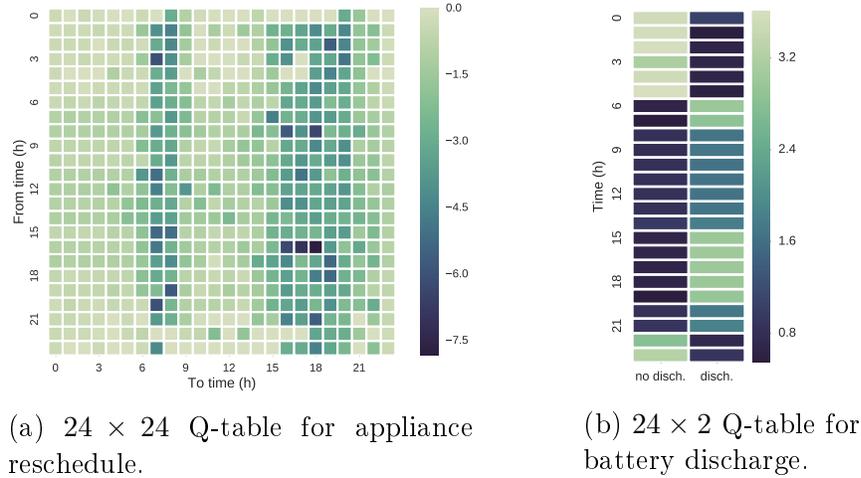


Figure 4.9: Q-tables. The units of the legend correspond to reward values, derived from the corresponding reward function.

*Reward function.* The reward function  $\mathcal{R}$  for appliance reschedule consists of two components:

1. The inverse of a cost signal (effectively then an income signal), based on indicative prices for energy imports. A Time-Of-Use (TOU) tariff can be defined (specified for each case study). Values of the tariffs have been normalized  $[0, 1]$ , since only relative differences are useful for the learning algorithm.
2. Negative reward (punishment) when the service is not satisfied on time (the washing cycle is not complete within the delivery-time-constrained window).

## II. Battery discharge operation

Our battery has been implemented to charge whenever there is a surplus of PV power available (supply  $>$  demand), before it is exported to the grid (whilst storage capacity is surpassed). For the discharge process, Q-learning is used.

**Q-learning for battery discharge** For a BatteryLearning agent, we map the state (hour of the day) to an action (either discharge [1] or not discharge [0]). It leads to a mapping of 24 hours to 2 actions, and a Q-table space of 48 combinations in size.

A graphical example of such  $24 \times 2$  Q-table is presented in Figure 4.9b, in which each pixel is false coloured according to the q-value of the action  $a^{(1)}$  of discharging and  $a^{(0)}$  for not discharging, for each hour of the day.

*Reward function.* The reward function  $\mathcal{R}$  for the discharge process is dictated by either one or both of these two considerations:

1. *Alleviate energy intense periods.* The battery has been configured to relieve peaks of high demand that require electricity imports. The battery learns when the highest hourly peak demand  $P_{grid\_import\_MAX}$  is and when imports from the grid exceed some threshold  $\delta$  (e.g. 70% if  $\delta = 0.7$ ) of this demand:

$$P_{grid\_import} > \delta P_{grid\_import\_MAX}. \quad (4.5)$$

Otherwise, it learns to turn itself off in that hour, storing the remaining energy for peak periods.

2. *Minimise high-cost grid imports.* The battery can also recognize favourable price conditions. It will learn when to discharge based on the price signal, discharging in high-price periods and not discharging during cheaper periods.

This mechanism is commonly known as arbitrage.

The activation of such considerations is defined in the reward function for each case, which needs to be tuned accordingly. The specificities of the reward functions used for this work are described in the corresponding case studies.

### 4.7.3 Self-consumption maximization

Self-consumption (SC), also referred to as a load matching index [70, 71], can be defined as the ratio of energy use from on-site generation to the total energy used, as expressed in equation 4.6. It is inversely related to the amount of PV (or other sources) power exported to the grid.

$$\text{Self Consumption}(\%) = \frac{E_{\text{demand from PV}}}{E_{\text{demand}}} \quad (4.6)$$

Although maximising self-consumption is our goal, this is not formally optimised by our Q-learning algorithm (eq. 4.4), as it is not explicitly expressed in our reward function. But this is indirectly achieved through the tariff signal, (Figure 5.1), through which we pay a low price whilst using off-peak centrally-generated energy and close to zero for locally generated energy. In this way, smart appliances learn to turn on during low-cost periods.

## 4.8 Summary

In this chapter, we have described a methodology extending the multi-agent stochastic simulation platform No-MASS to support the simulation of demand response strategies. The prior focus of No-MASS was on the integration of models of occupants' behaviours with building simulation software, in particular with EnergyPlus. The hypothesis is that No-MASS's underlying software architecture, and many of the modelling techniques already utilised within it, are readily extendible to handle DR simulation (simulating device-agents in addition to occupant-agents).

A set of six requirements have been formulated (Objective II.1 of this thesis) that we believe should be supported by any DR simulation tool. They include

modelling power loads from various type of devices, and managing demand with a combination of local generation, stored electricity and traditional supply systems, using mathematical optimization. Apart from this device-to-device interaction, we also suggest person-to-device interactions as a significant functionality for DR software. Various interesting mathematical formulations and model candidates were found in the literature search, but none of the simulation platforms studied aimed at addressing all of the requirements in a comprehensive fashion.

The proposed framework, No-MASS/DR, has the potential to do so. The details of No-MASS/DR and its methodology have been presented here, and allows to address objectives II.2 and II.3, which will be demonstrated in the next Chapter. Key components include the use of agent reinforcement learning to simulate two different DR mechanisms (load shifting and battery discharge) to improve energy use of renewable local sources. In the next Chapter, two different scenarios are evaluated: a single house and a community of houses.



## Chapter 5

# Application of the extended No-MASS/DR framework

The previous chapter introduces ideas to solve the DR problem using Multi-Agent Simulation. A specific software architecture has been presented. Machine learning algorithms, in particular Q-learning, have been proposed as good candidates to handle load shifting optimisation of automated appliance use, as well as the discharge process of electric storage. Its mathematical formulation requires a *reward function* which will be key to achieve the corresponding DR goals. In this chapter, all those concepts are put into practice. First, we present the application of the framework to a single-building case study, comparing improvements in the rate of renewable self-consumption when load shifting and storage capabilities are introduced. Second, we proceed to model a community of buildings synthetically populated, which brings new software design and computational challenges. Again, we test the effects in electricity self-consumption when new technologies are introduced.

### 5.1 Improving Self-Consumption

The two case studies presented in this Chapter consider the application of the No-MASS/DR framework for the study of renewable energy self-consumption, SC (see Section 4.7.3). Two main arguments support the efforts in improving on-

site power consumption. First, SC increases the economic value of distributed generation [72]. In a context of constantly increasing retail electricity prices, and declining investment costs for renewable technology, improving the rates of use of self-generated electricity contributes directly to achieve grid parity<sup>1</sup>. For the case of PV, which has the highest share in DG, grid parity has already been achieved in different European countries, and this is expected to become more widespread in the coming years [74].

Second, SC can lower the stress on the electricity distribution grid, caused by high levels of penetration of DG. Integration of intermittent power sources is intrinsically challenging, due to the disparity between power generation and power demands. Moreover, it can cause power faults such as voltage rise due to the peaks in generation during low-demand periods. Increased rates of SC can help to mitigate such effects, as well as reducing transmission losses.

However, the profitability of SC is unavoidably attached to the existing regulatory framework of grid-connected systems, which is different in each country. In the EU, there is not a specific regulation or legislation on self-consumption [75]. Subsidies or attractive policies are necessary that make SC financially beneficial for prosumers. Some models that consider remuneration for on-site electricity use are Feed-In-Tariffs (FIT's)<sup>2</sup>, net-metering<sup>3</sup> or Feed-In-Premiums (FIPs)<sup>4</sup>. All these policies support SC directly.

---

<sup>1</sup>For PV, grid parity is defined as “the coming of age moment, when electricity from PV will be cost competitive with that from conventional generation sources, without subsidies so that deployment will take off driven by economic fundamentals”[73].

<sup>2</sup>Prosumers pay the retail price for the power they use from the grid [75], and the supplier pays a generation tariff for any electricity generated and, where applicable, an export tariff for any surplus electricity exported to the grid [76].

<sup>3</sup>Prosumers feed excess electricity into the grid and consume it later when they need to, paying only for the net difference [75].

<sup>4</sup>Remuneration for RES generation, which introduces short-term market exposure of RES electricity [75].

On the other hand, SC is supported indirectly through the promotion of DR and electricity storage technologies [75]. SC is limited without technical enhancements of these two technologies. An average figure between 17% and 44% SC is suggested by [71] without any additional technology, depending on household size and irradiation exposure. To reach the full potential of on-site electricity use, we need to look into solutions that include demand response and electricity storage.

### **Self-consumption and community energy**

SC can be further achieved in a community energy framework. Becoming popular in the UK<sup>5</sup>, community energy projects “cover aspects of collective action to reduce, purchase, manage and generate energy”[77]. Distributed generated electricity, from collective or individually-owned resources, can be used for demands across the community, instead of being fed directly into the grid, enhancing the use of on-site generation. This idea, can be extended also to storage devices. In a community energy model, storage systems could be charged or discharged with generation surplus or demands from multiple households.

In this chapter, we test our tool for self-consumption maximization of *i*) a single building and *ii*) local neighbourhood (not necessarily detached from the city), that can operate as a community. In this, we assume that occupants are willing to adopt new technology and/or to modify their energy-using practices to improve upon their energy performance (albeit expressed through financial rewards).

---

<sup>5</sup>Interesting source to develop community energy projects: <http://www.planlocal.org.uk/>

## 5.2 Case study I - maximise self-consumption for a single building

### 5.2.1 Scenario description

In this section we present the results of simulating a household with a single professional resident occupying a detached house using No-MASS. Our objective is to maximise the utilisation of energy converted by a 3.8kW-peak PV panel installed on the roof by a set of large<sup>6</sup> and small appliances<sup>7</sup>, including electrical heating, connected to a 2kWh electric storage system (with a power limit of 1kW). Following notation in equations 4.1, 4.2 and 4.3 for the battery model:  $Q_n = 2\text{kWh}$ ,  $P_n = 1\text{kW}$  and  $\eta = 0.8$ .

To this end we use price-incentive based DR control strategies to: *a)* re-schedule autonomously controlled washing machine and dishwasher appliances; *b)* discharge the battery. A Time-Of-Use (TOU) tariff is tested, with three different pricing periods: on-peak (between 7-9h and between 16-21h), off-peak (at night between 23-7h), and mid-peak for the rest of the day. In Figure 5.1, the TOU signal is related to an averaged Real Time Price (RTP) signal<sup>8</sup>. Values of the tariffs have been normalized  $[0, 1]$ , since only relative differences are useful for the learning algorithm. For both DR mechanisms, device operation can be tuned using this tariff signal as a driver, that fosters the switch on of autonomous appliances whilst sunshine is likely, and the discharging of the battery whilst there is no PV generation. This operation is aimed at maximising self-consumption over time, while

---

<sup>6</sup>Cooker, dishwasher, washing machine, fridge and TV.

<sup>7</sup>Set of audio-visual appliances, small kitchen appliances, computing appliances and other appliances. The set size depends on the simulation, in which they are allocated following a Monte Carlo sampling.

<sup>8</sup>Data available from [http://bmreports.com/bsp/bsp\\_home.htm](http://bmreports.com/bsp/bsp_home.htm)

reducing on-peak demand (as defined in Figure 5.1).

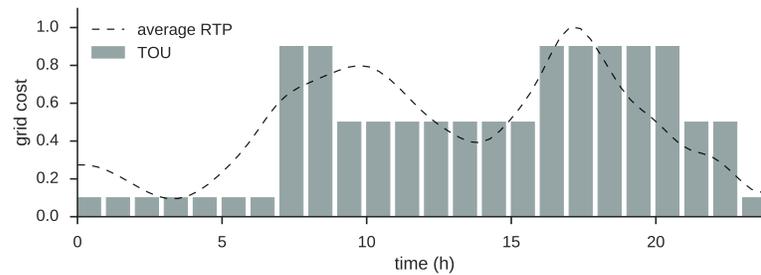


Figure 5.1: Reward function uses price signals.

## 5.2.2 Reward functions

As in other machine learning algorithms, a trial and error process is necessary to identify suitable algorithm parameters. In Reinforcement Learning, this is the case for the definition of the reward function. However, in contrast with other black-box machine learning algorithms (such as neural networks), Q-learning allows the modeller to incorporate "expert knowledge" about the operation of the system, through the explicit formulation of the reward function. Thus, different formulations lead to more or less efficient learning. Two ways of evaluating this are: to look at the relative reward values, which should increase with time, as the agents learn the more profitable actions (example in Figure 5.2); or to look at the expected behaviour of the system (for instance, checking when the appliance switch on events occur).

### Appliance re-schedule

The reward function for load shifting was explained in Section 4.7.2-I; being a combination of the inverse cost signal (from the TOU) and a punishment when the service is not satisfied within the defined delivery time  $t_D$  (scheme in Figure

4.8).  $t_D$  is specified for each appliance in the simulation configuration file. The value for the delay punishment  $p \geq 0$  is also input through the configuration file. For this simulation, we have set  $p = 1$ .

$$\mathcal{R}_{a(t)} = \begin{cases} -C_{cycle}, & t < t_D \\ -(C_{cycle} + p), & t \geq t_D \end{cases} \quad (5.1)$$

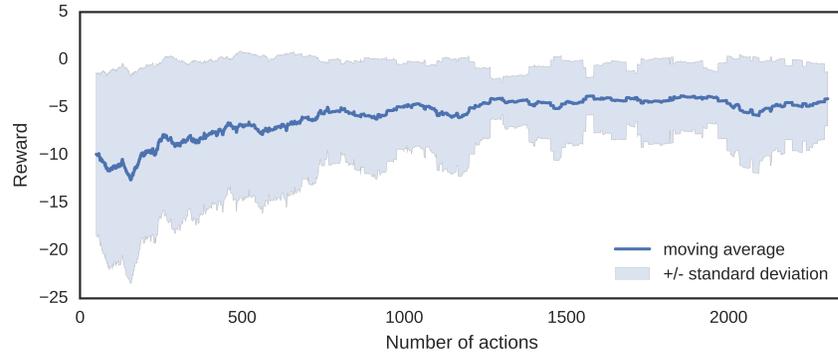
where  $C_{cycle}$  refers to the cost of running once the appliance  $a$  starting at time  $t$ , and depends on the electricity use and price.

An example of the learning period is presented in Figure 5.2. It shows a moving average of the reward values for the actions taken by the washing machine agents (5.2a) and the dishwasher agents (5.2b), and their standard deviation. In both cases, reward values increase at the beginning of simulation and the standard deviation reduces, meaning the agents are learning. This trend is more severe up to approximately the first 1000 actions; after that, it becomes more steady. For this work, we have selected 1000 actions as a minimum recommended value for the appliance re-schedule learning process. For each particular case (depending on the number of devices or households considered), the number of learning actions can be increased to 1500-2000 actions.

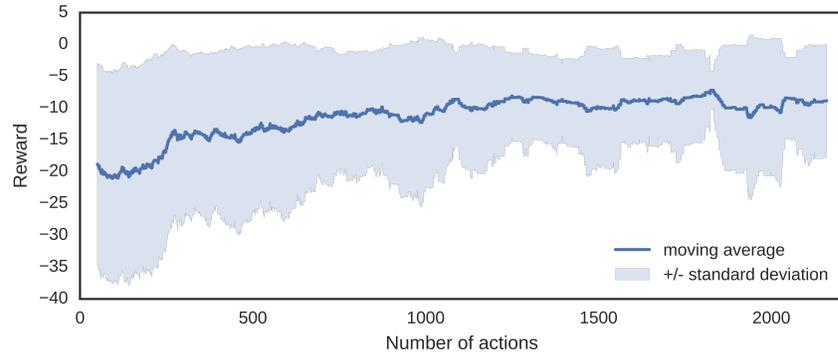
### Battery discharge

The battery discharge in this case study takes into account the first mechanism of those defined in Section 4.7.2-II: alleviation of demand-intense periods. Following equation 4.5, we define  $\chi$  as:

$$\chi = \frac{P_{grid\_import}}{P_{grid\_import\_MAX}}. \quad (5.2)$$



(a) Washing machine



(b) Dishwasher

Figure 5.2: Appliance learning: improvement of reward.

For each hourly time-step,  $\chi$  gives the ratio between current grid imports  $P_{grid\_import}$  and the highest peak demand  $P_{grid\_import\_MAX}$ . For this case study, we select the threshold for  $\chi$  to be  $\delta = 0.7$  (see Section 4.7.2-II). Again, a trial and error process was performed to obtain the parameters values.

Depending on the action taken at each time step, either  $a^{(1)}$  for discharge or  $a^{(0)}$  for not discharge, the reward  $\mathcal{R}$  varies. Therefore, the reward when discharging is:

$$\mathcal{R}_{a^{(1)}} = \begin{cases} \chi, & \chi \geq \delta \\ -\chi, & \chi < \delta \end{cases} \quad (5.3)$$

and when not discharging:

$$\mathcal{R}_{a^{(0)}} = \begin{cases} -\chi, & \chi \geq \delta \\ \chi, & \chi < \delta. \end{cases} \quad (5.4)$$

When the reward is set to a negative value,  $-\chi$ , it will act as a punishment, occurring when discharging at low-demand periods (eq. 5.3) or when not discharging at high-demand periods (eq. 5.4).

### 5.2.3 Results and discussion

Three scenarios are presented for comparison: first, the base case, where the models and systems are run without shifting demand and without considering a battery. Second, appliance re-scheduling capabilities are added to the base case. Third, a battery system is also considered. These three scenarios are simulated for winter and summer (heating not necessary).

Each simulation runs for a period of one week. To account for stochasticity in the calculations, results are presented as a distribution from a set of replicates; 100 for each scenario. Where agent learning is involved, the Q-learning algorithm requires training to populate the Q-tables. For the work presented here, a learning period of 125 weeks was necessary (learning process was explained in the previous section). This may seem long, but it is understandable and in accordance with the previous section, given the nature of the events: a turn-on of the washing machine or dishwasher will seldom occur more than once a day, and very rarely more than twice. Efficient learning needs a large sample of events (as was shown in Figure 5.2), and will thus need a relatively long learning period. However,

this period could be reduced by defining fewer states and actions, thus reducing the dimensionality of the Q-table. In relation with the number of actions in the previous section, considering roughly 10 switch on events per week, around 1250 actions were taken into account.

One day of a winter simulation is presented in Figure 5.3. The electrical demand consists of large appliances, small appliances and heating (delivered by an electric heating system). Demand arising from shifted devices (washing machine and dishwasher) and heating demands are indicated by shading. It can be seen that the total load is comparable to the amount of energy generated on-site in winter. However, occupants' schedules and their main use of electrical devices does not always occur during sunshine hours. Nevertheless, it can be seen that part of the flexible load has been shifted to the middle of the day where it can operate using locally converted power.

When a battery is available, the excess of solar energy is stored. Thus, the charging power (limited to 1kW) is shown in the middle of Figure 5.3.

