THE UNIVERSITY
of EDINBURGH

Analysis, Modelling and State Estimation for Large Scale Electric Demand Response

Alexandros Kleidaras

THE UNIVERSITY
of EDINBURGH

Thesis submitted in fulfilment of the requirements for the degree of
Doctor of Philosophy to the University of Edinburgh — 2021

Declaration

I declare that this thesis has been composed solely by myself and that it has not been submitted, either in whole or in part, in any previous application for a degree. Except where otherwise acknowledged, the work presented is entirely my own.

Alexandros Kleidaras

February 2021

Abstract

The need for additional reserves increases alongside the intermittency of generation and whilst rotating (conventional) generation is replaced, the system's inertia reduces and balance volatility increases. Conceptually, any regulation measure from the "generation side" has an equivalent countermeasure from the "demand side". One of the emerging technologies to provide such balancing services is Demand Response (DR). DR is commercially used, mainly via industrial loads combined with small scale diesel and gas generators. However, there is a lot of potential for DR from residential and commercial loads that remains untapped due to implementation costs, lack of technology expertise, load pattern complexity and the need to simultaneously control numerous sources.

The main focus of this thesis is to explore the potential of loads, mainly residential and small commercial, to provide DR services and develop methods focused on accuracy for the most challenging services (frequency regulation), whilst aiming for minimal infrastructure and implementation costs. The main points include analysis of common residential and commercial loads for DR services, focusing on thermostatically controlled loads (TCLs). TCLs are thermal loads which operate via thermostats on a duty cycle (on and off state), between two temperature settings in order to maintain an average set temperature. They use electricity as a primary energy source or for their control and pumps.

The next part includes analysis and creation of realistic bottom up models to study aggregated behaviour of TCLs during DR actions, as well as the effect of external factors. Afterwards, a distributed State Estimation algorithm is proposed to increase accuracy of aggregated models and track aggregation models from limited information. A new aggregation framework is proposed, specifically designed for heterogeneous populations, whilst being universal for all TCL types.

As such, different TCL types can be aggregated together (e.g. cooling and heating).

The results of this thesis show that with proper aggregation modelling, state estimation and dynamic updating in time, accuracy of stochastic aggregated models is improved compared to existing frameworks without the need for expensive thermal sensors. This suggests that with relatively limited information the use of residential and commercial TCLs for DR balancing services, is feasible.

Acknowledgements

Firstly, I would like to thank my supervisors, Dr Aristides Kiprakis and Prof John Thompson, for giving me the opportunity to be part of the ADVANTAGE FP7 project, for their continuous assistance and for having faith in me to take my own path. Also, I would like to thank Dr Aristides Kiprakis separately for his understanding and support during difficult times.

Special thanks to Prof Dejan Vukabratovic and Dr Mirsad Cosovic for assisting me to cover a lot of material, regarding network theory, in a short period of time and welcoming me in Novi Sad for my secondment. Furthermore, I would like to acknowledge the advice, comments and support of Dr Konstantinos Markakis.

My years in Edinburgh and PhD experience would have not being the same without the moral support of the people close to me, Dr Manolis Falagkaris, Dr Dimitris Tsikritsis, Dr Konstantinos Markakis, Mr Kenny Duffy and Prof Slawek Staworko. Last but certainly not least, I want to thank my family for their support and understanding throughout the many years where I stayed away from them in pursuit of my dreams; I promise to devote more of my time to my family from here on.

Nomenclature

Symbol	Unit	Description
β	^oC	Constant
γ	$^oC \cdot min^{-1}$	constant
δ	^oC	Temperature dead-band (of TCL)
ϵ	-	Constant
ζ	-	Constant
θ	^oC	Temperature
θ_a	^oC	Ambient temperature (of TCL)
θ_e	^oC	Temperature gain (external factor)
θ_g	^oC	Temperature gain during *on* state
θ_{off}	^oC	TCL switch off temperature
θ_{on}	^oC	TCL switch on temperature
θ_{set}	^oC	Temperature set-point
I	-	Temperature state bin
κ	-	Constant
λ	min^{-1}	Thermal time constant 1/RC (R, C expressed in min)
μ	-	1 for *on* state, 0 for *off* state
Ξ	-	Temperature state bin
ϱ	$^oC/min$	Rate of temperature change (rate of bin state progress)
ϱ_{on}	$^oC/min$	ϱ for *on* state
ϱ_{off}	$^oC/min$	ϱ for off state
σ	-	Variance
Φ	-	Gaussian Distribution with 0 mean

Symbol	Unit	Description
τ	min	Simulation time steps
υ	-	1 for external factor, 0 otherwise
A	-	Transition probability matrix
a	-	Constant $e^{-\tau/RC}$
B	-	Matrix
C	$kWh/^oC$	Thermal capacitance
c_p	Watts	Constant
D	-	Duty cycle (t_{on}/T)
E	m^2	Heat transfer surface area
f	-	Factor node
g	-	Constant
h	$W/(^oC \cdot m^2)$	Newton's heat transfer coefficient
H	^oC	Temperature change due to external factors
k	-	Iteration
m	-	BP message
$\mathcal{N}$	-	Gaussian Distribution
N_{TCL}	-	TCL population
N_{on}	-	TCL population in *on* state
N_{off}	-	TCL population in off state
N_{bin}	-	Number of bin states
P	Watts	Power rate
P_{ave}	Watts	Average power (of TCL population)
P_{tot}	Watts	Total aggregated power (of TCL population)
Pr	$\Re \in [0, 1]$	Probability
Q	Joules	Heat transfer
R	$^oC/kW$	Thermal resistance
sw_{off}	-	Fraction of units switching off in 1 time step
sw_{on}	-	Fraction of units switching *on* in 1 time step

Symbol	Unit	Description
t_{off}	*minutes*	*off* state duration
t_{on}	*minutes*	*on* state duration
T	*minutes*	Cycle duration ($t_{on} + t_{off}$)
$\mathcal{U}$	-	Uniform Distribution
$\mathbf{u}$	-	Control input vector
$\mathbf{v}$	-	Noise vector
$\mathbf{w}$	-	Noise vector
W	-	Wiener process
w_f	$\Re \in [0, 1]$	weight factor, freezer compartment
w_r	$\Re \in [0, 1]$	weight factor, refrigerator compartment
$\mathbf{x}$	-	State vector
X_{off}	-	Fraction of TCL population in off state bins
X_{on}	-	Fraction of TCL population in *on* state bins
$\mathbf{x_s}$	-	Steady state vector
$\mathbf{y}$	-	output (demand/consumption) vector
z	-	BP message mean