As a consequence of the introduction of the TOU tariff, we can simulate how the DR algorithm adjusts when electricity is consumed from the grid. For our three-case scenario, results of the weekly energy grid import at the different tariffs, as well as PV export, are represented for summer and winter in Figure 5.4 (the bars represent average values for a set of 100 simulations). The first observation is the dramatic difference between winter and summer, due to the amount of solar energy available in summer, which visibly increases PV exports. The introduction of the tariff scheme with appliance and battery technology is efficient in reducing on-peak and mid-peak grid imports, in both summer and winter periods. The effect of load shifting is to reduce on and mid-peak imports by increasing the use

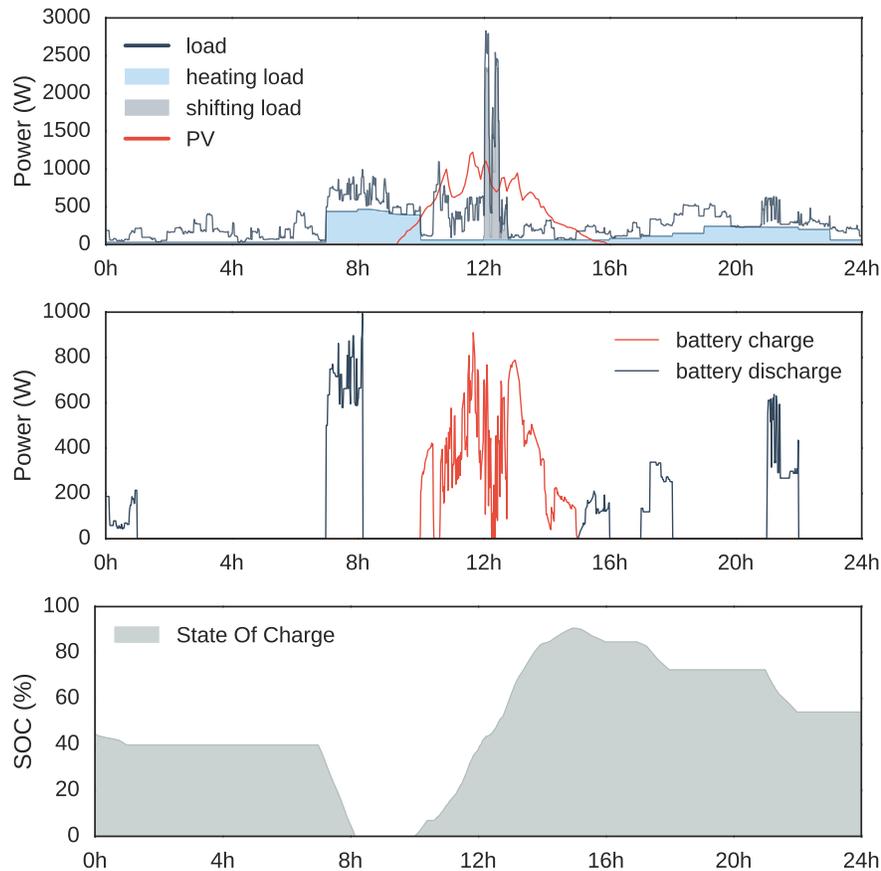


Figure 5.3: 24h of simulation for a winter day. PV profile and electrical demand on top; battery charge and discharge in the middle; SOC level at the bottom.

of cheaper electricity during off-peak hours (at night). This is specially relevant in winter, as solar energy may not be enough to run the complete programme, ending up paying high prices. As a consequence, the smart appliances learn to switch on predominantly at night. Furthermore, a mean of around 7kWh of solar energy is exported in winter when there is no storage system available. A dramatic reduction close to 98% is obtained when the battery model is introduced. This shows that storage systems can play a significant role when considering the electrical autonomy of distribution networks. In summer, there is not such reduction on the PV exports, suggesting that the algorithm for discharge learning needs some revision.

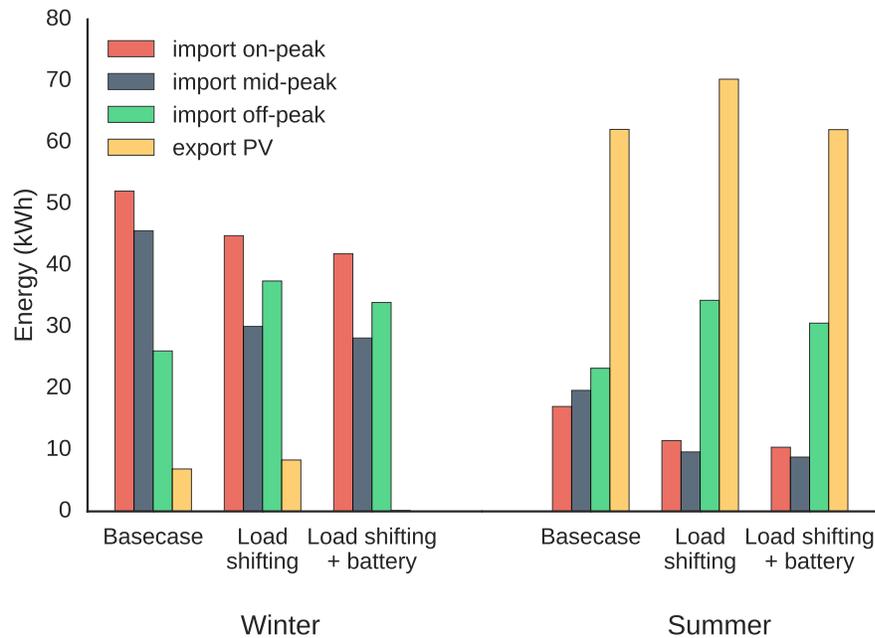


Figure 5.4: Import and export for the three scenarios, for summer and winter.

Also, although the algorithm employed is relatively simple in its formulation, it is nevertheless efficient in reducing on-peak and mid-peak grid imports, whilst maintaining cheaper off-peak imports and optimizing the use of solar energy.

Improvements in self-consumption (equation 4.6) are specially relevant in winter (in summer this effect is neutralized as solar energy is highly available in the basecase as well), with an increase of 52% with respect to its initial value, as presented in Figure 5.5. This figure in summer represents 8.1% increase in SC. However, the difference obtained only with the shiftable appliances is negative in summer. This is partly because the shifting occurs to night time hours, and partly because the two appliances considered represent a small proportion of the total energy demand, suggesting that further shiftable autonomous devices should be considered to better understand the potential for this strategy, and/or that new tariff structures are explored that encourage this behaviour.

This first case study was particularly useful for understanding the application of the algorithms and their behaviour with the tariff structure. Refinements to them, in terms of the reward functions for the battery discharge process, are thus applied and tested in the next section.

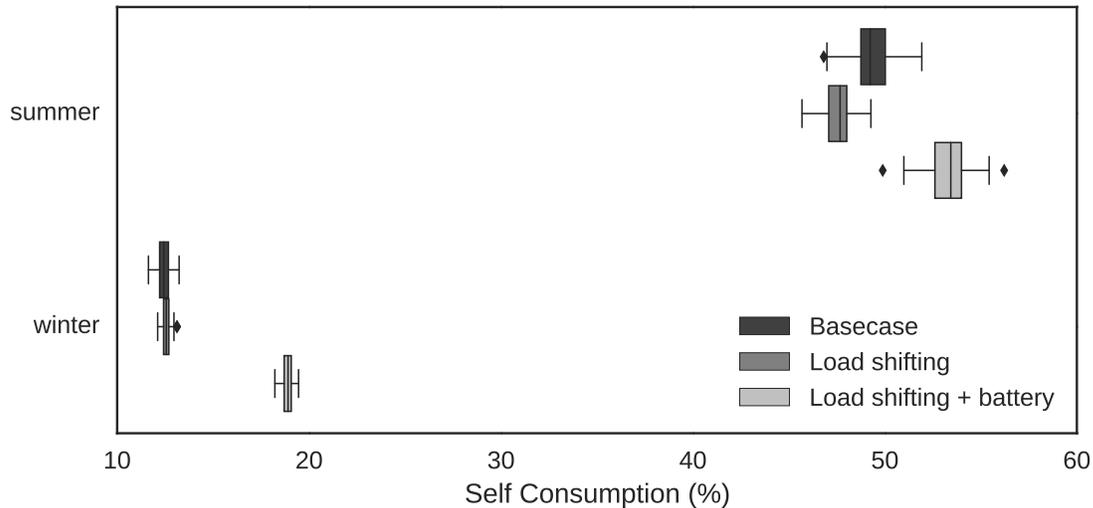


Figure 5.5: Self consumption index for the scenarios, for summer and winter.

## 5.3 Case study II: community of buildings

### 5.3.1 Scenario description

A small residential neighbourhood of 6 households is considered for this case study. Four of them have installed 3.8kW-peak PV panels on their rooftops, and two have installed 1.9kW-peak panels, following a smaller investment. As a community, they are willing to adopt smart appliances (washing machine and dishwasher) and individual electrical storage in their homes. No-MASS/DR enables to simulate the effects of the introduction of these technologies for this neighbourhood, including an estimation of the improvements in self-consumption, both individually and

collectively.

### 5.3.2 Implementation: multilevel coordination

When generalising No-MASS/DR from a single building to a multi-building simulation, new software challenges arose. There were limits in computational speed, data storage and data exchange, which had to be overcome, using data compression methods, or by incorporating parts of the post-process analysis into the simulation, so that less data had to be written. In that sense, although the multi-building simulation did not involve major conceptual changes, the software architecture had to be revised and enhanced.

The current operation of No-MASS/DR for multi-building simulation requires three-level coordination. The agent interaction and negotiation occurs then in three steps (see Figure 5.6):

1. *Local level.* Each house individually attempts to match power demand and supply.
  - Appliance agents request power for their operation.
  - If solar power is available:
    - PV agents supply power to the appliances (following their priority of use, defined in the configuration file), until all the demands are fulfilled.
    - The excess of PV is used to charge the battery.
    - After charging the battery, any surplus of PV is made available for the neighbourhood level.

- If solar power is not available, or it is not enough to fulfil appliance power demands, the battery discharge algorithm is run, and a decision to discharge or not discharge is taken.
    - If discharging, appliance demands are satisfied by the battery.
    - If the battery supply is not enough to fulfil the appliance power demands, information on the power shortage is communicated to the neighbourhood level.
2. *Neighbourhood level.* At this level, information is available for each house concerning: any appliance demand shortage, any PV surplus and any battery which is uncharged. A second process of power matching occurs.
- (a) Demand shortages are satisfied by solar energy from the neighbourhood.
  - (b) PV power surplus is used to charge the uncharged batteries (which were not used for discharging during the same time step).
3. *Global level.* Unsatisfied demands are supplied by grid power imports. Excess solar power is exported to the grid.

### 5.3.3 Reward functions

#### Appliance re-schedule

Shiftable appliances follow the same reward function as for Case Study I, explained in Section 5.2.2, formulated as:

$$\mathcal{R}_{a(t)} = \begin{cases} -C_{cycle}, & t < t_D \\ -(C_{cycle} + p), & t \geq t_D \end{cases} \quad (5.5)$$

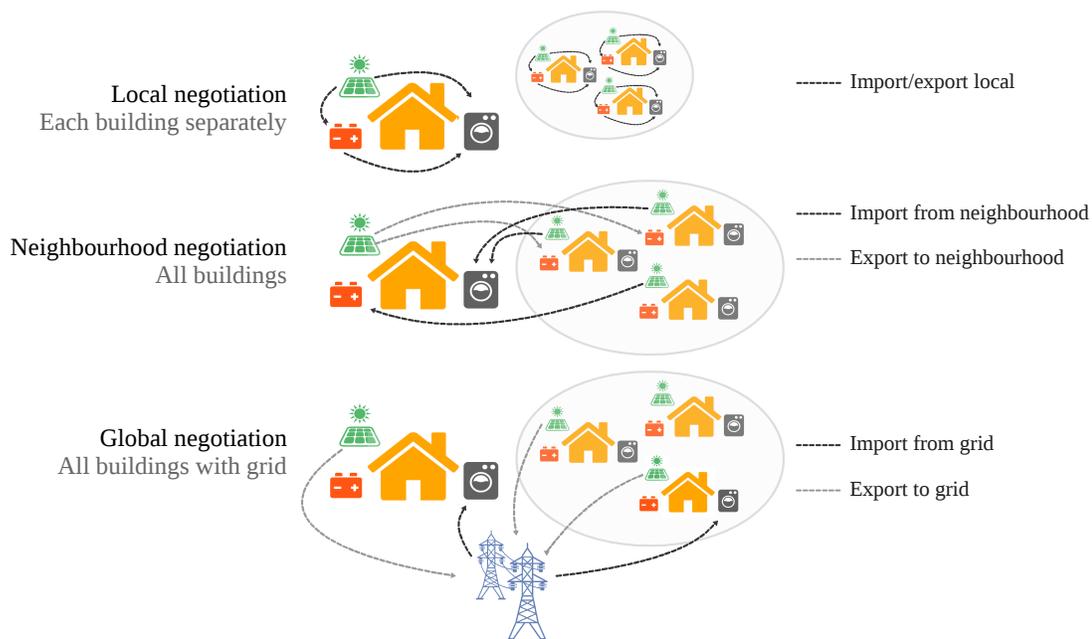


Figure 5.6: Multilevel coordination.

where  $p$  and  $t_D$  are specified in the simulation configuration file.

### Battery discharge

Battery discharge considers the two mechanisms defined in Section 4.7.2: a combination of alleviation of demand-intense periods and an arbitrage strategy. Thus, the reward values depend on the variable  $\chi$  (defined in equation 5.2) and the time-varying cost of grid electricity  $C_{grid}$ . Analogous to the case of a single-building, the reward function when discharging is:

$$\mathcal{R}_{a^{(1)}} = \begin{cases} 2\chi + C_{grid}, & \chi \geq \delta_1 \\ \chi + 2C_{grid}, & \delta_1 > \chi \geq \delta_2 \\ 3C_{grid}, & \chi < \delta_2 \end{cases} \quad (5.6)$$

and when not discharging:

$$\mathcal{R}_{a^{(0)}} = \begin{cases} 1/(3\chi), & \chi \geq \delta_1 \\ 1/(\chi + 2C_{grid}), & \delta_1 > \chi \geq \delta_2 \\ 1/(3C_{grid}), & \chi < \delta_2, \end{cases} \quad (5.7)$$

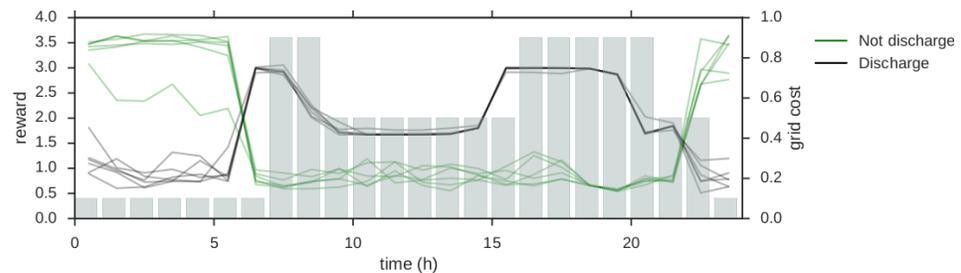
where  $\delta_1 = 0.8$  and  $\delta_2 = 0.7$ . In the discharging process specified in equation 5.6, bias towards  $C_{grid}$  increases as  $\chi$  decreases: when there is a peak in demand (high  $\chi$ ), the cost of the grid  $C_{grid}$  is less important, therefore the battery sees a greater reward the greater the peak if it chooses to discharge, independently on the cost of the grid imports  $C_{grid}$ , learning to alleviate peaks. On the other hand, when power demand is not as intense (low  $\chi$ ), the battery obtains greater reward for discharging the higher the grid cost  $C_{grid}$ . These processes are similar for the non-discharging reward, in equation 5.7, but in this case, rewards are inversely proportional to  $C_{grid}$  and  $\chi$  values.

As mentioned earlier, Q-learning is a flexible approach to model adaptable behaviours. In this case, a variety of options for the Equations (5.6) and (5.7) were implemented and tested, with different weights for  $\chi$  and  $C_{grid}$ , leading to different learning behaviours for the discharge process. Equations (5.6) and (5.7) were found to function satisfactorily for the battery.

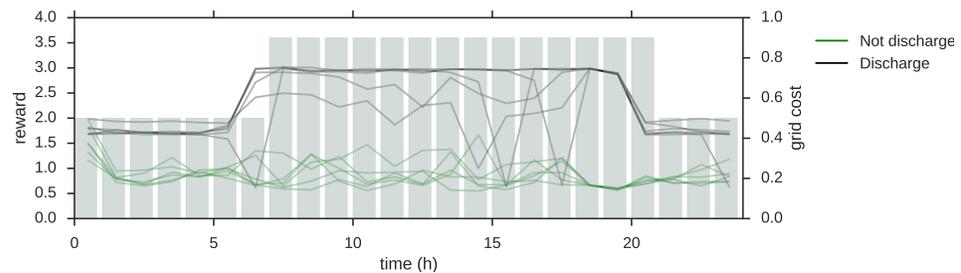
Figure 5.7a shows the selected grid tariff, and the consequent reward values for each of the batteries in the six buildings. The other two figures, 5.7b and 5.7c, are plotted here to show the effect of the tariff signal in the learning process of the battery. Without needing to define a new reward function, the discharge process shows sensitivity to the price signal, while taking into account the intensity of the

power demands. For the reversed tariff in Figure 5.7c, discharge during the day is not triggered, as it will be more advantageous to satisfy demands with grid imports and save stored energy for later.

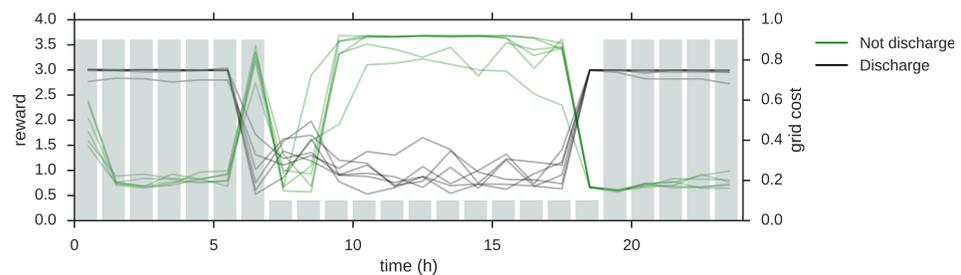
In all cases, the reward seems to be shifted with respect to the tariff values by one time step. This is because the effect of the action  $a_t$  taken at time step  $t$ , is evaluated at time step  $t + 1$ , but stored at  $t$ , when the action took place.



(a) Tariff I.



(b) Tariff II.



(c) Tariff III.

Figure 5.7: Reward associated to the discharge (black) or not-discharge (green) actions, taken by six batteries in six buildings. The shape of the curves depend on the tariff signal (in bars, on the right y-axis) and the demands in each case. Tariff I has been selected for the case study.

### 5.3.4 Results and discussion: Improving community self-consumption

In this section we present the simulation results for the 6-building neighbourhood detailed above, again considering three operation scenarios: the base case, the introduction of smart appliances in all houses and the additional introduction of electric storage, evaluated for a winter and a summer week. For each case a set of 100 week simulations<sup>9</sup> is considered, and the averaged results are presented here.

Figure 5.8 presents the weekly energy imports/exports from/to the grid. It shows results for winter (on the left) and summer (on the right), and for the three scenarios: base case on top, load shifting in the middle and load shifting plus battery in the bottom. The results are disaggregated for each building, with averaged values presented in table 5.1.

|                         | <b>Winter (kWh)</b> |          |          |           |
|-------------------------|---------------------|----------|----------|-----------|
|                         | on-peak             | mid-peak | off-peak | PV export |
| Base case               | 34.6                | 28.6     | 15.6     | 4.30      |
| Load shifting           | 27.6                | 14.9     | 25.1     | 7.60      |
| Load shifting & Battery | 22.3                | 12.5     | 25.4     | 0.013     |
|                         | <b>Summer (kWh)</b> |          |          |           |
|                         | on-peak             | mid-peak | off-peak | PV export |
| Base case               | 6.13                | 9.11     | 12.7     | 56.4      |
| Load shifting           | 3.14                | 4.79     | 13.1     | 59.6      |
| Load shifting & Battery | 0.31                | 0.689    | 11.8     | 50.2      |

Table 5.1: Average weekly import/export.

For the sake of comparison, the PV profile for the summer case corresponds to a bright and clear day, when solar energy is abundant. This is reflected in Figure 5.4, with large differences between the PV exports in summer and winter.

In winter, the top left graph shows that for the base case, grid imports occur

<sup>9</sup>The number of simulations considered takes into account when new simulations do not vary average values.

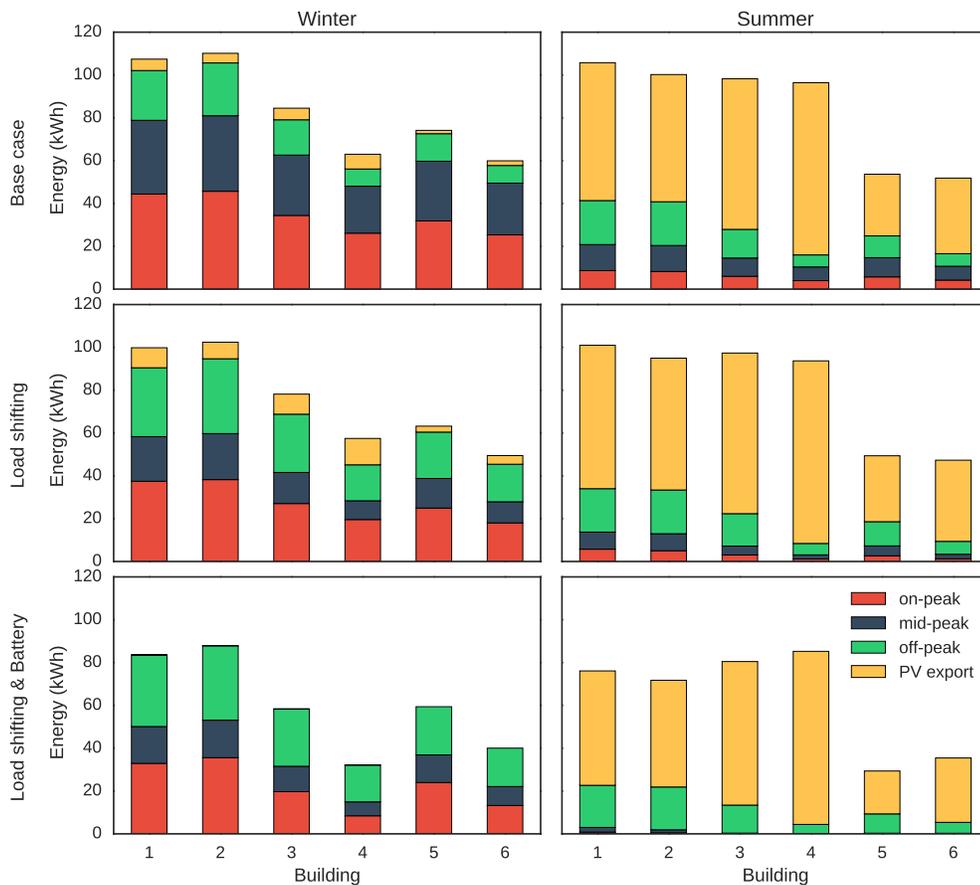


Figure 5.8: Imports and exports from the grid for winter (on the left) and summer (on the right), and for the three scenarios: base case on top, load shifting in the middle and load shifting plus battery in the bottom. Buildings 5 and 6 are those having 1.9kWp solar panels.

at peak and mid-peak times. The introduction of the load-shifting algorithm has two effects. Firstly, energy use is shifted to off-peak times. Secondly, PV exports increase a 77% when appliance re-scheduling takes place. This is due to the shifting process: smart appliances, learn that can be more beneficial to operate during cheap night time hours than during sunshine hours, as limited winter solar generation may not be enough to power the whole cycle, leading to expensive on-peak grid imports. When batteries are introduced in the community, practically no solar power is exported to the grid. In addition, total imports reduce to almost 24%,

with its main use being off-peak. The improvements in on-peak to off-peak shifting are maintained. This represents a success of the algorithm, as it achieves our two goals: alleviating energy intense periods and minimising high-cost grid imports.

In summer, the effect of the large amounts of PV available is clear, leading to a situation where grid imports are a third of the winter value. Also, total demands are lower, due to heating being switched off. Another consequence is that shifting appliances learn to switch to sunshine hours (slight decrease in PV exports) as well as to off-peak night periods (increase in off-peak imports). On average, on-peak and mid-peak imports are reduced over by 48% with load shifting alone. Again, a significant impact is achieved with the introduction of storage systems: local generation exports are cut by a 11%, and imports are shifted to off-grid periods even more, with a reduction over 93% for on- and mid-peak periods.

Results for summer and winter show that the reward functions used work well in shifting demand from expensive grid imports towards locally generated energy. This has been achieved by imposing a grid import tariff, which proves to be successful for our purposes. A powerful potential of the Q-learning technique used is that we could try a different price structure, and the system would adapt to make the necessary changes to achieve a relatively optimised operation.

Figure 5.9 shows the boxplots for the distribution of self-consumption for each simulation of each building. In this case, we differentiate SC at two levels: local and neighbourhood. They relate to the energy use arising from local resources (their own PV/battery system) and from their neighbours' resources (PV/battery systems owned by neighbours), representing individual and collective self-consumption. Mean values for the different groups are given in Table 5.2.

Winter values for individual SC in the base case average to 13.6%. This figure

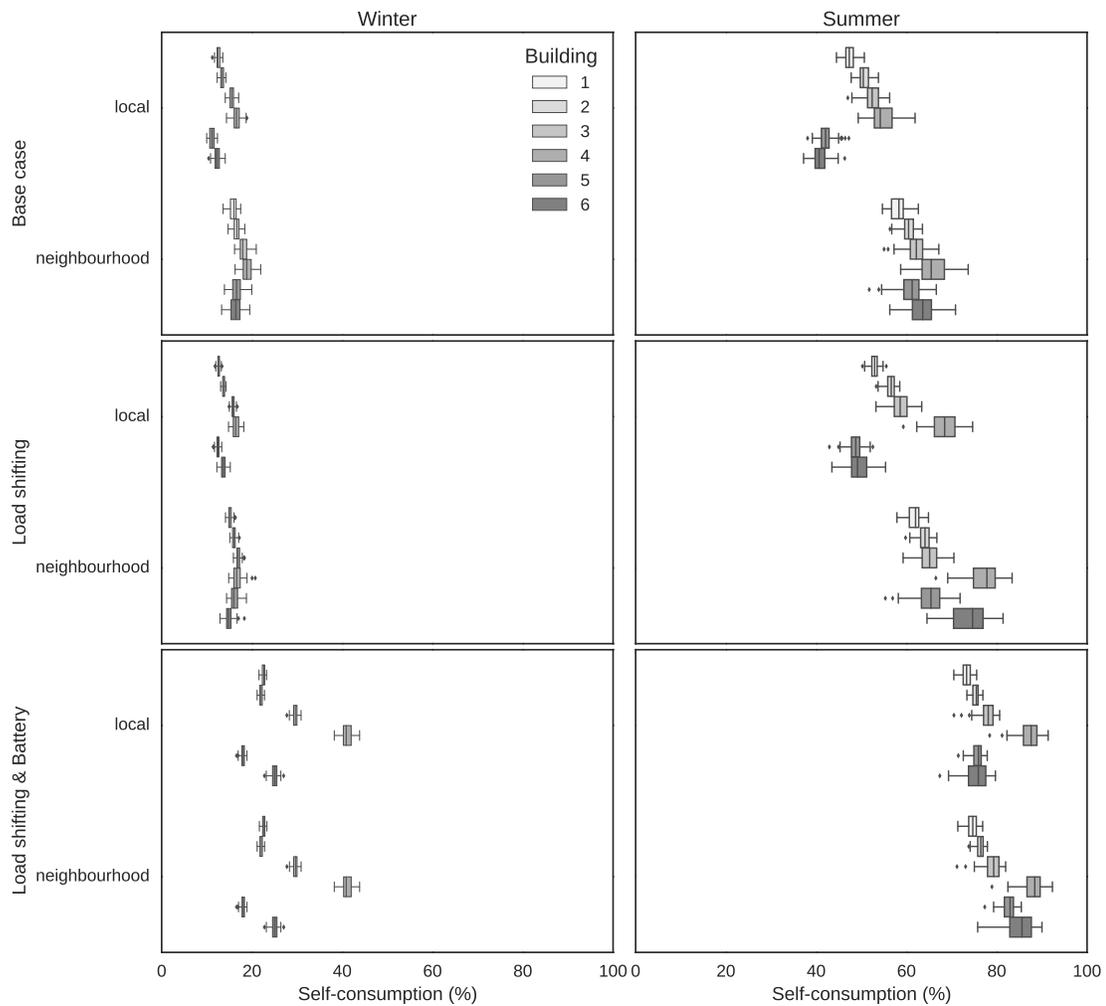


Figure 5.9: Self-consumption index for the three scenarios, in winter and summer.

increases to 17.1% for the neighbourhood case. For the reasons explained above, there are not major changes in local and neighbourhood SC when the load shifting algorithm is applied: local SC is slightly larger and collective SC is slightly smaller, although buildings 5 and 6 are benefited from their neighbours resources. One way to improve SC more dramatically with load shifting would be to consider a more expensive off-peak tariff, so that appliances learn to shift to sunshine hours, even if grid imports are necessary. As expected, the introduction of batteries has a significant impact, doubling the value of local SC. On the other hand, individual

|                         |        | <b>Winter</b> | <b>Summer</b> |
|-------------------------|--------|---------------|---------------|
| Base case               | Local  | 13.6          | 48.0          |
|                         | Neigh. | 17.1          | 61.9          |
| Load shifting           | Local  | 14.1          | 55.7          |
|                         | Neigh. | 16.0          | 67.8          |
| Load shifting & Battery | Local  | 26.4          | 77.5          |
|                         | Neigh. | 26.4          | 81.0          |

Table 5.2: Average self-consumption (%).

batteries available for all buildings neutralizes the positive effect of collective use, as the excess of solar energy is individually stored for later use instead of shared. Allowing the batteries to discharge to other houses' appliances could be useful for improving neighbourhood SC, given an appropriate business model.

In summer, SC increases in all cases. When the load shifting algorithm is in use, local SC increases from 48.0% to 55.7%, and neighbourhood SC from 61.9% to 67.8%. Maximum levels of SC are achieved with the storage systems in place during summer, reaching values of over 75%. With respect to the base case, this represents a 61% and a 31% increase in individual and collective SC respectively.

In both summer and winter, households achieve larger values in SC when allowed to share their resources (except in winter battery case, that same results are obtained). This suggests the need for a new way of looking at tariffs. Different pricing systems could be put into practise, which regard not only the interaction with the electricity grid of each individual customer, but also the interaction between customers, creating for example some sort of incentive structure to share power resources between members of the same neighbourhood. For example, in our simulation, households 5 and 6 made a smaller investment in their PV system, reducing their local self consumption. However, the base case and the load shifting scenarios in Figure 5.9 show how strongly they benefit from participating in the

community, with comparable collective SC values.

As it was mentioned above, SC is not explicitly maximised in our DR formulation, but it is indirectly affected by the price signal considered. Another way to evaluate the effectiveness of the Q-learning algorithm is looking at the cost reduction for the different scenarios. Figure 5.10 shows the total weekly cost of running appliances and devices. It shows a boxplot for each scenario with averaged cost for the 6 buildings. In both cases, for winter and summer, the algorithm predicts

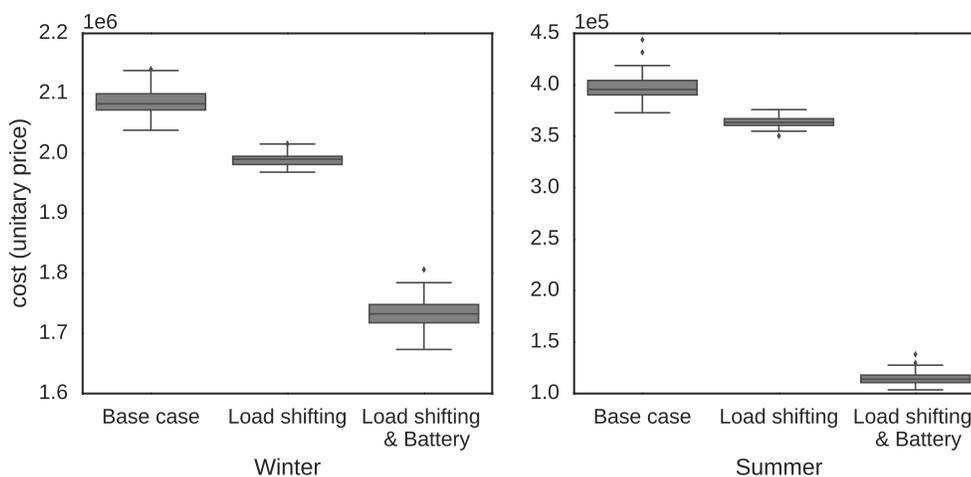


Figure 5.10: Cost reduction when load shifting and battery are introduced. Values based on unit tariff signal proposed.

reductions in cost. In the case of summer, when there is more PV energy available, costs are cut in half. The graph allows for comparison between the scenarios, but given that the cost is calculated using a normalized tariff signal, it is not possible to understand it as monetary price.

## 5.4 Conclusion

In this chapter, we have demonstrated the integration of price-based Demand Response strategies to optimise for load-shifting and the charging and discharging of

a battery, by incorporating learning abilities into our device agents. Q-learning has been proven to be a successful candidate for that task. Different reward functions have been explained and put into practise. Single and multi-building case studies have been presented; the latter requiring an extended negotiation logic.

More specifically, the learning algorithms have effectively: shifted demand from on- and mid- peak periods to off-peak periods, particularly when combining shiftable demand devices with battery storage; increased remarkably the self-consumption percentage (the fraction of energy demand that is satisfied by on-site generation). The multi-building approach demonstrates the positive effects in SC when communities of buildings are able to share their resources.

Regarding the Q-learning approach, it is important to note that the current implementation used in this thesis is unable to respond to irregular events (e.g. weather) or to act well with extreme values caused by particular situations. Being able to handle such circumstances would be possible if specific “extreme” state-action pairs are defined.

Although we believe that we have successfully demonstrated this proof-of-concept there is considerable scope for improvement to this framework. In particular with respect to (and still only considering requirements R1 to R3):

1. The battery model should be improved to model state-of-charge dependent losses.
2. In this work, battery and heating demand have been combined to reduce on-peak grid imports. In a more realistic situation, the heating system can be considered as an additional learning device in its own right, that contributes to the global goals.

3. The negotiation for the multi-building system should be upgraded to consider battery discharge to appliances in neighbouring households. Battery agents would need to take into consideration not only the operation status of their own local system (household), but also the global system (other households). Different reward structures would be necessary. Another upgrade could be to allow the batteries to charge from the grid. This could be achieved by providing the battery agents with the ability to learn charging strategies.
4. For islanded scenarios in which power cannot be drawn upstream of the mains network, a formal bidding mechanism needs to be integrated, favouring devices of relatively high priority when reserves are limited (e.g. emergency lighting, or freezers that have undergone a long delay since the compressor was last enabled).
5. Alternative tariff structures should be explored that better favour self-consumption.

Looking further to the future (R4 to R6):

6. No-MASS/DR should be generalised to solve for multiple buildings interconnected via our LV network model. That way, the above tariffs might also consider local network integrity, when the LV model is utilised as intended. The next chapter follows this premise.
7. Empirical evidence from DR field trials should be used to predict the extent to which users are willing to: *a)* devolve control to autonomous devices, and *b)* adjust their behaviours in response to information (e.g. tariff and/or CO<sub>2</sub> emissions) feedback.

Finally, field trials to test the validity of this new more general multi-agent stochastic simulation framework would be of considerable value.



## Chapter 6

# Low Voltage Network modelling

Previous chapters have elaborated on methodologies for describing energy use in homes, and potential efficiency improvements through DR, included application to a small energy community. An explicit model of the Low Voltage (LV) network has been developed which can be coupled with the DR framework. It could potentially enable the evaluation of network operation effects due to the integration of renewable sources; it could also be added as a control variable for the DR algorithms. In this chapter, the proposed model for power-flow analysis of a general low-voltage distribution network is presented, which uses an electrical circuit-based approach, implemented as a novel recursive algorithm, and can efficiently calculate the voltages at different nodes of a complicated branched network.

### 6.1 Introduction

The goal of the electricity transmission and distribution system is to efficiently and stably transport energy from a generation site to consumers. To do so, they use a network of networks, operating at different voltage levels (see Figure 6.1). At the highest level, the national high-voltage transmission network works at 400 or 275kV (using higher voltages results in lower transmission and distribution losses). In the UK, this network is owned and maintained by Transmission System Oper-

ators (TSO): National Grid for England and Wales, Scottish Power and Scottish Hydro-Electric in Scotland. The transmission network connects power stations and major substations, and has connections with France, the Netherlands and Northern Ireland. It feeds Grid Supply Points and major substations that step down the voltage [78], 400/132kV or 275/132kV, to the regional distribution systems. The 132kV circuit feeds the Bulk Supply Points (BSP), which step the voltage down to 66kV or 33kV. The primary distribution circuit, operating at 66 or 33kV, is connected to primary substations of 66/11kV or 33/11kV, leading to the secondary circuit. The 11kV network is radially connected to a set of secondary substations that transform voltage to the Low Voltage (LV) network, usually 400V<sup>1</sup>.

The LV network is generally arranged as multi branched radial feeders and consists of underground cables and overhead lines (although underground cable are used for new connections). According to [78], a maximum of 100 customers are connected to each LV circuit. As mentioned, individual domestic users have a single phase supply, and these connections are evenly distributed across the 3 phases, in order to avoid creating an unbalanced system. Blocks of flats have a three-phase service installed centrally and lateral connections are provided to flats. Commercial and industrial buildings are supplied with a three-phase service.

Correct operation of the power system involves engagement from different entities. A list of the actors taking part in the transmission and distribution grid is presented in Table 6.1; together with the roles they play and the main services they provide. Such services describe the interconnections among the different actors, configuring a complex system of technical, economic and regulatory bodies. Table 6.1 incorporates an *independent aggregator*, as part of a future smart grid formula-

---

<sup>1</sup>Note that these are line-line voltages for a 3-phase distribution system, and the line-neutral commonly seen at household level is  $400\sqrt{3}$  i.e. 230V

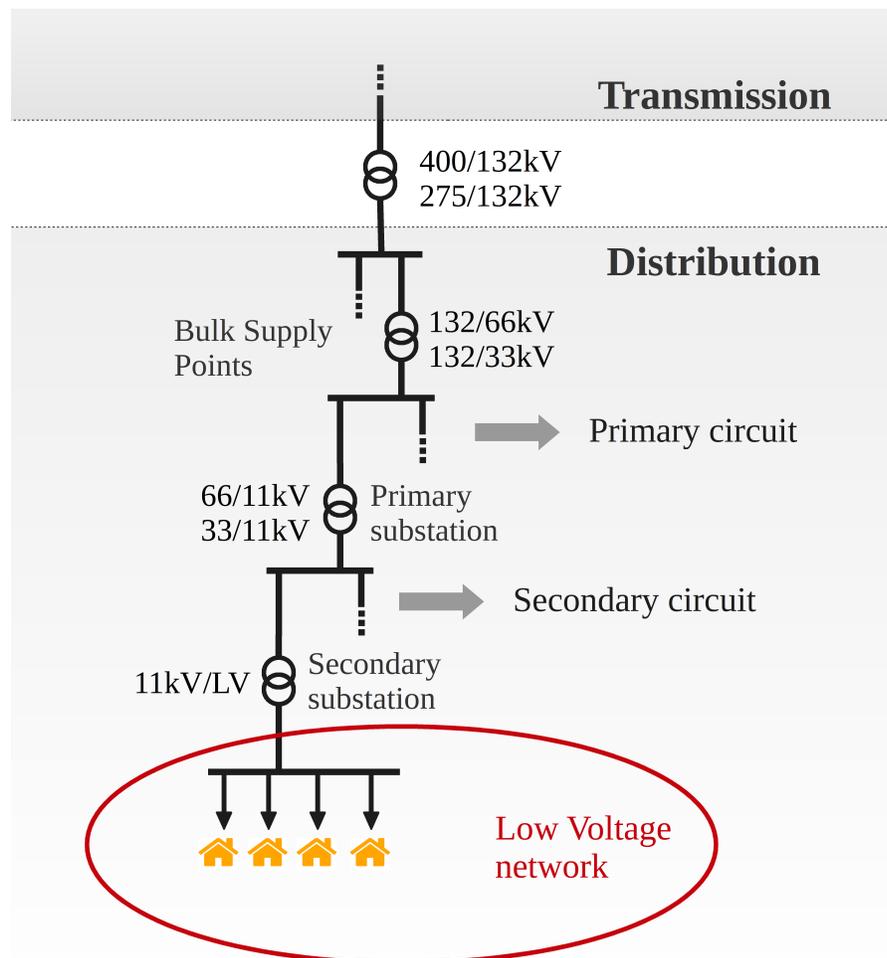


Figure 6.1: Power system

tion, whose function is to facilitate flexible demand for consumers and prosumers (producing-consumers).