List of Terms and Abbreviations

Abbreviation	Description
AMI	Advanced Metering Infrastructure
BP	Belief Propagation
CAPEX	Capital Expenditure
CFPE	Coupled Fokker-Planck Equations
CHP	Combined Heat and Power
CI	Confidence Interval
CPP	Critical Peak Pricing
DECC	Department of Energy & Climate Change
DG	Distributed Generation
DR	Demand Response
DSM	Demand-side Management
DTU	Demand Turn Up
EFR	Enhanced Frequency Response
EKF	Extended Kalman Filter
EV	Electric Vehicles
ESS	Energy Storage System
FFR	Firm Frequency Response
FG	Factor Graph
FR	Fast Reserve
GBP	Gaussian Belief Propagation
HVAC	Heating Ventilation and Air Condition
ICT	Information and Communication Technology
ID	Identification
KF	Kalman Filter
MC	Monte Carlo

Abbreviation	Description
MPC	Model Predictive Control
ODE	Ordinary Differential Equations
OPEX	Operational Expenditure
PMF	Probability Mass Function
PTR	Peak Time Rebate
RES	Renewable Energy Sources
RMS	Root Mean Square
RoCoF	Rate of Change of Frequency
RTP	Real-time Pricing
TCL	Thermostatically Controlled Load
TOU	Time of Use
TPM	Transition Probability Matrix
TSO	Transmission System Operator
S-PDE	Stochastic Partial Differential Equation
STOR	Short-term Operating Reserves
VPP	Virtual Power Plant
ZOH	Zero-order Hold

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Background and Motivation

Electricity networks around the globe are facing rapid changes driven by the consensus to reduce greenhouse emissions. Specific targets have been set and agreed by developed countries, starting with the Kyoto Protocol (signed 1997), United Nations Climate Change Conferences are held annually to assess progress in dealing with climate change. In Power Systems this is reflected by integrating more Renewable Energy Sources (RES), mainly intermittent RES. RES are energy resources naturally replenished, such as sunlight, wind, rain, tides, waves, and geothermal heat. RES are commonly used for electricity generation and intermittent RES describes sources such as wind, wave, tidal and solar power whose primary energy input is uncontrollable. Distributed Generation (DG) is a significant part of it as well, which includes besides RES, co-generation units, micro-turbines, biomass, geothermal plants, hydrogen plants, energy storage and more. The term DG describes generation connected in Distribution Networks instead of Transmission Networks, close to consumption, which is usually of much

smaller size. DG units have limited balancing capabilities (i.e. small inertia) and limited support capabilities for Transmission Networks' stability.

Among intermittent RES, the highest integration rates are shared between wind and solar power with plans to further increase their penetration in grids (such as in UK [1, 2]). These are characterised by inherent variability and no inertia which have detrimental effects to the grid's stability [3]. Intermittent RES are non-dispatchable, resulting in variable generation output depending on weather conditions, which in turn results in dispatch of fast ramping conventional generators and more frequent mismatches between supply and demand due to forecast errors.

Traditionally, Power System operation assumes that electricity generation is fully dispatchable and controllable, able to meet the systems requirements whilst maintaining stability. The key technology for stability is the rotating synchronous generators, which store kinetic energy, giving the system rotational inertia, the cornerstone of frequency regulation [3]. Power system inertia is defined as the ability of a power system to oppose changes in system frequency due to the kinetic energy stored in Synchronous Generators' rotating masses. The electrical rotating frequency of those is normally 50Hz or 60Hz, which is equal to the mechanical frequency times the pair of poles of the Synchronous Generator. Mismatch in real time between Power Demand and Supply causes a change in frequency. Frequency deviations are absorbed via this inertia, and the system's ability to do so is directly linked to the sum of its sources' inertia. Deviations should be kept within small limits (1% for UK National Grid [4]) of the nominal value, to avoid harmful vibrations in generators and load shedding. Larger deviations may lead to disconnection of generators for safety, which may ultimately lead to a chain reaction if unresolved and cause a black-out with massive economic and technical consequences. When the system's frequency diverges beyond set limits, Balancing Services are deployed to provide power stability (increase or decrease)

to ensure the security and quality of electricity. Load manipulation can be used in a similar manner to provide such Balancing Services, where for example instead of turning on a Back-Up Generator for a few minutes, loads can be used to reduce demand for the same time. UK's and EU's Balancing Services also include such schemes under the term Demand Response [5].

The intermittent RES impact in power system's stability is significant because they do not provide rotational inertia, whilst replacing conventional generators and their rotating machinery. Though lately, some of the larger scale RES (Wind Farms and Photovoltaic Farms) include small virtual inertia via the use of power electronics, energy storage and/or adjusting wind blades [6, 7]. It is important to note that in the case of pitch angle control for wind blades, there is an inherent reduction in efficiency. The traditional assumption that grid's inertia is sufficiently high with only small variations over time is thus not valid for power systems with high RES shares. This has implications for frequency dynamics and power system stability and operation. Frequency dynamics are faster in power systems with low rotational inertia, making frequency control and power system operation more challenging [3]. Not only does high penetration of RES and DG, reduces the system's inertia but also amplifies active power mismatches and thus frequency deviations as mentioned above. A system inertia's is commonly measured as the time duration during which the system can meet its demand solely via its stored (kinetic) energy. Studies indicate 3 primary drawbacks in power systems with high penetration of RES [6]:

1. Reduction in system inertia depending on intermittent RES generation,

2. Variable system inertia, higher frequency deviations, larger and unpredictable power deviation in small intervals (ramping of RES and unobservable DG)

3. The assumption of having the same inertia over a multi-area system is no longer valid

Research in the German Power System shows that during high intermittent RES generation the inertia drops from about 6s to 3s [6], which greatly diminishes the system's ability to absorb frequency deviations. Adding the fact that RES contribute to the increase of active power mismatch between supply and demand, therefore amplifying frequency deviations, the system is at a higher risk during periods of high RES generation. System operators are looking for new ways to increase system reliability and introduce new balancing services to tackle this issue. These services are used to balance demand and supply, and to ensure the security and quality of electricity supply across the transmission system. A metric of demand and supply balance is frequency, when the Rate of Change of Frequency (RoCoF) is zero then demand and supply are equal. An analysis by Imperial College London and Electricite de France concluded that the cost of reserve and response services in 2030 may increase up to 1.23 B£ and 1.04 B£, respectively for high intermittent RES penetration [1].

Currently, Transmission System Operators (TSOs) rely on Primary, Secondary and Tertiary reserves in the form of balancing services to stabilize frequency after a disturbance. Providers can offer one of these or a combination. In the UK, the Primary response provided must be provided within 10 seconds of an event, which can be sustained for a further 20 seconds. The Secondary response must be provided within 30 seconds of an event, which can be sustained for a further 30 minutes. The Tertiary response comes in the form of Fast Reserve (FR) and Short-term Operating Reserve (STOR). FR must respond within 2 minutes and sustained for at least 15 minutes, STOR must respond within 240 minutes (preferably though 20 minutes) and sustained for 2 hours at least.

In an event of fault which results in loss of infeed power, the Primary and Secondary response are designed to maintain frequency within the specified limits. Tertiary response's purpose is to assist in bringing the frequency back to nominal value. Conceptually, any regulation measure from the "supply side" has an

equivalent countermeasure from the "demand side". The idea of using loads for regulation has its roots in 1980's [8, 9]. Lately, such concepts have taken form under the DR and are been integrated as vital parts of Power Systems operation, not only for stability but also for regulation and energy management. Demand Response in this context means a change in demand (i.e. a load such as a refrigeration unit) in response to an external signal. That signal can be an automated command or a response to a change in price which normally does not require the intervention of the end user. DR can participate in a number of balancing services, which fall under the categories described above. The exact services in which DR can participate in UK will be discussed in Chapter 2.

DR is part of Demand-side management (DSM), not to be confused as one and the same. DR refers to the modification of demand for various reasons, such as balancing services described above or dynamic pricing, in response to a signal/event. DSM is a general term which describes the optimal management of assets from the demand side for some objective. Energy efficiency is part of DSM but not DR (Figure 1.1). For instance, optimising the temperature set point settings of refrigeration units and space heating in supermarkets, during day and night, in order to avoid conflicting settings and energy waste. A DR example is a case of a grid fault, in which demand responds to an external signal for 30 seconds and provides a balancing service to the grid. This is part of DSM but it is not energy efficiency. Main DSM categories are summarized in Figure 1.1 [10].

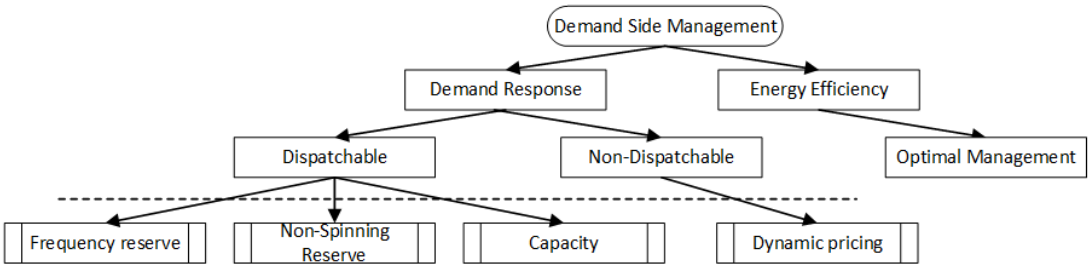


Figure 1.1: Demand Side Management categories

DR in the context of this thesis: There are limits to what can be achieved on the supply side, especially given the higher penetration of RES, because of the physical limits of generating units, but also cost. Demand can at times be greater than the capacity of all the available power plants put together, such as in contingency cases (infeed loss). This is the reason for Reserves and Balancing Services' existence. Note that Balancing Services are Power Services, to manage frequency, short in duration and fast in response, whilst Reserves are Energy Services, with slower response but higher duration, which usually follow after Balancing Services have been dispatched. DR can be provide Balancing Services (response to a contingency) and in some cases the shorter duration Fast Reserves. This type of DR requires specific response in Power (MW), in short time frames (usually a few seconds) and with specific requirements on Ramping Rate (accuracy is important). All Balancing Services providers are coordinated (dispatched) by the TSO in terms of response time, power, ramping rates etc. Their market is usually on a tender basis with fixed payments for availability, whether utilized or not. This type of DR is the main motivation and focus for this thesis.

Another type of DR is based on dynamic pricing/tariffs. A simple example of that is the usual dual tariff system, where different prices are used overnight and during the day. New types of more dynamic pricing are being introduced the last few years with the adoption of smart meters. This type of DR is indirect, through pricing, to mould the shape of demand, in order to reduce peak demand or follow RES output etc. It is energy arbitrage, not to be confused with energy efficiency or curtailment, where demand is deferred in other periods of time. It is energy and cost driven, has relatively longer duration and no high accuracy/response requirements etc., but based on consumers' behaviour.

1.2 ADVANTAGE Project

The research presented in this thesis is a part of a wider scale, inter-disciplinary and inter-sectoral project, ADVANTAGE FP7 (Advanced Communications and Information processing in smart grid systems) which was funded by the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 607774.

The main research focus was on smart grid operation and particularly on providing architectural solutions for smart grid systems with the contribution of innovative information and communication technology (ICT) solutions. ADVANTAGE is divided into 4 distinctive but correlated work packages [11]:

"Work package 1: Smart Homes: progressing the development of ICT solutions and applications in household environments to intelligent, customer-friendly, efficient and incentive-responsive home energy management solutions for smart grid household consumers.

Work package 2: Neighbourhood/Industrial Area Networks: providing efficient wide-area ICT solutions for advanced smart grid data exchange, gathering, processing and decision making in larger consumer conglomerations and environments.