### The Distribution System and Demand Response

Distribution networks are gaining increasing relevance in the energy transition landscape. They make up to 90% of the total electricity system network length [35] and a large percentage of all electrical demand, and they accommodate a growing penetration of Distributed Generation (DG) and Renewable Energy Sys-

| ACTOR  | ROLE                 | MAIN ASSOCIATED SERVICES  |
|--|----------------------|---|
| <b>Transmission System Operator (TSO)</b><br>Responsible for provision of infrastructure, information and operation on the main high voltage electric networks.  | Data provider        | Sharing market information in a transparent way.  |
|  | Grid operator        | Transmission grid maintenance and expansion through investments, while securing operation.  |
|  | Grid access provider | Stablish conditions of fair access to parties connected; administration and maintenance of access points.   |
|  | System operator      | Responsible for stable operation; definition of technical requirement; data exchange for forecast; capacity allocation and congestion management; maintain balance through frequency control. |
|  | Market operator*     | Setting market energy price and imbalance price; reserve allocation.  |
| <b>Distribution System Operator (DSO)</b><br>Responsible for operating, maintaining and developing the distribution system, and for ensuring long-term ability for the system to meet electricity demands. | Grid operator        | Distribution grid maintenance and expansion through investments, while securing operation.  |
|  | Grid access provider | Connecting between producers and consumers.   |
|  | System operator*     | Responsible for stable operation; definition of technical requirement; data exchange for forecast; capacity allocation and congestion management; maintain balance through frequency control. |
|  | Market operator*     | With increasing DG and DR, DSOs gain new roles in balancing local markets.  |
|  | Meter responsible*   | If not held by an independent aggregator, responsible for meter operation and data collection.  |
| <b>Independent aggregator</b><br>New actor needed in smart grids. Coordinates between market and grid, administrating flexibility.   | BSP <sup>a</sup>     | Provides balancing services to TSO, as consequence of differences between metered consumption / generation and actual bought / sold electricity.  |
|  | Meter responsible    | Sends signals to customers and demand changes to TSO.   |
|  | Party connected      | Provides flexibility to customers connected.  |
| <b>Supplier, retailer, trader</b><br>Contractors for residential customers to buy and sell electrical energy. Manage generation plants and demand acquisition for retailers.                               | BRP <sup>b</sup>     | Financial and legal responsible for imbalances between nominated and consumed / generated electricity.  |
|  | BSP                  | Provide balancing services to TSO.  |
|  | Resource provider    | Providing raw materials and electricity generation; maintenance of power plant.   |
|  | Party connected*     | Transfer of electricity to the grid at accounting point.  |
|  | Meter responsible*   | Administrative aspects of electricity supply to customers.  |
|  | Technology provider* | Responsible for setting price to customers: high influence on the success of DR.  |
| <b>Regulatory authority</b><br>Independent institution to develop competitive markets.   | Control function     | Guaranteeing a secure, cost-effective, efficient and customer oriented system, regulating decisions applied by grid, market and system operators.   |
| <b>Customer</b><br>Consumers and prosumers, becoming a more active actor.  | Party connected      | Receive electricity they need; can provide flexibility through DR.  |

Table 6.1: Actors and roles in the transmission and distribution power system. Sources: [34, 79].

\*Under certain circumstances.

<sup>a</sup>Balancing Service Provider

<sup>b</sup>Balancing Responsible Party

tems (RES). As discussed in previous chapters, DR is proposed as a beneficial and affordable technology to help address the challenges of intermittent supply.

As suggested in [34], DR use cases appear at different parts of the distribution system and at different scales, ranging from individual households (nano level), residential areas (micro level: LV-microgrid), operational areas of DSOs (meso level: MV-grids), through to nationwide areas (macro level: HV-grid). Thus, the value of DR and DG-RES applications is seen by all the actors involved in the electricity system at the different scales.

- For the **markets**: introduction of flexibility aggregators could raise competition, especially in real-time markets, for example for imbalance settlement.
- For the **TSO**: the cost and volume of generation/consumption that needs to be balanced depends on the Balancing Responsible Parties (BRP's). DR allows BRP's to balance themselves, requiring less intervention from the TSO. On the other hand, load-shifting can effectively increase available capacity, reducing grid investment costs.
- **DSOs** can also benefit from reduced grid investment costs due to lower peak loads and improved security of supply by means of DR. Moreover, DSOs can cut down costs due to network losses, as these are also dependent on peak loads.
- **Customers** are offered the possibility to participate actively in their energy use, becoming aware of their electricity consumption, and engaging with efficiency measures, lowering their electricity bill. Also, the uptake of DG-RES technology allows users to become less grid-dependent, reducing capacity costs.

## 6.2 LV network modelling

The modelling described in this chapter is concerned with the domestic scale, at individual and community levels. In a power system, this corresponds to the LV network, the circuits streaming down from the secondary substation.

Previous chapters described a methodology for DR simulation, modifying local demand and generation patterns. This variation has an impact on the electricity grid. At the same time, the performance of the grid should shape part of the DR solutions applied. For instance, voltage rising over statutory limits may trigger DG curtailment. For this reason, we consider it very important for a DR simulation software to have the means to describe the operation of the network. Ideally, the operational state could be fed back into the optimisation procedure of the DR.

To achieve this a load-flow algorithm has been developed and coded, that can be used as stand-alone software for power flow simulations or integrated within the No-MASS/DR framework. This load flow model can calculate the voltage at all of the nodes of a complex network. For the No-MASS case, the load-flow simulation is performed at each time step as a post-process, using demand and supply power values through the network (calculated using the corresponding models). This means that resulting operation is not fed back into the simulation, but only saved for later analysis. Future forms of No-MASS/DR will integrate in-the-loop functionalities for the LV network.

The rest of the Chapter is organised as follows. The next section contains a review of major load-flow analysis methods, and the requirements for temporal granularity based on the electrical phenomena studied. Then the proposed algorithm is presented, based on recursive circuit analysis, and its applicability in a multi-building network, containing over 120 homes is demonstrated. The chapter

ends with remarks and conclusion.

### 6.3 Load-flow analysis

Power-flow or load-flow analysis is a well known problem that consists of determining the steady state behaviour of a power system, finding steady state voltage and current values throughout a network. Power flow is a fundamental calculation for the analysis of any power system, largely used for planning and operation applications.

The power-flow problem is based on Kirchoff's Laws, by which the sum of all the currents or powers flowing in every node of the network are equal to zero:

$$\sum_{k=1}^m I_k = 0, \quad (6.1)$$

for a node with  $m$  branches. When power-flow equations are expressed in terms of power magnitudes (as the loads are considered to be constant power loads), it becomes a non-linear problem. For a single line distribution network with  $n$  nodes, modelled as an electrical circuit (see Figure 6.2), the load currents  $\mathbf{I} = [I_1, \dots, I_n]$  are written as

$$\begin{bmatrix} I_1 \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} Y_{11} & \dots & Y_{1n} \\ \vdots & \ddots & \vdots \\ Y_{n1} & \dots & Y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}, \quad (6.2)$$

for which each of the terms  $Y_{kj}$  of the admittance matrix  $[\mathbf{Y}]$  relates to the line impedances between nodes  $j$  and  $k$  as  $Y_{kj} = \frac{1}{Z_{kj}}$ . When a power load is connected

to the  $i$  node, load current is given by

$$I_i = \left( \frac{S_i}{V_i} \right)^*, \quad (6.3)$$

considering the following notation:

|              |   |
|--------------|---|
| $V_j$        | node voltage.   |
| $I_{line,j}$ | current between nodes $j$ and $j - 1$ .                       |
| $I_j$        | current flowing towards/from load/source.                     |
| $S_j$        | apparent power <sup>2</sup> flowing towards/from load/source. |
| $Z_j$        | cable impedance between nodes $j$ and $j - 1$ .               |

Combining equations 6.2 and 6.3, the power-flow problem is expressed as a set of  $2n$  non-linear equations (for complex magnitudes have real and imaginary parts). Solving this problem, thus, requires approximate numerical solutions.

The first algorithms to solve power flow appeared with the introduction of digital computers, with initial automatic digital solution dating from 1956 [80]. Developments for industry in load-flow analysis led mainly to studies in transmission networks. They usually have parallel lines in a meshed structure with multiple redundant paths from generation to load areas. In these types of networks, traditional numerical methods, such as the Newton-Raphson methods and its variants became popular, as they had powerful convergence properties [81]. These techniques rely on an admittance matrix (in equation 6.2) inversion process, and therefore, they were specifically developed towards efficient processing of large matrices.

---

<sup>2</sup>Real and reactive components.

However, the distribution network generally presents a very different structure, with a weakly meshed radial topology and highly resistive impedances (high  $R/X$  ratio). This leads to a sparse (or even ill-conditioned<sup>3</sup>) admittance matrix and the traditional methods (Newton-Raphson, Gauss-Seidel) become inefficient, unstable or divergent [82, 83]. Alternative algorithms have been proposed to deal specifically with distribution networks, the most popular ones being forward/backward sweep methods (also called ladder network methods) that use basic circuit theories to calculate power-flow. Their basic operation process for a network with  $n$  nodes, using the notation specified above, works as follows:

1. Terminal node voltages  $V_n$  are approximated to the slack bus voltage  $V_s$ :

$$V_n = V_s, \text{ if } j \text{ is the terminal node.} \quad (6.4)$$

2. Terminal load and line currents are calculated using equations 6.1 and 6.3:

$$I_{load,n} = \left( \frac{S_n}{V_n} \right)^* \quad (6.5)$$

$$I_{line,n} = I_{load,n} \quad (6.6)$$

3. From there, all voltages, and currents are subsequently calculated towards the slack bus:

$$V_j = V_{j+1} + I_{line,j+1} Z_{j+1} \quad (6.7)$$

$$I_{load,j} = \left( \frac{S_j}{V_j} \right)^* \quad (6.8)$$

$$I_{line,j} = I_{load,j} + \sum_{k \neq j} I_{line,k} \quad (6.9)$$

---

<sup>3</sup>In ill-conditioned systems the output is highly dependent on the input arguments.

4. Calculated and known slack voltages  $V_0$  and  $V_s$  are compared with a preset convergence tolerance value. If the difference is smaller than the selected tolerance, the algorithm stops: convergence has been achieved. Otherwise, node voltages are recalculated during the backward sweep, taking the known voltage value  $V_s$  instead of  $V_0$ , from the slack bus towards terminal nodes:

$$V_j = V_{j-1} + I_{line j} Z_j \quad (6.10)$$

5. The process continues until  $V_s$  and  $V_0$  differ by less than the given tolerance.

The advantages of these methods include: they have a simple formulation, are less sensitive to  $R/X$  ratio and they are robust to heavy loads, reflecting the dependency of the node voltage on the load level. The main limitation of these methods is that they are not suitable for meshed layouts in their fundamental formulation, because they consider there is a unique path from any given bus to the source. Extensions to them have been proposed to overcome this problem, for instance, using a combination of ladder network and Newton-Raphson analysis or a compensation method [82].

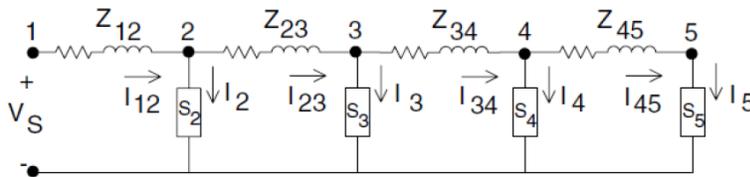


Figure 6.2: Description of circuit-based methods. Source [83].

For this thesis, a ladder network approach has been implemented to represent a low voltage distribution network.

### 6.3.1 Time granularity for load-flow simulation

Power grid dynamics occur at a large span of temporal scales (Figure 6.3), from the  $\mu s$  interval of operation of solid-state switching devices [84] to the tens of years in which long-term strategic planning takes place. In this range, myriad physical phenomena are influenced, where each needs to be analysed and understood for a safe and stable operation of the power grid.

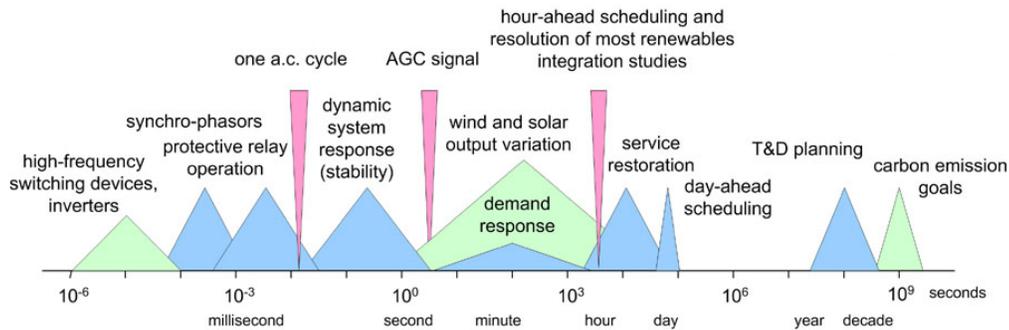


Figure 6.3: Power grid time-scales. Source: [84].

When simulating power-flow in a network, the selected time granularity should balance the required description of the system dynamics and computing weight [85]. Load-flow analysis can be used for studies in energy balance, short-circuit fault analysis, transient stability studies or economic dispatch, and each of them is characterised with a specific time scale. Moreover, selecting the wrong time-step for simulation could lead to major errors in the calculations.

In this work, the focus is on energy analysis more than in network system integrity, which dictates a target range of candidate temporal resolutions. Several studies [85–87] on the effects of time-resolution for energy modelling and renewable energy integration propose one minute resolution as a good trade-off between computing weight and an accurate description of the energy flows in the network. A one minute time step simulation allows the calculation of voltage and power values

through the nodes; higher resolution phenomena such as transients or harmonics are out of the scope of this project.

It is worth mentioning that the temporal granularity for power flow analysis is also constrained by the resolution of the models and data available for loads and generation, which in turn is dependent on the time resolution of the data available to build such models.

## 6.4 Load-flow analysis as a recursion algorithm

A forward/backward sweep solver has been used for power-flow analysis. It has been implemented using Object Oriented Programming (OOP) and more specifically, as a recursive algorithm. The details of the model and the algorithm are explained in this section.

### 6.4.1 The model: pre-requisites and assumptions

As explained in Section 6.3 there are multiple techniques and solving methods when addressing the load-flow problem. The selection of one method over another is highly dependent on the characteristics of the network under consideration, and on the particular objectives of the modelling process.

When selecting the appropriate method, there were two main modelling objectives:

1. Describe a residential distribution network (potentially around 100 houses), which typically follows a radial layout. The aim is not to solve for a single specific layout, but to solve a general case.

2. The availability of source code, as the aim is to implement the selected load-flow algorithm with the building simulation software. This remark makes the use of existing power systems software more challenging, especially for those which are not open source.

A single line distribution network is modelled as an electrical circuit (see Figure 6.2), and solved using an iterative forward/backward algorithm, following the method describe in previous section 6.3. Apparent power flows  $S_k$  to/from loads/sources are represented as ideal current sources, and the cable losses are entered with values for cable impedance.

Finding a solution for the power flow requires that one nodal voltage (slack bus)  $V_s$  is known and constant during the time-step simulation, acting as a reference value for the iterations. It corresponds to the feeder in the network, represented as an ideal voltage source, which connects the LV area with the rest of the grid.

### 6.4.2 The load flow algorithm

The forward/backward sweep algorithm has been implemented as a recursive algorithm. In Computer Science, *recursion* is a method employed when the solution to a problem requires the solution of smaller instances of the same problem. In practice, it is implemented by defining functions that are allowed to call themselves. In the case of a distribution network, recursion constitutes a technique to efficiently solve branched networks, as it eliminates the need for any iterative loop.

During the forward sweep, the currents and voltages at each node depend on the currents and voltages at the child nodes connected to them (one if a linear network, several if branched); whereas during the backward sweep, currents and voltages depend on the parent node (see figure 6.4 for clarification).

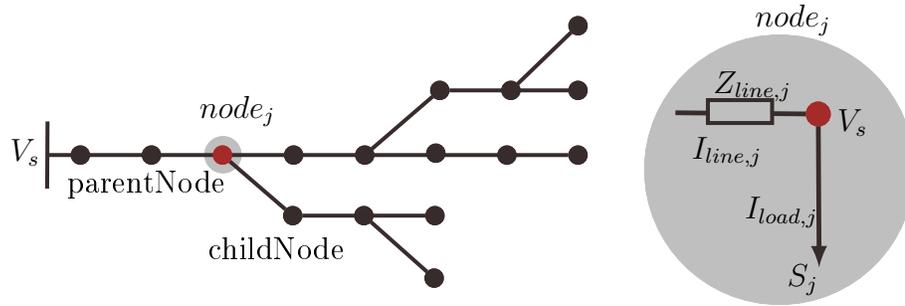


Figure 6.4: Diagram of network, relating parent and child nodes. An object  $node_j$  has attributes: voltage  $V_j$ , complex power (drawn or injected)  $S_j$ , load current  $I_{load,j}$ , line current  $I_{line,j}$  towards the parent node and cable impedance  $Z_{line,j}$  through this line.

A Node object (figure 6.4) has five variables attached, two of them are input arguments given to the algorithm:  $V_j$ ,  $I_j$ ,  $I_{load,j}$ ,  $S_j$  (input variable) and  $Z_j$  (input variable), following the definitions used in the section above. Defining  $Z_{line,j}$  as part of the node object may seem confusing, but it is required as the network has to be a series of identical node units. Power flow analysis is carried out in order to determine  $V_j$ ,  $I_{load,j}$ , and  $I_{line,j}$  throughout the network. The pseudo code description of the implementation is presented in Section 6.4.3.

This approach presents some advantages:

- The main feature of implementing the forward/backward loop as a recursive algorithm is that it provides a very efficient procedure for analysing branched layouts of the network. Using *for-loops* instead of recursion would make the process of defining an arbitrary multi-branched network and attaching child nodes to parent nodes extremely tedious. With recursion, this process is simple and fast for any multi-branched (not meshed) network.
- Evaluation of complex magnitudes. If values of active/reactive power and resistance/reactance are input into the model, the algorithm will determine the magnitude and phases of all variables.

- Flexibility to define specific line impedance values between nodes. As the nodes are defined individually, it is trivial to simulate a network with different impedance values between houses (due to different lengths or cables).
- Nodes can represent specific devices or aggregated loads (at household level, for example), thanks to the flexible line impedance (see bullet point above). This property allows for the impact that a single device (either an appliance, a generator or a storage device) has in the network to be evaluated.

Although the method proposed has been proved to fit the requirements and modelling objectives set for this work, this approach has limitations: the algorithm so far is not able to deal with meshed layouts; other elements of the power system and their effect on the network are ignored (such as transformers, converters, etc.); it cannot consider higher temporal resolution phenomena occurring in the network (such as transients).

### 6.4.3 Pseudo-code

The pseudo code of the implementation is presented in algorithms 1, 2 and 3. Algorithms 1 and 2 present the forward and backward sweep, respectively; whereas Algorithm 3 is the code run for converging to a solution. In order to set the *tolerance* value in Algorithm 3, a sensitivity analysis was carried out, and the selected value corresponds to the one for which the converged solution does not vary when decreasing the tolerance.

---

```

1: for  $\forall node_k$  do
2:   ForwardSweep( $node_k$ ) ▷ Use of recursion
3: end for
4:
5: if  $node_j$  is terminal and it is first iteration then
6:    $V_j = V_0$  ▷ Assume  $V_0$  at terminal nodes.
7: else
8:    $V_j = V_{j+1} + I_{linej+1} \times Z_{j+1}$ 
9: end if
10:  $I_{loadj} = \left(\frac{S_j}{V_j}\right)^*$ 
11:  $I_{linej} = I_{loadj} + \sum_{k \neq j} I_{linek}$ 

```

---

Algorithm 1: FORWARDSWEEP()

---

```

1:  $V_j = V_{j-1} - I_{linej} \times Z_j$ 
2: for  $\forall node_k$  do
3:   BackwardSweep( $V_j$ ) ▷ Use of recursion
4: end for

```

---

Algorithm 2: BACKWARDSWEEP( $V_{j-1}$ )

---

```

1: while  $Error = \|V_0 - V_{slack}\| > tolerance$  do
2:   ForwardSweep()
3:   BackwardSweep()
4: end while

```

---

Algorithm 3: POWERFLOW

## 6.5 Application to a case study

The proposed algorithm can deal with any branched network (without mesh). To test its effectiveness, a portion of a real low voltage network has been selected, situated in the area of the Meadows, in Nottingham (Figure 6.5). An approximated topology of four branches coming from the Wilford Crescent East Meadows secondary substation is selected. The four branches power 23, 54, 24 and 22 domestic buildings, respectively, coming to a total of 123 homes.

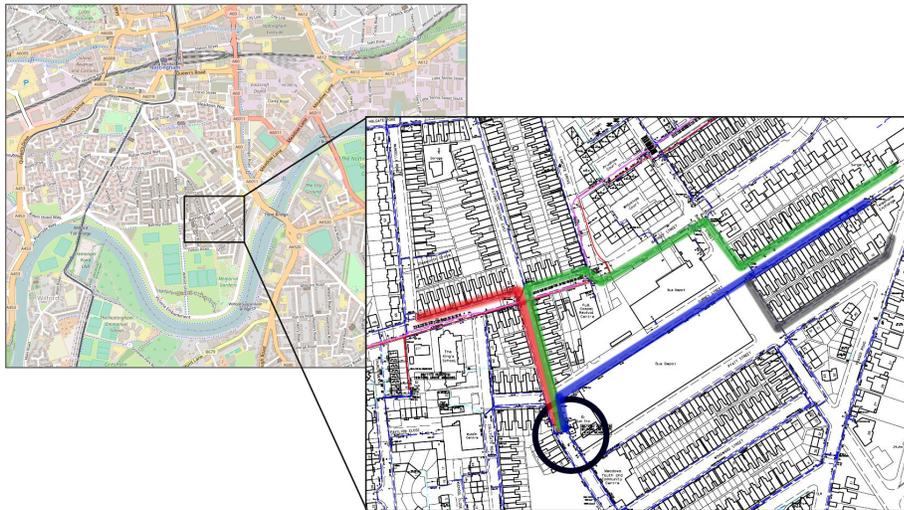


Figure 6.5: Meadows network layout.

A network scheme is presented in Figure 6.6. The 123 nodes are spread into branches A, B, C and D. The two latter ones connect buildings after a long cable stretch, that will be reflected in the power flow simulation. Each of the nodes, corresponds to the power load coming from a single building. Power demands coming from appliances and heating are calculated using No-MASS, as described in the previous Chapter 5. In this case, No-MASS/DR is used as a tool to generate demand profiles, including also the capabilities for load shifting. The stochastic energy profiles output from No-MASS/DR are then input to the LV network model.

PV technology has been deployed for all the households, with 80% of them owning a 3.8kW-peak system and the remaining 20% a 1.9kW-peak system. As in the previous Chapter, it is out of the scope of this thesis to develop models for solar generation, even though the platform would benefit from having them; this will be addressed in the future. Until then, solar power generation data is used as the generation sources.

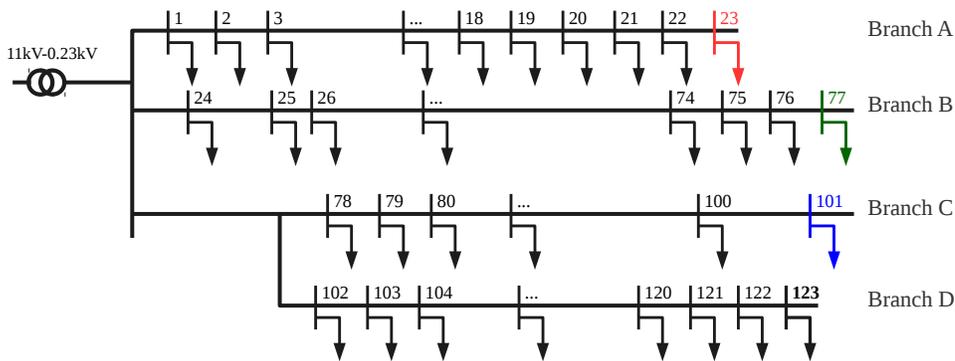


Figure 6.6: Network scheme.

The algorithm is flexible for defining cable impedance values, which in turn depend on the characteristics (material, configuration) and dimensions (cross section and length) of the cable. For this work, cable impedance is roughly approximated to  $0.15394 \Omega/km$ , following [88], considering a  $185mm^2$  aluminium core LV cable (cable reactance assumed to be negligible). For more detailed calculations, the values of the exact type of cable in the area can be input into the load-flow algorithm. Using Open Maps and Q-GIS software, the approximated distance between the buildings is measured and specified as given in Table 6.2.

With this configuration, a 24h load flow simulation for a winter day and a summer day has been carried out, with 1-minute time resolution. The results are presented in Figure 6.7, showing the voltages at steady-state for the terminal nodes

| Node      | $L_A$ (m) | $L_B$ (m) | $L_C$ (m) | $L_D$ (m)     | Node      | $L_B$ (m) cont. |
|-----------|-----------|-----------|-----------|---------------|-----------|-----------------|
| <b>1</b>  | 69.0      | 39.7      | 287.0     | 53.5 (+287.0) | <b>28</b> | 6.03            |
| <b>2</b>  | 4.35      | 5.93      | 6.54      | 6.26          | <b>29</b> | 6.03            |
| <b>3</b>  | 4.35      | 4.06      | 6.54      | 6.26          | <b>30</b> | 6.03            |
| <b>4</b>  | 4.35      | 4.06      | 6.54      | 6.26          | <b>31</b> | 6.03            |
| <b>5</b>  | 4.35      | 4.06      | 6.54      | 6.26          | <b>32</b> | 6.03            |
| <b>6</b>  | 4.35      | 4.06      | 6.54      | 6.26          | <b>33</b> | 6.03            |
| <b>7</b>  | 31.9      | 4.06      | 6.54      | 6.26          | <b>34</b> | 6.03            |
| <b>8</b>  | 5.16      | 4.06      | 6.54      | 6.26          | <b>35</b> | 6.03            |
| <b>9</b>  | 6.41      | 4.06      | 6.54      | 6.26          | <b>36</b> | 6.03            |
| <b>10</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>37</b> | 6.03            |
| <b>11</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>38</b> | 6.03            |
| <b>12</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>39</b> | 6.03            |
| <b>13</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>40</b> | 6.03            |
| <b>14</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>41</b> | 6.03            |
| <b>15</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>42</b> | 6.03            |
| <b>16</b> | 6.41      | 4.06      | 6.54      | 6.26          | <b>43</b> | 6.03            |
| <b>17</b> | 6.41      | 18.8      | 6.54      | 6.26          | <b>44</b> | 6.03            |
| <b>18</b> | 6.41      | 6.11      | 6.54      | 6.26          | <b>45</b> | 6.03            |
| <b>19</b> | 6.41      | 6.11      | 6.54      | 6.26          | <b>46</b> | 6.03            |
| <b>20</b> | 6.41      | 6.11      | 6.54      | 6.26          | <b>47</b> | 6.03            |
| <b>21</b> | 6.41      | 6.11      | 6.54      | 12.5          | <b>48</b> | 6.03            |
| <b>22</b> | 6.41      | 6.11      | 6.54      | 15.0          | <b>49</b> | 6.03            |
| <b>23</b> | 6.41      | 6.11      | 6.54      | -             | <b>50</b> | 6.03            |
| <b>24</b> | -         | 219.9     | 6.54      | -             | <b>51</b> | 6.03            |
| <b>25</b> | -         | 3.90      | -         | -             | <b>52</b> | 12.3            |
| <b>26</b> | -         | 12.1      | -         | -             | <b>53</b> | 31.5            |
| <b>27</b> | -         | 47.5      | -         | -             | <b>54</b> | 10.7            |

Table 6.2: Cable lengths on four branches.

of each branch: 23, 77, 101 and 123. The graphs also include a dashed line the voltage statutory limits<sup>4</sup> over and under 6% of 230V [78].

In both cases, winter and summer, the most affected branch in terms of voltage levels is B, followed by C and D. The power demands and micro generation in each node are fairly homogeneous (even though they are stochastic demands coming from similar types of appliances), which leaves the topology of the network as the main factor in creating the differences between branches. Firstly, branch B is

<sup>4</sup>Voltage regulation from LV busbars of the HV/LV transformer to any service cut-out shall not exceed: 6% of 230 volts when supplied from *Standard 11kV Feeders*; 4% of 230 volts when supplied from *Long 11kV Feeders*, defined as extending beyond the 15km radius of a Bulk Supply Point or Primary Substation.

the longest one, with twice as many nodes as the other branches. Secondly, the connection to buildings in branches C and D has a long cable stretch across a bus depot on the same street, with consequential power losses.

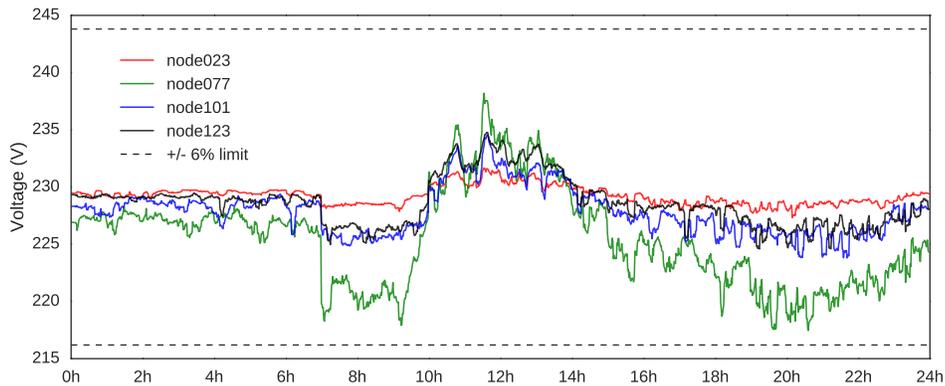
Voltage profiles in winter and summer also show differences. In winter (Figure 6.7a), there is the clear effect of heating systems (there has not been a stochastic model applied to them, but they follow a deterministic schedule). They cause voltage drops in the morning and also in the evening. Again, these drops are deeper for those branches more intensely loaded or with larger losses.

On the other hand, the same network has to deal with a very different voltage pattern on a summer day (Figure 6.7b). The absence of heating is visible in the summer profile. With longer periods of more intense solar irradiation, there are clear over-voltage issues, exceeding the upper +6% limit on several occasions between 10h and 18h.

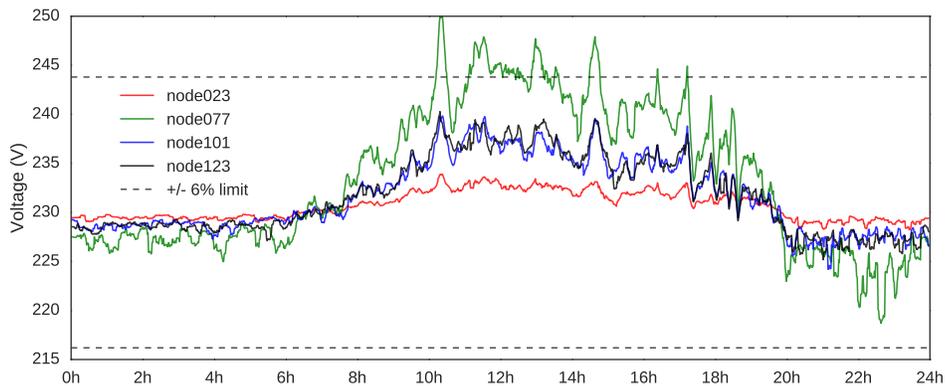
The issues described above, could be addressed with DR solutions. Load shifting and electricity storage can help to mitigate peak demands, and optimise the use of renewable energy generation to avoid curtailment measures. In any case, information of the steady state operation of the network (such as that presented in this section) should be taken into account, and provided to the DR algorithm to enable it to take more informed actions.

## Simulation performance

A key feature of the load-flow algorithm presented here is its ability to deal with multi-branch networks in a very efficient way, given the simplicity of the method used. To prove this, the original network topology in Figure 6.6 has been redefined, reconnecting its nodes in a more complex layout, depicted in Figure 6.8. Both



(a) Winter.



(b) Summer.

Figure 6.7: Voltage levels through a day in winter and a day in summer, for nodes at the end of the lines in each branch.

topologies, the original and the new modified one, given a  $tolerance = 10^{-5}$  (as used in Algorithm 3), require a low number of iterations, between 2 and 4, to achieve convergence. Thus, the increased complexity in the topology does not involve extra iterations for convergence.

The simulation time, however, is slightly larger in the case of the modified layout, with an increase of 3% in the simulation time of a single time-step, from an average of  $6.265 \cdot 10^{-2}s$  to  $6.458 \cdot 10^{-2}s$ . This value corresponds to the execution

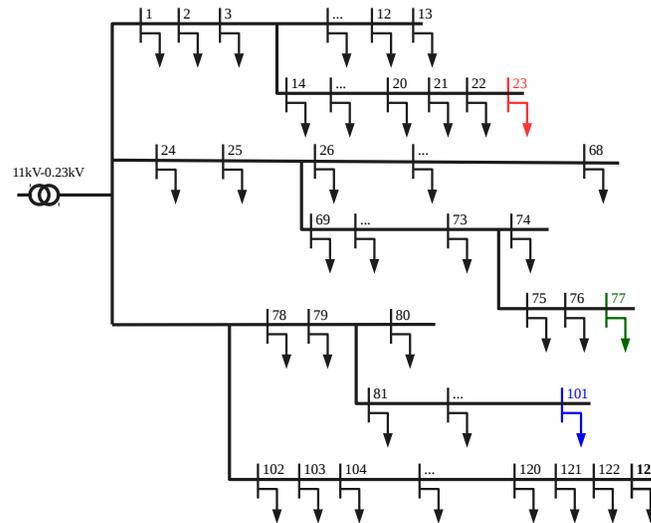


Figure 6.8: Imaginary modified topology for performance comparison.

time of the load-flow algorithm for a single time-step. The distribution of time values is presented in Figure 6.9.

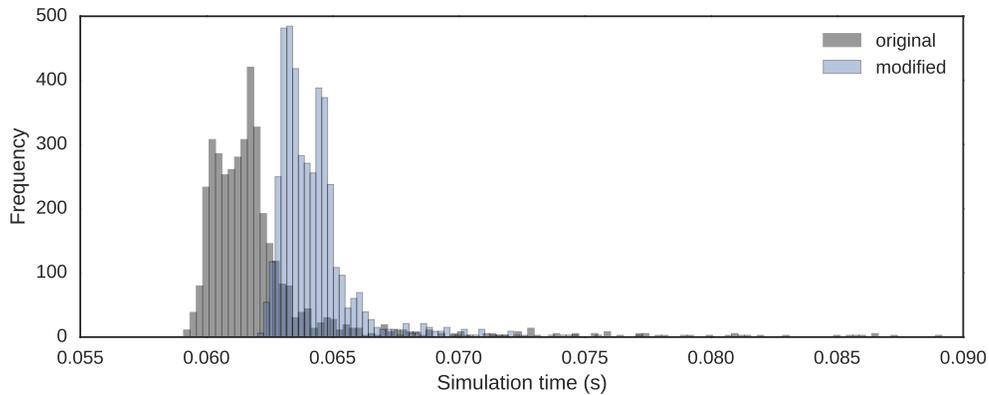


Figure 6.9: Distribution of simulation times to achieve convergence in one time-step, for both topologies.

## 6.6 Conclusion

In this chapter a low-complexity algorithm to efficiently solve radial distribution networks has been shown. Based on iterative forward/backward sweep methods,

and coded using object-oriented programming and recursion, the power-flow algorithm can solve branched systems for any (non-meshed) topologies.

To show the potential of the model, a network containing over 120 demand nodes (from homes) and its corresponding generation nodes (PV panels) has been simulated. In total, the network was defined with 246 nodes.

The model and algorithm are highly flexible. Although aggregated demand coming from homes has been used, it would be straightforward to connect to power demands from individual appliances. This could be relevant if simulating in detail a smaller groups of houses, or even a single home, with all its electrical devices connected to it, given precise information on the cable impedances.

Losses in the case presented here are very roughly modelled in terms of cable impedance, but more detailed models or real data could be used instead. The only thing required is a .csv file containing impedance values between the nodes.

Even though it was not used here, the algorithm is ready to deal with complex magnitudes, and therefore to calculate reactive power as well as real power. For that, impedance values should have a reactance terms as well, and the demand and generation models should include some form of power factor measure.

When executing the algorithm, it provides information on the network according to the demand and supply conditions. The main differences between the two seasons are found on the voltage rise during a summer day, and the intense morning and evening power loads on the winter day.

Those events reflected here in a one-day simulation, the need for this information to be part of the control algorithm in any DR software. Information on the operational state of the grid should be fed into the optimization process. Including feedback of the grid is a priority for the future development of No-MASS/DR.

It was also shown, that increasing the complexity of the network layout, leads to a slight increase of the order of just 3% in the algorithm's execution time for the same number of nodes. However, the number of iterations for the backward/-forward sweep remains the same, for a tolerance of  $10^{-5}$ .

It has also been shown that No-MASS/DR is able to simulate demand and supply, and also load shifting mechanisms, for over 100 buildings. The major limitation in running No-MASS/DR for such large networks is writing and saving data of the results. If the saved data is limited to a minimum required, it can solve for multiple buildings simultaneously.

# Chapter 7

## Conclusions

### 7.1 Introduction

This thesis aimed to develop better unified modelling of Demand Response strategies, that required integrated modelling of energy systems, with a particular focus on the study of maximising locally generated renewable energy. Section 1.2 of the opening chapter of this thesis outlined a set of objectives to contribute to this endeavour. Those objectives are reviewed in this final chapter, followed by some recommendations to continue this line of research.