Work package 3: Micro Grids: developing ICT solutions for challenging problem of integration and distributed control of micro grids in smart grids.

Work package 4: Intelligent Distribution Networks: advancements towards intelligent distribution network that will maintain efficient distribution management, load clustering, demand side management, distributed micro grid control, and

two-way communication with customers and micro grids through Advanced Metering Infrastructure (AMI)."

The work presented in this thesis falls primarily under Work package 4. It focuses on analysis and modelling of residential loads for Demand Response. These loads are connected in Distribution Networks and when aggregated in large numbers, through AMI, may be used to provide ancillary services to the system. This thesis provided input to Work packages 1 and 4, through demand profiles on household level, in a break-down analysis per load type. Output from Work packages 1 and 2 regarding ICT solutions, specifically capabilities of smart meters and power metering, was used in the assumptions of this thesis. The outcome of this thesis shows that TCLs in Distribution Networks can be part of DR.

Thermostatically Controlled Load definition

A thermostat is a component which senses the temperature of a physical system and performs actions so that the system's temperature is maintained near a desired setpoint, with a specific temperature deadband. Thermostats are used in any device or system that heats or cools to a setpoint temperature, examples include building heating, central heating, air conditioners, HVAC systems, water heaters, refrigeration units and medical and scientific incubators. In literature, these devices are often broadly classified as thermostatically controlled loads (TCLs). Thermostatically controlled loads comprise roughly 50% of the overall electricity demand in the United States [12]. The "closed loop" within that deadband, with switch on and switch off temperatures, as well as the heat exchange time constants of those devices (due to thermal resistance and capacity characteristics) create an equivalent of "thermal storage". As such, TCLs can be seen as having an equivalent "state of charge" to batteries, where switch on temperature would be equal to 0% "state of charge" and switch off equal to 100% "state of charge", which prompted authors in [13] to model TCLs as "leaky batteries".

1.3 Smart Grid Technology and DR

The Smart Grid concept is described as "*an electricity network that can intelligently integrate the actions of all users connected to it - generators, consumers and those that do both - in order to efficiently deliver sustainable, economic and secure electricity supplies*" [14], based on intelligent monitoring, communication and control. Low cost computing and communication hardware is what enables the Smart Grid and therefore automated DR strategies can be developed. T. Bigler et al. describe the hardware components required to implement smart devices with bidirectional communication through internet with electricity providers [15]. Smart appliances are becoming commercially available from a variety of producers in UK with a 20% penetration.

There are various studies on DR's potential, such as providing primary response services by electric vehicles [16], energy arbitrage to defer demand to times of lower pricing [17], reactive power regulation from distributed energy storage [18] (note that EVs and distributed energy storage, especially in form of batteries, can technically provide the same services). DR is estimated to be able to provide up to 54% of the operating reserve requirements of the power system depending on the time of day [19]. This context motivates the development of frameworks such as the one described in this thesis, to coordinate large populations of such devices to provide DR services beyond the old simplistic load-shedding paradigm [20]. Smart Grid concepts are about non-intrusive DR, which aims not to disturb the quality of service provided and/or the comfort of the consumer. The flexibility potential of DR, seen as the maximum amount of time and power increase/decrease within the comfort requirements of the user, varies during the day, and the potential for increasing or decreasing the power consumption is in general not equal, as shown by R. D'hulst et al in a pilot project with 186 household participants in the Flanders region in Belgium [21]. Major benefits of DR:

1. System stability: the ability to regulate frequency enhances the system's inertial-primary response increasing the system security [19, 22]

2. Cost effective: reduces need for on-line part-loaded generators, also increases the efficiency of the system. DR may be used to enhance the dispatch of generators or to follow intermittent RES production or replace fast ramping reserves [1, 17].

3. Environmentally friendly: by replacing conventional generators for balancing services but also assisting in higher levels of RES penetration [1, 17].

Some of the most prominent loads to provide DR are the thermostatically controlled loads (TCLs) and electric vehicles (distributed batteries as well). A lot of research is focused on TCLs due to their thermal storage capabilities; they represent a substantial base load connected 24/7, ubiquitously around the grid, thus reliably available for DR at all times. Their actual ability to provide balancing, reserve or arbitrage services highly depends on device-level temperature constraints and intrinsic thermal inertia. In recent years other loads have been considered too, such as EVs [23], distributed batteries [7] and wet loads [19]. Similar models developed for TCLs can be used for other types of loads with storage capabilities [23]. Another interesting DR concept comes from voltage regulation, which affects active power consumption (short-term). Trials of this concept were conducted by electricity north west, using online tap changing transformers [24].

1.4 DR Services: Literature Review of State of the Art Models

The study of DR for services has its roots in studies of loads with some form of storage capabilities in '80-'90. The original target was to model the unexpected increased demand and power oscillations after load shedding and interruptions (brown-outs and black-outs). A significant part of that came from TCLs, due to starting all together (synchronized). In early works of Ihara and Schweppe [8] and Malhame and Chong [9] studied this behaviour of TCLs. The proposed TCL model assumed a "snapshot" of TCLs current state with some noise to add stochasticity, whilst using first order ODE to describe TCLs' dynamics. A TCL's thermal storage capability stems from its operational cycle, switching between two states, "on" and "off". Its objective is to maintain temperature within a small dead-band and control can be used to alter it without exceeding said dead-band (non-intrusive control) (Figure 1.2) [25].

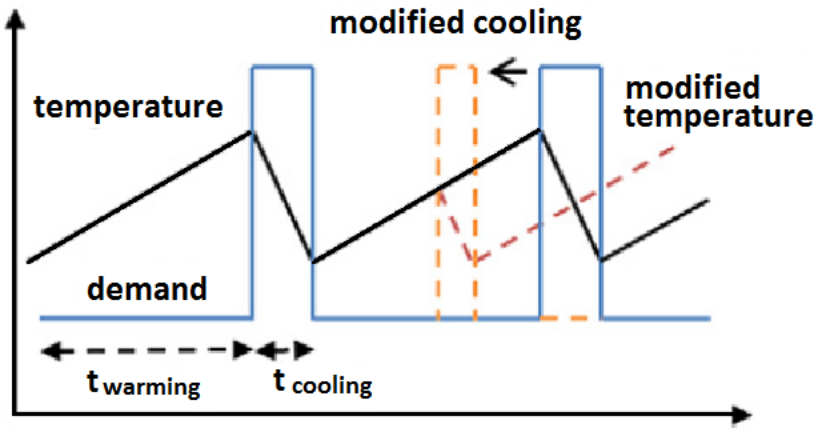


Figure 1.2: Cooling cycle example with and without DR. Under normal operation the TCL is alternating between power on and power off cycles. When DR is used this cycle is altered. This showcases the concept of short-term power manipulation.

Several works in literature describe the temperature dynamics of TCLs by means of a first order ODE [26, 13, 27, 28, 29]. The main advantages of that approach are the reduced computational burden in simulations involving a large number of devices and the simplicity in modelling. Higher order formulations are used in [30, 31, 32] for higher accuracy; for instance, the use of second order models allows to model the dynamics of multi-compartment TCLs (e.g. freezer and refrigeration compartments) or those affected by various factors [31]. Though, control design and computation are not as trivial as in the case of first order ODE models, while the different modelling also means inability to cluster them with those of lower order. The last drawback is very important when developing statistically based frameworks.

Totu et al. [27] focused on developing stochastic control algorithms assuming a centralised and a distributed approach to provide short-term reserves. Their work is based on the classic first order ODE model (Ihara & Schweppe [8]) for individual units and Coupled Fokker-Planck Equations (CFPE) (Malhame & Chong [9]) to model aggregated population behaviour. For the distributed approach with local controllers, some coordinator (observer) is required to analyse the available units for DR, as well as power. A stochastic signal is computed centrally and then sent [33].

More specifically, 2 control approaches are proposed, initially switch fraction and later switch rate, a modification of the first one. The switch fraction signal is sent every few seconds, composed of two rational numbers e_0, e_1 (probabilities). The signal triggers a percentage e_0 of units in "on state" to switch off (similar for e_1 and units in "off state"). While the switching fractions are given at the population level, the actual switch is decided at the unit level based on the result of a binomial trial with success rate equal to the broadcast fraction value. For a large enough population, by the law of large numbers, the response of the population will be close to the desired one [34]. The switch rate actuation introduces small

stochasticity to the time of activation, in order to avoid synchronization across the population [35].

The main issues that arise from CFPE are due to the fact that they can only be assumed for a completely homogenous population, which introduces errors and is unrealistic. Moreover, in such models detailed information of units' characteristics are required. The issue of homogeneity and approximate CFPE solutions was address by D. Callaway in [28]. Callaway worked on the diffusion approximation proposed by Malhamé and Chong [9] and derived a new exact solution to a well known hybrid state aggregated load model. The proposed control algorithm was based on set-point actuation, in which the set operational point is switched through an external signal. Additionally, the effect of low heterogeneity was examined, it was concluded that small heterogeneity improves the dynamics associated with control strategies and that some amount of heterogeneity is required for realistic dynamics, though the CFPE model's accuracy is reduced [28]. This implies that higher heterogeneity, as the one observed in real world, would have significant errors, as such, clustering would be required in smaller groups, increasing stochasticity (Kolmogorov's Law of Large Numbers). Last but not least, set-point actuation can be intrusive, unless proportionally switched with small changes within the operational dead-band.

One model developed by Imperial College London focuses on a decentralised framework, where units can be seen as agents belonging to a population (distribution) [36]. It is based on the classic first order ODE model (Ihara & Schweppe [8]) and CFPE (Malhame & Chong [9]) for aggregated population behaviour [37]. Trovato et al modified the model for frequency control, short-term reserves and a variation of it for energy arbitrage as well [38]. The model was developed for populations with low heterogeneity (independent parameter variation by $\pm 15\%$) and focused mostly on control [26].

Two types of controllers were developed for that purpose, a linear controller and a pre-set shape controller [22]. The first one controls aggregated power as a linear function of locally measured frequencies and their rates of change (RoCoF), while the second one follows a predefined reference power profile depending on the estimated infeed loss. As mentioned though by the authors "realistic and effective power system applications proposed in this paper may require a communication infrastructure". Accurate calculation of frequency and RoCoF is required at device level and a sliding window of 500ms is suggested [22]. The model also assumes knowledge of system's damping (MW/Hz), droop characteristics (MW/Hz), base power (MW) and system inertia (s). Due to these restrictions, they conclude "The risk of under- or over-response from the TCL population rises, affecting the demand-side flexibility and controllability", a potential alternative is thus suggested; a semi-autonomous mode between a central point and the agents, by using a communication protocol to update these parameters in hourly intervals [22].

The drawbacks of the approach of such a decentralised method are:

1. Lack of knowledge regarding the availability of other agents in the system (hence response); agents must have prior knowledge of installed and available loads for DR [2].
2. Sending system information to agents every hour, as suggested, does not capture dynamics of the shorter time scales (e.g. generation changes every 30 minutes via day ahead dispatch, 5min changes via intraday dispatch etc.).
3. Most importantly on-fault and post-fault systems conditions are different and unknown by the agents. These include the system's damping (MW/Hz), generators' droop characteristics (MW/Hz), base power (MW), system inertia (s). As such the agents' response is likely to be inaccurate.
4. Some of the issues seen in the models created by Totu et al. and Callaway, due to the use of first order ODE and CFPE models.

5. Finally, requirements of local processing, control and execution, means there has to be a device capable of those plus measuring the rate of frequency change (sub-second sampling) accurately. This would add a significant cost per appliance, when considering a population in the order of hundreds of thousands it can be prohibiting.

A final remark for the above models is that they require knowledge of physical TCL (load) parameters and precise sensors. As stated by Vinther et al.: "A model abstraction is required with a level of information that can represent the essential characteristics of the subsystems with minimum dependence on the number and physical dimensions of specific components/units used in each subsystem" [39].