### 7.2 Achievements

#### 7.2.1 Objective I

The individual goals of objective I have been satisfied as follows:

- I.1 A range of modelling strategies were tested and evaluated. They considered two different types of stochastic approach. In the first, states of the fractional

power demand were considered to be a discrete-time random process, and were modelled with a Markov chain approach. In the second, they were considered to be a continuous-time random process, and survival analysis was employed to model them. For these two approaches, seven different models were evaluated.

I.2 Data clustering techniques were deployed, to search for a parsimonious model structure, by minimising the dimensionality of state transition probability matrices.

Supported by a comprehensive validation procedure, a survival multistate model was selected, which satisfies our modelling objectives:

- a. It describes temporal dependency of the electricity demand, as demonstrated in Figure 3.4 and table 3.4.
- b. Small appliances are modelled by their typologies, following four categories: audio-visual, computing, kitchen and other appliances.

The novelty of this work lies in the effectiveness of the methodology to model groups of small appliances, even those of a different type. The mathematical formulation to reduce Markov's transition matrix' dimensionality, uses a state-of-the-art density clustering algorithm, with potential applications beyond the energy modelling field.

## 7.2.2 Objective II

II.1 A comprehensive set of requirements to be met by DR software were articulated in section 4.2.

II.2 Multi Agent Simulation has proven successful in encapsulating the multifaceted nature of the DR problem. Appliances, generation units and storage systems were modelled as software agents that are able to communicate and negotiate with each other to achieve energy matching objectives at every time step. Moreover, it was shown that a Multi Agent Simulation approach could be easily extended in scope, from a single household to a small community. Computational limitations in running No-MASS/DR arise from the amount of data that needs to be written out (which can be selected in the configuration file). As long as the number of columns is well selected and kept to the minimum necessary, No-MASS/DR can be used to consider larger neighbourhoods.

II.3 For handling DR strategies, machine learning is useful. In particular, Q-learning algorithms have been shown to be both computationally simple and highly effective for DR optimisation, with system information easily integrated in its reward function.

II.4 A proof of concept of No-MASS/DR was demonstrated for the case of a single building and a community of buildings. Our algorithms are effective in simulating the effects of technology uptake (smart appliances and electric batteries) on the self-consumption of electric energy.

No-MASS/DR constitutes a novel contribution as a simulation framework that can simultaneously handle *i)* stochastic power demands, *ii)* device-to-device interactions for energy balancing and DR of electrical equipment and *iii)* load flow analysis. It complements the existing functionality to handle *iv)* occupants' behaviour and *v)* human-to-human interaction, and sets the foundations to consider

*vi*) human-to-device interactions.

### 7.2.3 Objective III

A model for the Low Voltage network was developed and presented in Chapter 6.

In this we have shown that:

III.1 The source code is available, and successfully implemented with No-MASS/DR.

III.2 The algorithm efficiently simulates a branched radial network containing over 100 nodes. Increasing the complexity of the branched layout leads to a slight increase in the simulation time.

III.3 The nodes may be defined either for the aggregated demand of the household, or for the PV solar panels. However, it is trivial to consider the nodes as individual appliances, if a more detail simulation was required.

III.4 The network model is highly flexible, and can include different types of local generation or storage.

A novel implementation to perform load-flow analysis with a forward/backward sweep method is put into practise. It uses Object Oriented programming and recursion in an algorithm that efficiently handles branched radial networks and is specifically developed to be integrated into a Demand Response simulation framework.

## 7.3 Further work and recommendations

As part of objective II, a set of six requirements for DR software were identified and formulated in Section 4.2, briefly repeated here:

- R1. Simulating stochastic demands of appliances.
- R2. Satisfying demands from local generation and storage, and the local or national grid.
- R3. Optimizing the decision of where power should be drawn from/diverted to, in order to satisfy an objective function.
- R4. Presenting information to the user and emulating the users' decision making rationale regarding the rescheduling of user-controlled devices.
- R5. Accounting for diversity in the extent to which users are willing to relinquish control and to actively engage in behavioural change.
- R6. Facilitating the above for communities of buildings which can communicate to achieve individual homeowners' requirements and also low voltage network requirements.

Of those, we have contributed, to a greater or lesser degree, to R1, R2, R3 and R6, providing the foundations for such software. However, there is scope for improvement both in terms of the success with which these requirements are met and how the remaining requirements could also be comprehensively addressed.

### **Ideas to improve R1-R3**

- R1. Considerable effort has been invested in modelling appliance use in a way that describes the stochastic nature of power loads. The contribution of this thesis to R1 consisted of small appliances modelling, which complemented existing models of large appliances.

Such models have been exploited in this work to study the effect of Demand Response programmes, whose primary focus is on a specific type of large

appliance, those under category *ii*) switched on by the user and off when a programme is complete, as their shifting imposes little disruption for the user (washing machine, dishwasher and tumble dryer). However, there is no reason why DR mechanisms (load shifting or others) couldn't be extended to other types of device and appliance in the future. In particular, small appliances could play a more evident role here. Although the practicability of using small appliances for load shifting is arguable, there are available options. For instance, users could be asked to charge computing and audio-visual equipment with an incorporated battery at certain times of the day, knowing when they are more/less likely to be used, and when they need to be fully charged. To be able to simulate this, available models of use of small appliances are necessary. The challenge here would be to develop a strategy that models DR responses to appliances that are represented as an aggregate (that information describing the composition of the aggregate typology is represented).

Finally, extensions to this work should consider the effects of electric vehicles. The amount of energy needed to charge car batteries may be comparable to the energy required to run household appliances. Considering the likely increase in EV uptake, DR scenarios need to take into account the growing penetration of electrified transport.

R2. A software architecture has been developed that handles the distribution of power from local sources and the grid to satisfy demands. Information regarding generation and storage has been obtained with simplified models or directly with data, and only considered local generation of solar energy. This was sufficient for the demonstration of our proof-of-concept. Nevertheless, for

a robust operation of DR software, the following is necessary:

- A more comprehensive palette of local generation technologies should be modelled, including wind turbines and co-generation plants. Building simulation software include radiation models and weather data to facilitate this, and inverter models are also sometimes available.
- The electric battery model is highly simplified. Storage losses with time are not explicitly handled. If they were, the discharge algorithm should penalise long term storage. Likewise, refinements of the model should take into account battery degradation over time, as it may have important cost impacts.

R3. No-MASS/DR would benefit from an exhaustive investigation of alternative learning algorithm candidates. One promising candidate is a Distributed W-Learning (DWL) algorithm [49]. Other options for reinforcement learning, such as genetic algorithms, could also be explored. As explained in [89], different type of algorithms do not necessarily consistently outperform their counterparts, so that hybrid solutions are sometimes more robust.

### **Ideas to address R4 and R5**

A large part of the motivation for adopting and extending No-MASS, lies in its ability to model users' behaviours and their interactions with energy systems. In that sense, some of the modules in No-MASS already contain models that relate socio-demographic characteristics of occupants with their activities (cooking, washing, sleeping, etc.), their actions on the building envelope (opening/closing windows and shading devices) or the social interactions between occupants and the corresponding energy implications.

An obvious next step would be to model of the interactions between people and devices, as expressed in R4 and R5. Properly informed by empirical evidence, this would provide invaluable support in testing the viability of alternative DR schemes. So far, the only attempt to model human intervention available in No-MASS/DR is through the definition of a delivery time of re-scheduled appliances. This is fixed for the duration of the simulation, and assumes that the users always agree with the new schedule. A more refined setup of the selection of the delivery time would be appropriate. For example, when a re-schedule event is predicted, the user could make a decision either to accept, modify or refuse the action, based on a pre-defined level of engagement in the DR scheme (informed by the corresponding rewards).

This level of engagement could be put into numbers with a linear parametrization of the probability of reaction dependent on the (financial) incentive, as in equation 7.1:

$$P(x) = ex + g, \tag{7.1}$$

where  $e$  relates to the flexibility of the user to change its behaviour (elasticity), and ranges from inelastic behaviours when users do not want to relinquish control, to elastic behaviours, when users are relatively easily influenced by incentives. The intercept  $g$  is related to "green awareness", a willingness to change behaviour even when no incentive is guaranteed. This simplified linear model is depicted in Figure 7.1.

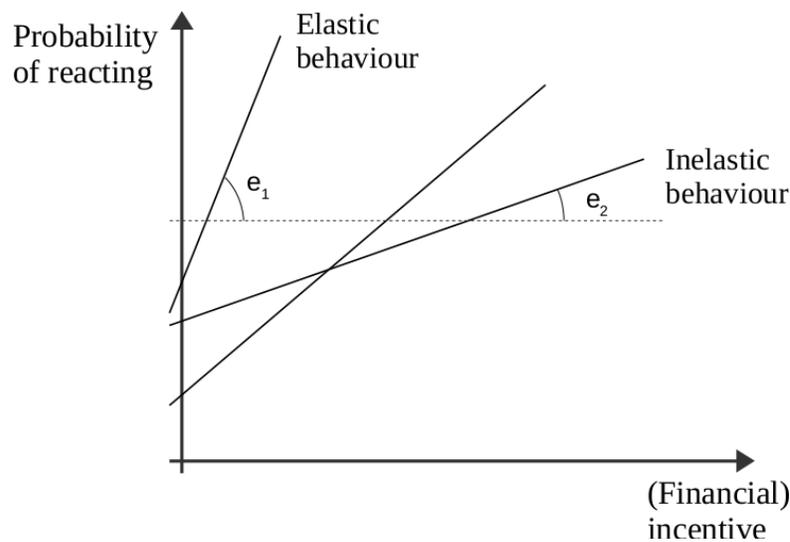


Figure 7.1: Linear behavioural model.

### Ideas to improve R6

R6 was only partially addressed in this work, leaving scope for improvement. Although the model was implemented into No-MASS/DR, the real potential in modelling the LV network comes from integrating its results in the DR algorithm as a driver variable. This means that for each time step, the steady-state conditions of the distribution network are known, and actions can be taken accordingly, for network stability and safety. One way of doing this could be to absorb network information into the existing reward functions of the learning appliance agents and battery agents. Another possibly more effective way, could be to add a new learning grid agent. It could learn about situations that compromise network safety and be able to force corrective action from other agents when such events occur.

The existence of the explicit grid agent could be useful for upscaling the model to consider a group of LV networks and its connection to a primary substation. The operating point of the slack bus, which connects the LV to the rest of the grid, considered in our simulations has to be constant during the iteration process

at each time step. However, it can change during the simulation, opening the possibility of simulating the varying state of the grid with time, allowing for a more realistic simulation.

Finally, No-MASS/DR could be extended to also consider a detached islanded (e.g. remote rural) network. If a grid agent is not defined, or for any reason is not able to provide the required power, appliances may have a priority of service defined (this feature is already available in No-MASS/DR), allowing the modeller to specify a (time-dependent) hierarchy of services to be followed in the case of a state of limited supply. Associating a degree of discomfort every time a service is not provided, No-MASS/DR could be used to explore the impacts of different options for such hierarchical representations .

## A Small Appliance Modelling parameters

For purposes of implementation of the model, tables [A.1](#), [A.2](#), [A.3](#), [A.4](#) and [A.5](#) contain the values of the parameters obtained from the dataset. The survival multistate model is presented in Algorithm 4 as pseudo-code.

---

```

1:  $s = 0$  ▷ Assume initial state  $s_0$ .
2:  $t = t_{START}$ 
3: while  $t < t_{END}$  do
    % Calculate duration at state  $s$ 
4:    $R1 = random(0, 1)$ 
5:    $\lambda, k, \gamma = \lambda_s, k_s, \gamma_s$  ▷ Use table A.1.
6:    $t_s = \gamma + \lambda [-\ln(R1)]^{1/k}$ 
7:    $t = t + t_s$ 
    % Calculate next state
8:    $R2 = random(0, 1)$ 
9:    $s = MinIdx [(cdfP(t) - R2) > 0]$  ▷ Use table A.3 and A.4.
10:  Append  $s$  to  $s_{arr}$ 
11: end while
12: % Transform into fractional energy array.
13:  $f_{arr} = \tilde{F}_s \times s_{arr}$  ▷ Use table A.2.
14: % Transform into energy use array.
15:  $E_{arr} = \tilde{E}_{max} \times f_{arr}$  ▷ Use table A.5.

```

---

Algorithm 4: SIMULATE SMALL APPLIANCE USAGE ( $[\lambda, k, \gamma]_s, P(t, \mathbf{s})$ )

| Energy state | Audio-visual |       |           | Computing |       |           | Kitchen  |      |           | Other    |       |           |
|--------------|--------------|-------|-----------|-----------|-------|-----------|----------|------|-----------|----------|-------|-----------|
|              | $\gamma$     | $k$   | $\lambda$ | $\gamma$  | $k$   | $\lambda$ | $\gamma$ | $k$  | $\lambda$ | $\gamma$ | $k$   | $\lambda$ |
| $s_0$        | 8.92         | 0.743 | 10.74     | 7.52      | 1.529 | 12.52     | 7.80     | 1.37 | 4.29      | 7.46     | 0.930 | 9.61      |
| $s_1$        | 8.29         | 0.916 | 7.52      | 7.29      | 1.110 | 4.87      | 8.93     | 1.17 | 6.57      | 8.25     | 1.148 | 5.20      |
| $s_2$        | 8.38         | 1.096 | 9.82      | 8.03      | 0.889 | 5.37      | 8.39     | 1.25 | 4.29      | 8.83     | 1.147 | 7.16      |
| $s_3$        | 8.33         | 0.965 | 6.95      | 8.83      | 0.607 | 20.14     | 8.13     | 1.06 | 5.35      | 9.34     | 1.201 | 9.03      |
| $s_4$        | 8.95         | 0.648 | 12.20     | 8.52      | 0.977 | 6.19      | 8.02     | 1.26 | 4.50      | 8.86     | 0.872 | 7.34      |
| $s_5$        | 9.45         | 0.980 | 15.59     | 8.13      | 0.870 | 6.36      | 8.34     | 1.27 | 4.20      | 9.03     | 1.148 | 13.35     |
| $s_6$        | 9.02         | 0.747 | 13.06     | 8.68      | 0.790 | 7.76      | 8.72     | 1.12 | 5.21      | 8.89     | 1.070 | 6.92      |
| $s_7$        | 9.03         | 1.065 | 11.72     | 8.76      | 0.854 | 10.97     | 8.22     | 1.33 | 4.01      | 8.23     | 1.046 | 5.32      |
| $s_8$        | 8.40         | 1.005 | 7.25      | 8.95      | 0.657 | 13.82     | 7.63     | 1.28 | 4.58      | 8.39     | 1.214 | 4.47      |
| $s_9$        | 8.79         | 0.805 | 13.93     | 8.70      | 0.872 | 15.63     | 8.01     | 1.24 | 4.36      | 8.64     | 0.989 | 6.30      |
| $s_{10}$     | 8.24         | 1.051 | 5.56      | 8.30      | 1.093 | 5.72      | 8.73     | 1.16 | 7.12      | 8.65     | 0.963 | 8.62      |

Table A.1: Survival distribution parameters ( $\gamma$ : location,  $k$ : shape,  $\lambda$ : scale) for four categories.

| Energy state | $\bar{F}_{Audio-visual}$ | $\bar{F}_{Computing}$ | $\bar{F}_{Kitchen}$ | $\bar{F}_{Other}$ |
|--------------|--------------------------|-----------------------|---------------------|-------------------|
| $s_0$        | 0.0000                   | 0.0000                | 0.0000              | 0.0000            |
| $s_1$        | 0.0402                   | 0.0333                | 0.0007              | 0.0265            |
| $s_2$        | 0.1429                   | 0.1297                | 0.1382              | 0.1587            |
| $s_3$        | 0.2500                   | 0.2222                | 0.2473              | 0.2513            |
| $s_4$        | 0.3333                   | 0.3500                | 0.3570              | 0.3684            |
| $s_5$        | 0.4667                   | 0.4286                | 0.4494              | 0.5000            |
| $s_6$        | 0.5525                   | 0.5581                | 0.5495              | 0.5450            |
| $s_7$        | 0.6660                   | 0.6526                | 0.6597              | 0.6565            |
| $s_8$        | 0.7708                   | 0.7717                | 0.7463              | 0.7500            |
| $s_9$        | 0.8750                   | 0.8462                | 0.8405              | 0.8333            |
| $s_{10}$     | 0.9571                   | 1.0000                | 0.9684              | 0.9167            |

Table A.2: Median fractional energy for transforming energy states into fractional energy profiles.

| Hour of day | Audio-visual |            |            |            |            |            |            |            |            |            |               |
|-------------|--------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------|
|             | $P(s=s_0)$   | $P(s=s_1)$ | $P(s=s_2)$ | $P(s=s_3)$ | $P(s=s_4)$ | $P(s=s_5)$ | $P(s=s_6)$ | $P(s=s_7)$ | $P(s=s_8)$ | $P(s=s_9)$ | $P(s=s_{10})$ |
| 0h          | 0.3587       | 0.1437     | 0.1174     | 0.0792     | 0.0406     | 0.0392     | 0.0293     | 0.0443     | 0.0470     | 0.0710     | 0.0295        |
| 1h          | 0.3799       | 0.1566     | 0.1172     | 0.0800     | 0.0353     | 0.0378     | 0.0305     | 0.0446     | 0.0459     | 0.0517     | 0.0205        |
| 2h          | 0.3844       | 0.1667     | 0.1171     | 0.0803     | 0.0332     | 0.0372     | 0.0384     | 0.0423     | 0.0442     | 0.0408     | 0.0153        |
| 3h          | 0.3859       | 0.1715     | 0.1157     | 0.0820     | 0.0302     | 0.0375     | 0.0441     | 0.0403     | 0.0441     | 0.0366     | 0.0122        |
| 4h          | 0.3883       | 0.1756     | 0.1128     | 0.0813     | 0.0298     | 0.0398     | 0.0440     | 0.0400     | 0.0436     | 0.0336     | 0.0113        |
| 5h          | 0.3879       | 0.1797     | 0.1088     | 0.0805     | 0.0298     | 0.0398     | 0.0448     | 0.0412     | 0.0420     | 0.0342     | 0.0114        |
| 6h          | 0.3853       | 0.1719     | 0.1115     | 0.0785     | 0.0303     | 0.0379     | 0.0419     | 0.0424     | 0.0449     | 0.0404     | 0.0151        |
| 7h          | 0.3781       | 0.1638     | 0.1117     | 0.0799     | 0.0336     | 0.0391     | 0.0456     | 0.0404     | 0.0451     | 0.0436     | 0.0193        |
| 8h          | 0.3607       | 0.1704     | 0.1020     | 0.0368     | 0.0412     | 0.0412     | 0.0506     | 0.0412     | 0.0460     | 0.0481     | 0.0223        |
| 9h          | 0.3437       | 0.1709     | 0.1029     | 0.0845     | 0.0390     | 0.0423     | 0.0509     | 0.0427     | 0.0481     | 0.0499     | 0.0251        |
| 10h         | 0.3311       | 0.1682     | 0.1057     | 0.0860     | 0.0397     | 0.0438     | 0.0521     | 0.0426     | 0.0475     | 0.0542     | 0.0291        |
| 11h         | 0.3283       | 0.1657     | 0.1044     | 0.0859     | 0.0377     | 0.0419     | 0.0518     | 0.0414     | 0.0502     | 0.0595     | 0.0333        |
| 12h         | 0.3287       | 0.1610     | 0.1022     | 0.0876     | 0.0332     | 0.0413     | 0.0502     | 0.0410     | 0.0542     | 0.0648     | 0.0358        |
| 13h         | 0.3255       | 0.1554     | 0.1032     | 0.0872     | 0.0318     | 0.0434     | 0.0485     | 0.0400     | 0.0549     | 0.0701     | 0.0400        |
| 14h         | 0.3202       | 0.1540     | 0.1049     | 0.0860     | 0.0322     | 0.0436     | 0.0467     | 0.0417     | 0.0567     | 0.0718     | 0.0421        |
| 15h         | 0.3144       | 0.1496     | 0.1027     | 0.0865     | 0.0346     | 0.0480     | 0.0441     | 0.0465     | 0.0566     | 0.0730     | 0.0440        |
| 16h         | 0.3073       | 0.1408     | 0.1035     | 0.0879     | 0.0374     | 0.0505     | 0.0426     | 0.0494     | 0.0560     | 0.0753     | 0.0492        |
| 17h         | 0.3021       | 0.1319     | 0.1039     | 0.0905     | 0.0387     | 0.0512     | 0.0443     | 0.0515     | 0.0564     | 0.0786     | 0.0510        |
| 18h         | 0.2994       | 0.1280     | 0.1020     | 0.0888     | 0.0399     | 0.0497     | 0.0454     | 0.0539     | 0.0565     | 0.0825     | 0.0539        |
| 19h         | 0.3001       | 0.1267     | 0.1022     | 0.0892     | 0.0416     | 0.0518     | 0.0417     | 0.0500     | 0.0563     | 0.0880     | 0.0523        |
| 20h         | 0.2976       | 0.1279     | 0.1068     | 0.0902     | 0.0425     | 0.0471     | 0.0400     | 0.0492     | 0.0555     | 0.0887     | 0.0545        |
| 21h         | 0.2953       | 0.1276     | 0.1054     | 0.0904     | 0.0438     | 0.0461     | 0.0409     | 0.0470     | 0.0565     | 0.0921     | 0.0550        |
| 22h         | 0.3008       | 0.1289     | 0.1046     | 0.0854     | 0.0450     | 0.0475     | 0.0418     | 0.0458     | 0.0550     | 0.0933     | 0.0519        |
| 23h         | 0.3222       | 0.1340     | 0.1097     | 0.0817     | 0.0449     | 0.0457     | 0.0383     | 0.0450     | 0.0512     | 0.0862     | 0.0419        |