S. Koch et al. adopted an alternative approach, instead of CFPE aggregation models, using transition probability matrixes (TPM) to describe the aggregated behaviour of TCLs [29]. The individual TCL is described by the classic first order ODE ([8]). There are a few advantages with this approach. First of all, the population does not have to be assumed homogeneous, such as in the case of CFPE where such an assumption is necessary to derive a solution. A homogeneous population assumes identical loads (appliances) and identical use pattern, which is unrealistic. In contrast to assuming homogeneous populations (identical TCL parameters) it is more realistic to assume Gaussian distribution of parameters, as TCLs build for similar purpose should not vary widely, e.g. refrigerators will have similar operating set-points (subject to human preference) and insulation etc. In statistic models of unknown parameter distribution it is common to start with Gaussian (or similar) distributions of parameters and work towards more suitable ones, once more real-world data is available. Another advantage is that it is based on calculating approximately transition probabilities to represent the aggregated behaviour (response) of the population to external signals (or normal operation), instead of detailed characteristics. The proposed approach to derive said transition matrix though was based on heterogeneity

in only one parameter. Moreover, the authors concluded that calculating the transition matrix with heterogeneity in more than one parameter is impractical if not infeasible in some cases. Additionally, the resulting transition matrix has the same (average) values for all "on" state bins and the same applies for "off" state bins, which as shown later in this thesis but also in literature is not the case in reality. The control proposed is practically the equivalent of fraction actuation [34], where a portion of the population (probabilistically) switches from one state to the other.

Afterwards, Mathieu & Callaway in [40] introduced system identification, state estimation and feedback control, in order to improve accuracy and feasibility. Different levels of infrastructure and communication were explored from full state information of all units to only a subset. Assumptions include knowledge of total aggregated power consumption (substation or similar level) and varying levels of knowledge of TCLs parameters, ambient temperatures, and dead-bands (10%, 30%, 100% of the population). Discretizing the dead-band in bins, calculating the transition probability from one bin to another for a given time step and knowing the thermal state of TCLs (i.e. in which bin they are at every time step), is the backbone of this approach. Descretizing in more bins results in higher accuracy, but that also requires high precision sensors. Knowing the setpoint (and temperature deadband) is simple with smart appliances, as well as approximate ambient temperature, but very precise thermal state close to real time within $< 0.1\ ^oC$ or higher accuracy, requires very precise and expensive thermal sensors. Normal TCLs are not equipped with such sensors and it would be costly to do so for millions of domestic and commercial TCLs, let alone not practical for existing ones, which is the main obstacle for the realisation of [40] or any approach that depends on accurate knowledge of thermal states. The population's aggregated power consumption on substation level requires disaggregation techniques since it is part of the total power (mixed with other loads) [41]. To that end, a few

disaggregation techniques were investigated on feeder level. Disaggregation on substation level is also possible, yet feeders have less loads and thus discrete states are easier to distinguish resulting in higher accuracy. The total daily demand varied roughly between 4MW and 12MW, with air conditioning being about 1MW to 5 MW (the target of disaggregation) and lowest root mean square errors were between 0.2MW and 0.3MW [42]. Kalman Filter (KF), a centralised algorithm, was used for state estimation and the transition matrix was assumed known and accurate, which is not the case in reality. As such, a joint parameter-state estimation was also investigated with Extended KF, however as the authors stated it was unable to converge and requires real time measuring or some form of deriving data off-line (for individual unit consumption) [40].

Individual unit modelling instead of aggregated population models were explored by Vrettos et al., for electric water heating. To eliminate inherent numerical diffusion and inaccuracy in first-order ODE, the second-order, three level finite difference Crank-Nicolson scheme was utilized [32]. Afterwards, an MPC (model predictive control) approach was used instead aiming to optimise consumption on household level [43, 44]. These models are effective for household level and act as agents, but are not scalable to large populations (hundreds of thousands to millions) for provision of services such as frequency control or reserves. In subsequent work, the use of aggregation models with state estimation is similar to one used in [40], but with the additions of time delay and noisy estimates of the aggregated population's demand and infrequent perfect measurements of each unit's state [45].

Another similar decentralised model to Trovato's [22], is investigated by Ziras et al., considering start-up consumption spikes (similar approach to Totu's et al. work in switch rate actuation [35]). The control was based on dynamic dead-band manipulation (similar to Trovato et al. [22]) while respecting original dead-band limits. The decentralised model has the same problems as described earlier

(unknown DR power, population's state/availability, power system's condition etc.). To overcome this, the model assumes knowledge of system's damping (MW/Hz), generators' droop characteristics (MW/Hz), base power (MW) and system inertia (s) with a 10 second sampling interval [46]. Such knowledge of power system's state, accurately every 10 seconds, means some form of central information exchange. Still, in an event where Primary Frequency Control is required, those values will change, resulting in similar problems as those described above. Last but not least, the local "agents" require frequency measurements every 200ms [47] in order to be able to respond.

1.5 Research Objectives and Contributions

Load modelling is important for both Power Systems Planning and for DR, yet have different requirements. For instance, when voltage dependence is important, ZIP model accuracy per load type is sought after [48]. In such models, TCLs are depicted as steady loads within a population and only the number of TCLs in operation alters, resulting to cold loads (refrigeration units) been modelled as a steady 24/7 load [48, 49, 50]. In, reality though, as seen by real world examples, that is not the case. Cold loads in particular vary during the day (20% to 60% [51, 52]. This would result in a small error within the overall load profile for Power Systems Planning, but for an aggregator participating in Balancing Services using such loads, response error has to be within 2.5% [53]. The general aim of this thesis is to examine and model this change in TCL demand and create aggregation models for DR which can easily update in time and with low sensoring requirements.

Ultimately, the key to determining the state of a TCL population is tracking specific parameters; state *on* and off duration (t_{on}, t_{off}, i.e. duty cycle D &

period T), power when *on* (P_{on}) and number of units *on*, off in regular intervals. Also knowledge of temperature settings (θ_{set}) and ambient approximately (θ_a). These can infer the thermal state of the population, as well as be used to create aggregated models. Knowledge of those parameters **do not** require use of sophisticated, highly accurate and expensive thermal sensors but only use of cheap Smart Plugs and/or Smart Meters.