| Hour of day | Computing  |            |            |            |            |            |            |            |            |            |               |
|-------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|------------|---------------|
|             | $P(s=s_0)$ | $P(s=s_1)$ | $P(s=s_2)$ | $P(s=s_3)$ | $P(s=s_4)$ | $P(s=s_5)$ | $P(s=s_6)$ | $P(s=s_7)$ | $P(s=s_8)$ | $P(s=s_9)$ | $P(s=s_{10})$ |
| 0h          | 0.4084     | 0.2563     | 0.0450     | 0.0370     | 0.0266     | 0.0376     | 0.0405     | 0.0410     | 0.0534     | 0.0394     | 0.0148        |
| 1h          | 0.4232     | 0.2626     | 0.0479     | 0.0373     | 0.0237     | 0.0341     | 0.0349     | 0.0348     | 0.0520     | 0.0361     | 0.0134        |
| 2h          | 0.4306     | 0.2672     | 0.0544     | 0.0382     | 0.0230     | 0.0326     | 0.0273     | 0.0324     | 0.0496     | 0.0325     | 0.0124        |
| 3h          | 0.4348     | 0.2679     | 0.0575     | 0.0379     | 0.0225     | 0.0322     | 0.0237     | 0.0307     | 0.0480     | 0.0325     | 0.0122        |
| 4h          | 0.4361     | 0.2676     | 0.0578     | 0.0379     | 0.0226     | 0.0317     | 0.0240     | 0.0300     | 0.0388     | 0.0411     | 0.0123        |
| 5h          | 0.4364     | 0.2664     | 0.0575     | 0.0379     | 0.0230     | 0.0319     | 0.0238     | 0.0334     | 0.0474     | 0.0311     | 0.0113        |
| 6h          | 0.4331     | 0.2653     | 0.0565     | 0.0376     | 0.0234     | 0.0324     | 0.0257     | 0.0373     | 0.0486     | 0.0296     | 0.0105        |
| 7h          | 0.4259     | 0.2589     | 0.0526     | 0.0376     | 0.0247     | 0.0334     | 0.0323     | 0.0428     | 0.0499     | 0.0307     | 0.0113        |
| 8h          | 0.4138     | 0.2484     | 0.0461     | 0.0379     | 0.0265     | 0.0352     | 0.0409     | 0.0514     | 0.0544     | 0.0333     | 0.0122        |
| 9h          | 0.4001     | 0.2394     | 0.0432     | 0.0382     | 0.0303     | 0.0357     | 0.0445     | 0.0563     | 0.0607     | 0.0376     | 0.0141        |
| 10h         | 0.3884     | 0.2328     | 0.0417     | 0.0391     | 0.0322     | 0.0376     | 0.0487     | 0.0589     | 0.0627     | 0.0419     | 0.0161        |
| 11h         | 0.3818     | 0.2295     | 0.0405     | 0.0381     | 0.0322     | 0.0403     | 0.0517     | 0.0603     | 0.0629     | 0.0447     | 0.0181        |
| 12h         | 0.3773     | 0.2268     | 0.0389     | 0.0392     | 0.0327     | 0.0407     | 0.0539     | 0.0614     | 0.0642     | 0.0458     | 0.0191        |
| 13h         | 0.3723     | 0.2259     | 0.0376     | 0.0397     | 0.0327     | 0.0425     | 0.0548     | 0.0621     | 0.0642     | 0.0482     | 0.0201        |
| 14h         | 0.3660     | 0.2249     | 0.0376     | 0.0403     | 0.0329     | 0.0428     | 0.0558     | 0.0636     | 0.0655     | 0.0493     | 0.0214        |
| 15h         | 0.3548     | 0.2228     | 0.0376     | 0.0407     | 0.0340     | 0.0436     | 0.0559     | 0.0663     | 0.0691     | 0.0524     | 0.0227        |
| 16h         | 0.3428     | 0.2223     | 0.0375     | 0.0409     | 0.0367     | 0.0411     | 0.0587     | 0.0675     | 0.0719     | 0.0555     | 0.0251        |
| 17h         | 0.3356     | 0.2179     | 0.0386     | 0.0411     | 0.0385     | 0.0417     | 0.0604     | 0.0700     | 0.0736     | 0.0579     | 0.0266        |
| 18h         | 0.3292     | 0.2141     | 0.0396     | 0.0411     | 0.0394     | 0.0451     | 0.0619     | 0.0708     | 0.0765     | 0.0600     | 0.0263        |
| 19h         | 0.3196     | 0.2118     | 0.0395     | 0.0430     | 0.0380     | 0.0495     | 0.0633     | 0.0706     | 0.0784     | 0.0599     | 0.0266        |
| 20h         | 0.3162     | 0.2156     | 0.0391     | 0.0439     | 0.0384     | 0.0511     | 0.0627     | 0.0683     | 0.0775     | 0.0602     | 0.0270        |
| 21h         | 0.3129     | 0.2213     | 0.0391     | 0.0426     | 0.0359     | 0.0536     | 0.0615     | 0.0674     | 0.0732     | 0.0585     | 0.0276        |
| 22h         | 0.3430     | 0.2349     | 0.0413     | 0.0403     | 0.0310     | 0.0490     | 0.0555     | 0.0622     | 0.0666     | 0.0512     | 0.0250        |
| 23h         | 0.3785     | 0.2498     | 0.0433     | 0.0366     | 0.0286     | 0.0434     | 0.0469     | 0.0519     | 0.0581     | 0.0434     | 0.0195        |

Table A.3: Hourly state probability for Audio-visual and Computing categories.

| Hour of day | Kitchen      |              |              |              |              |              |              |              |              |              |                 |
|-------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|--------------|-----------------|
|             | $P(s = s_0)$ | $P(s = s_1)$ | $P(s = s_2)$ | $P(s = s_3)$ | $P(s = s_4)$ | $P(s = s_5)$ | $P(s = s_6)$ | $P(s = s_7)$ | $P(s = s_8)$ | $P(s = s_9)$ | $P(s = s_{10})$ |
| 0h          | 0.9403       | 0.0567       | 0.0006       | 0.0007       | 0.0006       | 0.0005       | 0.0004       | 0.0002       | 0.0000       | 0.0000       | 0.0000          |
| 1h          | 0.9418       | 0.0561       | 0.0003       | 0.0004       | 0.0005       | 0.0004       | 0.0003       | 0.0001       | 0.0000       | 0.0000       | 0.0000          |
| 2h          | 0.9408       | 0.0560       | 0.0004       | 0.0005       | 0.0007       | 0.0007       | 0.0005       | 0.0002       | 0.0001       | 0.0000       | 0.0000          |
| 3h          | 0.9357       | 0.0580       | 0.0008       | 0.0009       | 0.0017       | 0.0016       | 0.0008       | 0.0003       | 0.0001       | 0.0001       | 0.0001          |
| 4h          | 0.9328       | 0.0595       | 0.0016       | 0.0012       | 0.0022       | 0.0017       | 0.0004       | 0.0003       | 0.0002       | 0.0001       | 0.0001          |
| 5h          | 0.9209       | 0.0624       | 0.0033       | 0.0025       | 0.0049       | 0.0034       | 0.0017       | 0.0005       | 0.0002       | 0.0001       | 0.0001          |
| 6h          | 0.9041       | 0.0719       | 0.0061       | 0.0034       | 0.0054       | 0.0048       | 0.0026       | 0.0010       | 0.0003       | 0.0002       | 0.0001          |
| 7h          | 0.8768       | 0.0836       | 0.0090       | 0.0068       | 0.0086       | 0.0079       | 0.0038       | 0.0023       | 0.0007       | 0.0004       | 0.0001          |
| 8h          | 0.8683       | 0.0862       | 0.0114       | 0.0084       | 0.0100       | 0.0081       | 0.0081       | 0.0015       | 0.0006       | 0.0005       | 0.0004          |
| 9h          | 0.8656       | 0.0924       | 0.0096       | 0.0082       | 0.0084       | 0.0073       | 0.0041       | 0.0023       | 0.0009       | 0.0008       | 0.0004          |
| 10h         | 0.8740       | 0.0913       | 0.0077       | 0.0077       | 0.0072       | 0.0057       | 0.0030       | 0.0015       | 0.0006       | 0.0007       | 0.0005          |
| 11h         | 0.8759       | 0.0910       | 0.0069       | 0.0075       | 0.0075       | 0.0056       | 0.0028       | 0.0013       | 0.0006       | 0.0005       | 0.0004          |
| 12h         | 0.8854       | 0.0861       | 0.0056       | 0.0068       | 0.0070       | 0.0045       | 0.0026       | 0.0010       | 0.0005       | 0.0003       | 0.0002          |
| 13h         | 0.8832       | 0.0870       | 0.0056       | 0.0076       | 0.0076       | 0.0048       | 0.0028       | 0.0013       | 0.0006       | 0.0004       | 0.0002          |
| 14h         | 0.8839       | 0.0879       | 0.0053       | 0.0076       | 0.0072       | 0.0040       | 0.0022       | 0.0010       | 0.0004       | 0.0004       | 0.0002          |
| 15h         | 0.8858       | 0.0847       | 0.0056       | 0.0073       | 0.0077       | 0.0045       | 0.0021       | 0.0010       | 0.0005       | 0.0004       | 0.0004          |
| 16h         | 0.8863       | 0.0839       | 0.0058       | 0.0069       | 0.0066       | 0.0046       | 0.0029       | 0.0012       | 0.0011       | 0.0005       | 0.0003          |
| 17h         | 0.8839       | 0.0845       | 0.0062       | 0.0077       | 0.0065       | 0.0043       | 0.0025       | 0.0014       | 0.0017       | 0.0005       | 0.0004          |
| 18h         | 0.8932       | 0.0792       | 0.0055       | 0.0063       | 0.0059       | 0.0043       | 0.0021       | 0.0012       | 0.0012       | 0.0006       | 0.0004          |
| 19h         | 0.9016       | 0.0760       | 0.0045       | 0.0053       | 0.0049       | 0.0035       | 0.0020       | 0.0010       | 0.0006       | 0.0004       | 0.0002          |
| 20h         | 0.9072       | 0.0719       | 0.0039       | 0.0048       | 0.0042       | 0.0036       | 0.0015       | 0.0009       | 0.0005       | 0.0003       | 0.0002          |
| 21h         | 0.9171       | 0.0672       | 0.0026       | 0.0037       | 0.0044       | 0.0037       | 0.0012       | 0.0007       | 0.0004       | 0.0001       | 0.0002          |
| 22h         | 0.9248       | 0.0634       | 0.0021       | 0.0033       | 0.0030       | 0.0017       | 0.0007       | 0.0003       | 0.0002       | 0.0001       | 0.0000          |
| 23h         | 0.9323       | 0.0605       | 0.0017       | 0.0021       | 0.0012       | 0.0010       | 0.0006       | 0.0003       | 0.0001       | 0.0000       | 0.0001          |
| Hour of day | Other        |              |              |              |              |              |              |              |              |              |                 |
|             | $P(s = s_0)$ | $P(s = s_1)$ | $P(s = s_2)$ | $P(s = s_3)$ | $P(s = s_4)$ | $P(s = s_5)$ | $P(s = s_6)$ | $P(s = s_7)$ | $P(s = s_8)$ | $P(s = s_9)$ | $P(s = s_{10})$ |
| 0h          | 0.8724       | 0.0841       | 0.0012       | 0.0029       | 0.0097       | 0.0182       | 0.0001       | 0.0000       | 0.0112       | 0.0000       | 0.0001          |
| 1h          | 0.8730       | 0.0835       | 0.0013       | 0.0027       | 0.0096       | 0.0200       | 0.0000       | 0.0000       | 0.0097       | 0.0000       | 0.0000          |
| 2h          | 0.8722       | 0.0830       | 0.0013       | 0.0026       | 0.0101       | 0.0201       | 0.0001       | 0.0000       | 0.0105       | 0.0000       | 0.0000          |
| 3h          | 0.8750       | 0.0821       | 0.0014       | 0.0025       | 0.0098       | 0.0212       | 0.0002       | 0.0000       | 0.0107       | 0.0000       | 0.0001          |
| 4h          | 0.8725       | 0.0817       | 0.0013       | 0.0026       | 0.0097       | 0.0219       | 0.0002       | 0.0000       | 0.0101       | 0.0000       | 0.0001          |
| 5h          | 0.8718       | 0.0788       | 0.0015       | 0.0027       | 0.0109       | 0.0245       | 0.0004       | 0.0001       | 0.0092       | 0.0000       | 0.0001          |
| 6h          | 0.8692       | 0.0765       | 0.0022       | 0.0032       | 0.0120       | 0.0284       | 0.0004       | 0.0002       | 0.0077       | 0.0000       | 0.0001          |
| 7h          | 0.8674       | 0.0757       | 0.0026       | 0.0036       | 0.0127       | 0.0301       | 0.0008       | 0.0003       | 0.0066       | 0.0000       | 0.0001          |
| 8h          | 0.8644       | 0.0731       | 0.0029       | 0.0038       | 0.0136       | 0.0331       | 0.0011       | 0.0005       | 0.0069       | 0.0003       | 0.0003          |
| 9h          | 0.8631       | 0.0741       | 0.0026       | 0.0036       | 0.0141       | 0.0329       | 0.0010       | 0.0007       | 0.0071       | 0.0004       | 0.0004          |
| 10h         | 0.8577       | 0.0743       | 0.0034       | 0.0044       | 0.0143       | 0.0335       | 0.0015       | 0.0007       | 0.0084       | 0.0008       | 0.0003          |
| 11h         | 0.8556       | 0.0737       | 0.0040       | 0.0046       | 0.0149       | 0.0347       | 0.0013       | 0.0015       | 0.0086       | 0.0008       | 0.0003          |
| 12h         | 0.8571       | 0.0729       | 0.0038       | 0.0044       | 0.0146       | 0.0342       | 0.0014       | 0.0012       | 0.0089       | 0.0013       | 0.0003          |
| 13h         | 0.8565       | 0.0744       | 0.0037       | 0.0039       | 0.0154       | 0.0334       | 0.0013       | 0.0015       | 0.0085       | 0.0011       | 0.0003          |
| 14h         | 0.8579       | 0.0654       | 0.0130       | 0.0031       | 0.0151       | 0.0332       | 0.0015       | 0.0011       | 0.0094       | 0.0011       | 0.0003          |
| 15h         | 0.8600       | 0.0596       | 0.0197       | 0.0020       | 0.0143       | 0.0321       | 0.0019       | 0.0011       | 0.0084       | 0.0006       | 0.0002          |
| 16h         | 0.8614       | 0.0622       | 0.0199       | 0.0022       | 0.0136       | 0.0293       | 0.0015       | 0.0007       | 0.0073       | 0.0007       | 0.0010          |
| 17h         | 0.8615       | 0.0642       | 0.0198       | 0.0022       | 0.0131       | 0.0278       | 0.0015       | 0.0007       | 0.0070       | 0.0007       | 0.0010          |
| 18h         | 0.8622       | 0.0659       | 0.0201       | 0.0024       | 0.0125       | 0.0254       | 0.0014       | 0.0005       | 0.0074       | 0.0010       | 0.0010          |
| 19h         | 0.8639       | 0.0664       | 0.0204       | 0.0016       | 0.0125       | 0.0242       | 0.0011       | 0.0005       | 0.0075       | 0.0008       | 0.0012          |
| 20h         | 0.8639       | 0.0658       | 0.0203       | 0.0018       | 0.0127       | 0.0246       | 0.0010       | 0.0005       | 0.0075       | 0.0006       | 0.0013          |
| 21h         | 0.8681       | 0.0681       | 0.0195       | 0.0020       | 0.0111       | 0.0220       | 0.0009       | 0.0003       | 0.0072       | 0.0007       | 0.0022          |
| 22h         | 0.8624       | 0.0804       | 0.0088       | 0.0031       | 0.0103       | 0.0192       | 0.0007       | 0.0003       | 0.0114       | 0.0003       | 0.0030          |
| 23h         | 0.8701       | 0.0847       | 0.0018       | 0.0032       | 0.0098       | 0.0177       | 0.0002       | 0.0001       | 0.0119       | 0.0001       | 0.0003          |

Table A.4: Hourly state probability for Kitchen and Other categories.

| House | Audio-visual | Computing | Kitchen | Other  |
|-------|--------------|-----------|---------|--------|
| 1     | 1.8          | 0         | 0       | 0      |
| 2     | 18.0         | 0         | 0       | 0      |
| 3     | 27.0         | 0         | 0       | 0      |
| 4     | 30.0         | 0         | 0       | 0      |
| 5     | 43.2         | 6.0       | 0       | 0      |
| 6     | 50.4         | 58.8      | 0       | 0      |
| 7     | 54.0         | 63.0      | 0       | 0      |
| 8     | 85.8         | 99.0      | 0       | 0      |
| 9     | 97.2         | 105.0     | 0       | 0      |
| 10    | 98.4         | 135.0     | 0       | 0      |
| 11    | 145.2        | 138.0     | 0       | 0      |
| 12    | 148.8        | 139.2     | 0       | 0      |
| 13    | 153.6        | 143.4     | 9.6     | 0      |
| 14    | 210.0        | 150.0     | 91.8    | 0      |
| 15    | 297.6        | 151.2     | 258.0   | 0      |
| 16    | 312.6        | 156.6     | 267.6   | 0      |
| 17    | 430.8        | 197.4     | 580.8   | 0      |
| 18    | 459.0        | 228.6     | 580.8   | 2.4    |
| 19    | 518.4        | 417.0     | 691.2   | 30.0   |
| 20    | 657.0        | 424.2     | 723.0   | 45.6   |
| 21    | 754.2        | 628.8     | 834.0   | 241.2  |
| 22    | 892.2        | 816.6     | 1416.0  | 915.0  |
| 23    | 1037.4       | 1119.6    | 1662.6  | 1213.8 |
| 24    | 1499.4       | 1495.8    | 1755.6  | 2861.4 |
| 25    | 2035.2       | 2625.0    | 2850.0  | 3811.8 |

Table A.5: For each house in the dataset, sum of maximum powers (W)  $\sum_i^{N_k} P_{max\ i}^{(k)}$  for all low-load appliances  $N_k$  in each category  $k$ . Cases where power is 0.0W are households with devices that have been removed from the modelling (for reasons specified in the text) or households that do not own any of these appliances.



# Bibliography

- [1] Department of Energy and Climate Change and Intertek Testing and Certification Ltd. *Household Electricity Survey*. 2010-2011.
- [2] Department of Energy and Climate Change. *Energy consumption in the UK*. 2015.
- [3] S. Jaboob. *On the ownership and use of residential electrical appliances*. PhD thesis. University of Nottingham, 2015.
- [4] A. Grandjean, J. Adnot and G. Binet. *A review and an analysis of the residential electric load curve models*. In: *Renewable and Sustainable Energy Reviews* 16.9 (2012), pp. 6539–6565.
- [5] M. Stokes. *Removing barriers to embedded generation: a fine-grained load model to support low voltage network performance analysis*. PhD thesis. De Montfort University, 2005.
- [6] J. Paatero and P. Lund. *A model for generating household electricity load profiles*. In: *International Journal of Energy Research* 30.5 (2006), pp. 273–290. ISSN: 0363-907X. DOI: [10.1002/er.1136](https://doi.org/10.1002/er.1136). URL: ([Link](#)).
- [7] A. Capasso, W. Grattieri, R. Lamedica and a. Prudenzi. *A bottom-up approach to residential load modeling*. In: *IEEE Transactions on Power Systems* 9.2 (May 1994), pp. 957–964. DOI: [10.1109/59.317650](https://doi.org/10.1109/59.317650). URL: ([Link](#)).

- 
- [8] J. Tanimoto, A. Hagishima and H. Sagara. *Validation of probabilistic methodology for generating actual inhabitants' behavior schedules for accurate prediction of maximum energy requirements*. In: *Energy and Buildings* 40.3 (2008), pp. 316–322. DOI: [10.1016/j.enbuild.2007.02.032](https://doi.org/10.1016/j.enbuild.2007.02.032). URL: [\(Link\)](#).
- [9] J. Widén, M. Lundh, I. Vassileva, E. Dahlquist, K. Ellegård and E. Wäckelgård. *Constructing load profiles for household electricity and hot water from time-use data-Modelling approach and validation*. In: *Energy and Buildings* 41.7-1 (July 2009), pp. 753–768. ISSN: 0378-7788. DOI: <http://dx.doi.org/10.1016/j.enbuild.2009.02.013>. URL: [\(Link\)](#).
- [10] J. Widén and E. Wäckelgård. *A high-resolution stochastic model of domestic activity patterns and electricity demand*. In: *Applied Energy* 87.6 (June 2010), pp. 1880–1892. ISSN: 03062619. DOI: [10.1016/j.apenergy.2009.11.006](https://doi.org/10.1016/j.apenergy.2009.11.006). URL: [\(Link\)](#).
- [11] I. Richardson, M. Thomson, D. Infield and C. Clifford. *Domestic electricity use: A high-resolution energy demand model*. In: *Energy and Buildings* 42.10 (2010), pp. 1878–1887. DOI: [10.1016/j.enbuild.2010.05.023](https://doi.org/10.1016/j.enbuild.2010.05.023). URL: [\(Link\)](#).
- [12] J. Page, D. Robinson, N. Morel and J.-L. Scartezzini. *A generalised stochastic model for the simulation of occupant presence*. In: *Energy and Buildings* 40.2 (2008), pp. 83–98. DOI: [10.1016/j.enbuild.2007.01.018](https://doi.org/10.1016/j.enbuild.2007.01.018). URL: [\(Link\)](#).
- [13] F. Haldi and D. Robinson. *Adaptive actions on shading devices in response to local visual stimuli*. In: *Journal of Building Performance Simulation* 3.2

- (June 2010), pp. 135–153. DOI: [10.1080/19401490903580759](https://doi.org/10.1080/19401490903580759). URL: [\(Link\)](#).
- [14] F. Haldi and D. Robinson. *Interactions with window openings by office occupants*. In: *Building and Environment* 44.12 (2009), pp. 2378–2395. DOI: [10.1016/j.buildenv.2009.03.025](https://doi.org/10.1016/j.buildenv.2009.03.025). URL: [\(Link\)](#).
- [15] A. T. Bharucha-Reid. *Elements of the Theory of Markov Processes and their Applications*. Courier Corporation, (2012).
- [16] A. K. Jain, M. N. Murty and P. J. Flynn. *Data Clustering: A Review*. In: *ACM Computing Surveys* 31.3 (1999), pp. 264–323. DOI: [10.1145/331499.331504](https://doi.org/10.1145/331499.331504). URL: [\(Link\)](#).
- [17] M. Ankerst, M. M. Breunig, H.-P. Kriegel and J. Sander. *OPTICS: Ordering Points to Identify the Clustering Structure*. In: *Proceedings of the SIGMOD International Conference on Management of Data*. ACM, (1999), pp. 49–60. ISBN: 1-58113-084-8. DOI: [10.1145/304182.304187](https://doi.org/10.1145/304182.304187). URL: [\(Link\)](#).
- [18] M. Ester, H.-P. Kriegel, J. Sander and X. Xu. *A density-based algorithm for discovering clusters in large spatial databases with noise*. In: *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining*. Vol. 96. 34. Portland, Oregon, USA, (1996), pp. 226–231.
- [19] J. Sander, X. Qin, Z. Lu, N. Niu and A. Kovarsky. *Automatic Extraction of Clusters from Hierarchical Clustering Representations*. In: *Advances in Knowledge Discovery and Data Mining*. Vol. 2637. Springer Berlin Heidelberg, (2003), pp. 75–87. ISBN: 978-3-540-04760-5. DOI: [10.1007/3-540-36175-8\\_8](https://doi.org/10.1007/3-540-36175-8_8). URL: [\(Link\)](#).

- 
- [20] D. R. Cox and D. Oakes. *Analysis of survival data*. Vol. 21. CRC Press, (1984).
- [21] R. Kohavi. *A Study of Cross-validation and Bootstrap for Accuracy Estimation and Model Selection*. In: Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2. San Francisco, CA, USA: Morgan Kaufmann Publishers Inc., (1995), pp. 1137–1143. ISBN: 1-55860-363-8. URL: [\(Link\)](#).
- [22] G. Box and G. M. Jenkins. *Time series analysis: forecasting and control*. Hoboken, New Jersey: John Wiley & Sons, (2015).
- [23] C.-Y. J. Peng and T.-S. H. So. *Logistic regression analysis and reporting: A primer*. In: Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences 1.1 (2002), pp. 31–70.
- [24] O. Vallis, J. Hochenbaum and A. Kejariwal. *A Novel Technique for Long-Term Anomaly Detection in the Cloud*. In: 6th USENIX Workshop on Hot Topics in Cloud Computing. USENIX Association, (2014). URL: [\(Link\)](#).
- [25] R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. 2015. URL: [\(Link\)](#).
- [26] D. Starnes, J. Tabor, D. Yates and D. Moore. *The Practice of Statistics*. W. H. Freeman, (2014). ISBN: 9781464108730. URL: [\(Link\)](#).
- [27] D. Murray, J. Liao, L. Stankovic and V. Stankovic. *Understanding usage patterns of electric kettle and energy saving potential*. In: Applied Energy 171 (2016), pp. 231–242. ISSN: 0306-2619. DOI: <http://dx.doi.org/10.1016/j.apenergy.2016.03.038>. URL: [\(Link\)](#).

- 
- [28] M. Shabanzadeh and M. Moghaddam. *What is the Smart Grid? Definitions, Perspectives, and Ultimate Goals*. In: 28th International Power System Conference. (2013).
- [29] IEA. *Energy Technology Perspectives 2012 - how to secure a clean energy future*. OECD/IEA, 2012. URL: [\(Link\)](#).
- [30] IEA. *Global EV Outlook 2016*. IEA, 2016. URL: [\(Link\)](#).
- [31] IEA. *Technology Roadmap: Solar Photovoltaic Energy*. IEA, 2014. URL: [\(Link\)](#).
- [32] IEA. *Technology Roadmap: Energy Storage*. IEA, 2014. URL: [\(Link\)](#).
- [33] IEA. *Global EV Outlook 2013*. IEA, 2013. URL: [\(Link\)](#).
- [34] M. Stifter. *IEA DSM Task 17 - Demand Flexibility in Households and Buildings*. IEA DSM, 2016. URL: [\(Link\)](#).
- [35] IEA. *Smart Grids in Distribution Networks - Roadmap, development and implementation*. OECD/IEA, 2015. URL: [\(Link\)](#).
- [36] M. Stifter and R. Kamphuis. *Roles and Potentials of Flexible Consumers and Prosumers*. IEA DSM, 2016. URL: [\(Link\)](#).
- [37] M. Beckmann. *The challenge of shifting peak electricity demand*. Centre for Carbon Measurement, 2013.
- [38] C. Frantz and SEE Change Institute. *Dimensions of Energy Behavior: Psychometric Testing of Scales for Evaluating Behavioral Interventions in Demand Side Management Programs*. Southern California Edison, 2015.
- [39] NERC. *2011 Demand Response Availability Report*. North American Electric Reliability Corporation, 2013. URL: [\(Link\)](#).

- 
- [40] R. Deng, Z. Yang, M.-Y. Chow and J. Chen. *A Survey on Demand Response in Smart Grids: Mathematical Models and Approaches*. In: IEEE Transactions on Industrial Informatics 11.3 (2015), pp. 570–582. ISSN: 1551-3203. DOI: [10.1109/TII.2015.2414719](https://doi.org/10.1109/TII.2015.2414719). URL: [\(Link\)](#).
- [41] J. S. Vardakas, N. Zorba and C. V. Verikoukis. *A Survey on Demand Response Programs in Smart Grids: Pricing Methods and Optimization Algorithms*. In: IEEE Communications Surveys & Tutorials 17.1 (2015), pp. 152–178. ISSN: 1553-877X. DOI: [10.1109/COMST.2014.2341586](https://doi.org/10.1109/COMST.2014.2341586). URL: [\(Link\)](#).
- [42] P. Faria, Z. A. Vale and J. Ferreira. *Demsi—A demand response simulator in the context of intensive use of distributed generation*. In: Systems Man and Cybernetics (SMC), 2010 IEEE International Conference on. IEEE. (2010), pp. 2025–2032. URL: [\(Link\)](#).
- [43] A.-H. Mohsenian-Rad and A. Leon-Garcia. *Optimal residential load control with price prediction in real-time electricity pricing environments*. In: IEEE transactions on Smart Grid 1.2 (2010), pp. 120–133.
- [44] L. Chen, N. Li, S. H. Low and J. C. Doyle. *Two market models for demand response in power networks*. In: Smart Grid Communications IEEE International Conference. (2010), pp. 397–402.
- [45] A. Y. Saber and G. K. Venayagamoorthy. *Unit commitment with vehicle-to-grid using particle swarm optimization*. In: PowerTech, IEEE Bucharest. IEEE. (2009), pp. 1–8.
- [46] M. Zheng, C. J. Meinrenken and K. S. Lackner. *Agent-based model for electricity consumption and storage to evaluate economic viability of tariff arbi-*

- trage for residential sector demand response*. In: Applied Energy 126 (2014), pp. 297–306. DOI: <http://dx.doi.org/10.1016/j.apenergy.2014.04.022>. URL: ([Link](#)).
- [47] K. Clement, E. Haesen and J. Driesen. *Coordinated charging of multiple plug-in hybrid electric vehicles in residential distribution grids*. In: Power Systems Conference and Exposition, IEEE/PES. IEEE. (2009), pp. 1–7.
- [48] W. Shi and V. W. Wong. *Real-time vehicle-to-grid control algorithm under price uncertainty*. In: Smart Grid Communications, IEEE International Conference. IEEE. (2011), pp. 261–266. URL: ([Link](#)).
- [49] I. Dusparic, A. Taylor, A. Marinescu, V. Cahill and S. Clarke. *Maximizing renewable energy use with decentralized residential demand response*. In: IEEE First International Smart Cities Conference (ISC2). IEEE. (2015), pp. 1–6. URL: ([Link](#)).
- [50] Z. Wen, D. O’Neill and H. Maei. *Optimal demand response using device-based reinforcement learning*. In: IEEE Transactions on Smart Grid 6.5 (2015), pp. 2312–2324. URL: ([Link](#)).
- [51] N. Gudi, L. Wang, V. Devabhaktuni and S. Depuru. *Demand response simulation implementing heuristic optimization for home energy management*. In: North American Power Symposium (NAPS), 2010. IEEE. (2010), pp. 1–6.
- [52] S. D. J. McArthur, E. M. Davidson, V. M. Catterson, A. L. Dimeas, N. D. Hatziargyriou, F. Ponci and T. Funabashi. *Multi-agent systems for power engineering applications - Part I: Concepts, approaches, and technical challenges*. In: IEEE Transactions on Power Systems 22.4 (2007), pp. 1743–1752. DOI: [10.1109/TPWRS.2007.908471](https://doi.org/10.1109/TPWRS.2007.908471). URL: ([Link](#)).

- 
- [53] M. Vasirani, R. Kota, R. L. Cavalcante, S. Ossowski and N. R. Jennings. *An agent-based approach to virtual power plants of wind power generators and electric vehicles*. In: IEEE Transactions on Smart Grid 4.3 (2013), pp. 1314–1322.
- [54] S. Iacovella, P. Vingerhoets, G. Deconinck, N. Honeth and L. Nordstrom. *Multi-Agent Platform for Grid and Communication Impact Analysis of Rapidly Deployed Demand Response Algorithms*. In: (2016).
- [55] T. K. Wijaya, D. Banerjee, T. Ganu, D. Chakraborty, S. Battacharya, T. Papaioannou, D. P. Seetharam and K. Aberer. *DRSim: A cyber physical simulator for Demand Response systems*. In: Smart Grid Communications, IEEE International Conference. IEEE. (2013), pp. 217–222. URL: [\(Link\)](#).
- [56] A. Dimeas and N. Hatziargyriou. *Multi-agent reinforcement learning for microgrids*. In: Power and Energy Society General Meeting. IEEE. (2010), pp. 1–8. URL: [\(Link\)](#).
- [57] K. Dittawit and F. A. Aagesen. *Demand side focused simulation platform for the evaluation of demand side management approaches*. In: IEEE/PES Innovative Smart Grid Technologies Conference Europe (ISGT-Europe). IEEE. (2014), pp. 1–6. URL: [\(Link\)](#).
- [58] J. Chapman. *Multi-agent stochastic simulation of occupants in buildings*. PhD thesis. University of Nottingham, 2016.
- [59] J. Page, D. Robinson, N. Morel and J.-L. Scartezzini. *A generalised stochastic model for the simulation of occupant presence*. In: Energy and buildings 40.2 (2008), pp. 83–98.

- 
- [60] S. Jaboob. *Stochastic Modelling of Occupants' Activities and Related Behaviours*. PhD thesis. University of Nottingham, 2015.
- [61] C. F. Reinhart. *Lightswitch-2002: a model for manual and automated control of electric lighting and blinds*. In: *Solar Energy* 77.1 (2004), pp. 15–28.
- [62] F. Haldi and D. Robinson. *Interactions with window openings by office occupants*. In: *Building and Environment* 44.12 (2009), pp. 2378–2395.
- [63] F. Haldi and D. Robinson. *Adaptive actions on shading devices in response to local visual stimuli*. In: *Journal of Building Performance Simulation* 3.2 (2010), pp. 135–153.
- [64] J. Chapman, P. Siebers and D. Robinson. *On The Multi Agent Stochastic Simulation Of Occupants In Buildings*. In: *Journal of Building Performance Simulation* (Manuscript submitted for publication).
- [65] J. Chapman, P. Siebers and D. Robinson. *Data Scarce Behavioural Modelling and the Representation of Social Interactions*. (Unpublished manuscript).
- [66] A. Sancho-Tomás, M. Sumner and D. Robinson. *A generalised model of electrical energy demand from small household appliances*. In: *Energy and Buildings* 135 (2017), pp. 350–366. DOI: <http://dx.doi.org/10.1016/j.enbuild.2016.10.044>. URL: (Link).
- [67] M. Le Thellier, A. Sancho-Tomás, G. Sousa-Ramirez and D. Robinson. *Creation of synthetic households for multi-agent stochastic simulation*. (In preparation).
- [68] C. J. C. H. Watkins. *Learning from delayed rewards*. PhD thesis. University of Cambridge, 1989.

- 
- [69] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction*. Vol. 1. 1. MIT press Cambridge, (1998).
- [70] J. Salom, J. Widén, J. Candanedo, I. Sartori, K. Voss and A. Marszal. *Understanding net zero energy buildings: evaluation of load matching and grid interaction indicators*. In: Proceedings of building simulation. Vol. 6. (2011), pp. 2514–2521. URL: ([Link](#)).
- [71] R. Luthander, J. Widén, D. Nilsson and J. Palm. *Photovoltaic self-consumption in buildings: A review*. In: Applied Energy 142 (2015), pp. 80–94.
- [72] J. Widén. *Improved photovoltaic self-consumption with appliance scheduling in 200 single-family buildings*. In: Applied Energy 126 (2014), pp. 199–212.
- [73] <http://www.pvparity.eu/>.
- [74] J. Dehler et al. *Self-consumption of electricity from renewable sources*. INSIGHT-E, 2015. URL: ([Link](#)).
- [75] N. Sajn. *Electricity Prosumers*. European Parliamentary Research Service, 2016. URL: ([Link](#)).
- [76] <https://www.gov.uk/feed-in-tariffs>.
- [77] <https://www.gov.uk/guidance/community-energy>.
- [78] T. Haggis. *Network Design Manual*. EON, 2006.
- [79] M. e. a. Stifter. *IEA DSM Task 17 - Roles and Potentials of Flexible Consumers and Prosumers*. IEA DSM, 2016. URL: ([Link](#)).
- [80] B. Stott. *Review of load-flow calculation methods*. In: Proceedings of the IEEE 62.7 (1974), pp. 916–929.

- 
- [81] J. E. Van Ness. *Iteration methods for digital load flow studies*. In: Transactions of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems 78.3 (1959), pp. 583–586.
- [82] J. Liu, M. Salama and R. Mansour. *An efficient power flow algorithm for distribution systems with polynomial load*. In: International Journal of Electrical Engineering Education 39.4 (2002), pp. 371–386. URL: [\(Link\)](#).
- [83] W. H. Kersting. *Distribution system modeling and analysis*. CRC press, (2012).
- [84] A. Von Meier. *Challenges to the integration of renewable resources at high system penetration*. California Energy Commission, 2014.
- [85] S. Cao and K. Sirén. *Impact of simulation time-resolution on the matching of PV production and household electric demand*. In: Applied Energy 128 (2014), pp. 192–208. URL: [\(Link\)](#).
- [86] C. Bucher, J. Betcke and G. Andersson. *Effects of variation of temporal resolution on domestic power and solar irradiance measurements*. In: IEEE PowerTech Grenoble Conference. IEEE, (2013), pp. 1–6. DOI: [10.1109/PTC.2013.6652217](#). URL: [\(Link\)](#).
- [87] A. J. Urquhart and M. Thomson. *Impacts of Demand Data Time Resolution on Estimates of Distribution System Energy Losses*. In: IEEE Transactions on Power Systems 30.3 (2015), pp. 1483–1491. DOI: [10.1109/TPWRS.2014.2349157](#). URL: [\(Link\)](#).
- [88] A. Fazeli, M. Sumner, E. Christopher and M. Johnson. *Power flow control for power and voltage management in future smart energy communities*. In: (2014). URL: [\(Link\)](#).

- [89] D. E. Moriarty, A. C. Schultz and J. J. Grefenstette. *Evolutionary algorithms for reinforcement learning*. In: *Journal of Artificial Intelligence Research* 11 (1999), pp. 241–276.