Key points from the above are summarized below. The aim of this thesis is to address them to the largest extend possible and propose solutions. As such they can be seen as key objectives:

1. State of the art DR models assume "snapshots" of TCLs with some stochasticity instead of dynamic models.

2. First order ODE are inadequate to fully describe all TCLs. Yet higher order ODE formulations are complex and result to even more complex aggregation models.

3. Diverse modelling would result into smaller clusters and higher statistical errors. A "one for all" fit would be advantageous for aggregators.

4. CFPE compared to transition matrix based approaches assume homogeneity to be solved and therefore are prone to error under high heterogeneity.

5. Models developed so far require precise thermal sensor and/or frequency measurement equipment on load level and/or knowledge of power system's state on load level and/or precise load data in real time etc. All those would incur significant costs.

6. State estimation approaches on the other hand show great potential for reducing such costs and communications needs.

7. Unfortunately approaches without precise thermal sensors have not being studied extensively, more so combined with state estimation. Additionally,

mostly homogeneous or slightly heterogeneous cases (not realistic) were considered.

8. Lastly, most approaches rely only on centralised frameworks, which as pointed out by Trovato et al [22] and Ziras et al [47] tend to be slower in response compared to decentralised/distributed ones and not as robust. The use of privacy is also a concern.

Even though decentralised models exhibit faster response and cut down communication delay [22], central ones maintain knowledge of the system. As such, a distributed/hierarchical approach is more likely to be feasible, where:

- Knowledge of the system is central/distributed (as in substation or aggregator level) instead of local level. This includes monitoring the system's damping (MW/Hz), droop characteristics (MW/Hz), base power (MW) and system inertia (s) in real time or receiving this information from system operators.
- Aggregators update their state, receive system information and re-evaluate DR availability based on load information. This information can be based on smart meters (and limited additional metering infrastructure if needed [40]).
- A centralised (or hierarchical) signal, with system's knowledge in real time.
- Distributed/local unit decision in real time. Something proposed by Totu et al. as well [27, 35].

According to the points stated, this thesis tries to tackle them and provide potential solutions. It is worth pointing out that detailed conclusions are offered at the end of each chapter.

- Analysis of load types' potential for DR services (mainly focusing on TCLs).

This is done mainly in Chapter 2, but also in Chapter 3 with specific focus on cold loads.

- Examining the effects of external factors on TCLs and heterogeneous operation in time. Subsequently also on aggregation models and DR. This is done in Chapter 3, the case study being cold loads and addresses objective 1.

- Simplified equivalent model (first order ODE) of thermal loads described by higher order systems (usually second order). In Chapter 3 for cold loads with multiple compartments or cold loads operating next to each other (i.e. supermarkets). This addresses objective 2.

- Creation of a bottom up (Monte Carlo Markov Chain), heterogeneous, realistic model including external factors. The model has been validated against real world data (UK and EU cold load data) in Chapter 3 and addresses objective 1.

- Aggregation models which include external factors and heterogeneity during operation. One of which fit for high heterogeneity. These are detailed in Chapters 4, 5 and address objectives 3-5, 7. Key points are 1) that this method **does not focus on thermal sensors** but rather relies on duty cycle (time in power on and off states) to approximate relative state, 2) the transition matrix **does not have steady probabilities**, unlike in literature - introduced by Koch et al. [29] and later used by others (e.g. [54, 55]). Power on and off states are actually very easy to get from Smart Meters, some of which can give reading every second per appliance [56] or Smart Plugs (e.g. teckin, tp-link kasa, WeMo etc. [57]), which are also inexpensive or relatively low cost. These already come with apps, even on mobiles, which track real-time demand of each appliance (thus Power and duration of a TCL's on & off states) and can additionally command it to switch on/off [58]. Actually, they have inbuilt WiFi connection and are

becoming very popular due to reducing costs as well as their ability to be connected with devices such as Amazon Echo, Google Assistant and Apple's Siri for remote control.

- Introduction of a distributed state estimation framework (Gaussian Belief Propagation with Factor Graphs) from aggregate power measurements, with limited AMI. BP is highly flexible and can be used in hierarchical or central frameworks as well. This is introduced in chapter 4 and addresses objectives 6, 8.

- Methodologies for online updating and tracking of the aggregated population, without requirements of (precise) thermal sensor, but only power measurements from smart meters or disaggregation on feeder / substation level. This is described in chapter 5 and addresses objectives 5-8.

1.6 Thesis Structure

This thesis explores mainly DR in the residential sector and its loads (similar ones can be found also in the commercial sector). The focus is on large scale deployment of TCLs for DR, thus modelling aggregated behaviour of large clusters and state estimation. TCLs were chosen since they have the highest DR potential among loads [59]. The models developed in Chapters 3-5 are specifically for TCLs and not other loads. Similar aggregation models though, can be used for other loads and have been used in other studies (i.e. EVs [23]). Chapters are organized as follows:

- Chapter 2, analyses the residential sector and helps understand the role of residential loads in DR services. Loads' DR potential for both dispatchable and non-dispatchable DR is examined and based on their characteristics and human behaviour.

- Chapter 3 analyses the effect of external factors on TCLs demand. Equivalent first order ODE models are developed to express dynamics which normally require second order ODE. A dynamic model of individual units is created to reflect the dynamic behaviour of aggregated TCLs. The developed model is validated against real world UK and EU measured data using cold loads, via bottom-up Monte Carlo (MC) simulations.

- Chapter 4 Belief Propagation (BP) is introduced, a fully distributed algorithm, instead of KF for state estimation of TCLs. BP can also be used centrally or hierarchically and is compatible with KF. Such an approach can be used to avoid smart agents which require Advanced Metering Infrastructure (AMI) and knowledge of power systems' state. Distributed state estimation can be developed similarly for other loads if their state space model is known.

- Chapter 5 A novel aggregation methodology is proposed to calculate the Probability Transition Matrix (PTM) of highly heterogeneous TCL clusters, whilst being a simpler approach than the one proposed in literature. Additionally, it allows for different types of TCLs (or even loads) to be clustered together as long as they have similar PTM. Lastly, robust online updating methods are proposed, which do not depend on precise thermal sensor, but only power measurements (i.e. from smart meters or disaggregation), thus being able to track the changes during operation.

- Chapter 6 Finally, an overview of the main findings of the research and the contributions is presented. The implications and limitations of the research are discussed, while recommendations for further development and improvement, as well as next steps for this research are given.

1.7 List of Publications

List of the author's publications in relevance to this thesis:

Kleidaras, A., & Kiprakis, A. (2015, July). A roadmap for domestic load modelling for large-scale demand management within Smart grids. In International Conference on Wireless and Satellite Systems (pp. 33-47). Springer, Cham.

Kleidaras, A., Kiprakis, A. E., & Thompson, J. S. (2018). Human in the loop heterogeneous modelling of thermostatically controlled loads for demand side management studies. Energy, 145, 754-769.

Kleidaras, A., Cosovic, M., Vukobratovic, D., & Kiprakis, A. E. (2017, September). Demand response for thermostatically controlled loads using belief propagation. In Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2017 IEEE PES (pp. 1-6). IEEE.

Other publications:

Vagropoulos, S. I., Kleidaras, A. P., & Bakirtzis, A. G. (2014, September). Financial viability of investments on electric vehicle charging stations in workplaces with parking lots under flat rate retail tariff schemes. In Power Engineering Conference (UPEC), 2014 49th International Universities (pp. 1-6). IEEE.

Chapter 2

Residential Load Analysis for Demand Response

2.1 Introduction

This chapter's focus is to analyse the residential sector, the largest among all sectors, its loads and potential for DR services. Despite being the largest sector, it is the least utilized among the big three (industrial, commercial and residential). The principal downsides of the residential sector derive from its own particular structure, comprising of various small loads with high variety of sub-types, demand profile stochasticity, daily and seasonal consumption variation. Moreover, only part of the residential loads can be utilized for DR services, referred in literature as controllable or smart loads. These loads have different characteristics, constraints and thus suitability for DR services. Tracking and coordinating, in real time, large numbers of different loads is not feasible (if not improbable in some cases [40]).

The first part presents current DR services, separated in Non-Dispatchable and Dispatchable. The control strategies for DR can be divided into indirect (Non-Dispatchable) and direct control (Dispatchable). In Non-Dispatchable DR, consumers are incited to modify their behaviour and thus demand profiles [60, 61]. Dispatchable DR is usually central or based on an automated control [62], dispatched under emergency cases such as faults, lost generation, high RES fluctuation etc., in order to balance demand and supply. In some DR services, prompt reaction within a couple seconds is critical and the participating loads must have the capability to respond almost instantaneously [4].

Afterwards, an analysis of the residential sector is presented, focusing on the main issues of large-scale DR, which stem from residential sectors' large number of loads, with low demand and high heterogeneity. Basic load types are analysed and their potential for DR is examined. Among them, thermostatically controlled loads (TCLs) are the most potent.

The chapter concludes with highlighting the importance of loads with some form of energy storage (be it chemical, thermal or other form), especially for Dispatchable DR, since balancing services are becoming increasingly crucial [3]. Small commercial consumers have similar loads with residential ones and **less** stochasticity, thus can also be used in the same frameworks developed in later chapters of this thesis.

2.2 Control Strategies for Demand Response

2.2.1 Indirect Load Control

Generation cost varies depending on numerous factors, yet the most significant ones are the generation mix and the system's constraints at each moment. Generation mix is driven by demand, which changes throughout the day. In principle, unit commitment starts with the cheapest generation unit and the next cheapest available one (given constraints) is dispatched, until the demand is satisfied. In reality various factors affect the decision making in energy markets but in general cheaper units are procured first and the most expensive ones cover peak loads. Thus, market spot price during demand peaks tends to be much higher than base or minimum demand, with a difference of up to 2 orders of magnitude. High penetration of RES has significant consequences in energy markets' spot price, depending on market structure.

Indirect load control refers to dynamic pricing or other forms of incentive-based control. Essentially it is used in Non-Dispatchable DR, where energy users might choose to activate Demand Response on their own, without being dispatched by a third party, given some incentive. Non-Dispatchable DR focuses on energy markets, mainly peak price reduction through peak load reduction [62], price following services [63](a UK real world pilot project) and RES integration [28, 31, 64]. They are referred in literature as peak shifting/shaving, valley filling and RES following methods. The aim of time sensitive pricing is to support cheaper energy transmission and distribution, and promote generation from greener resources while operating power systems within limits (avoid overloads during peaks).

Dominant tariffs for indirect control are Time of use (TOU), Real-time pricing

(RTP), Critical peak pricing (CPP) and Peak time rebate (PTR) [61, 65, 66, 67]. A percentage of consumers is expected to alter the starting time of their appliances. The exact response (fraction of consumers) to dynamic tariffs and their demand depends on demand-price elasticity, the appliances (loads), time of the day and human behaviour. Consumers participating in such schemes benefit from lower market prices, normally during off-peaks and/or excess RES (i.e. cheap generation mix in general). Their primary target is to shift the demand to different periods (earlier or later) based on generation and demand profiles. [68]

TOU tariff: Pricing in fixed periods of time to reflect the higher cost of generation during peak periods, usually occurring around the same time every day seasonally, and lower cost during off-peak periods (e.g. UK's Economy 7). However, TOU cannot capture the real time cost of generation, such as intermittent RES, which varies based on weather conditions.

RTP tariff: RTP is based on the dynamic nature of the generation cost. Consumers pay electricity prices that are linked to the wholesale cost of electricity. This type of pricing can fluctuate on a sub-hourly basis, making it harder to track and maybe not as easily implemented in that form for residential and small commercial consumers.

CPP tariff: CPP reflects the true cost of power generation to electricity consumer during peak hours. A price signal is sent based on the cost of production. The rest hours of the day (off-peak), a discounted tariff exists for the consumers participating in the CPP. Customers can reduce their bills by shifting their consumption from more expensive hours (peak hours) to less expensive hours, thus reducing the overall demand peak.

PTR tariff: PTR is an alternative to CPP and works the opposite way. Instead of giving a discount during off-peak hours, consumers receive a rebate for reducing

their demand during peak hours. A set point of power can be used (in kW or kWh). The rebate is computed as the difference of the set point and the consumer's power demand.

From the above tariffs, RTP is obviously the most suitable for RES integration and to reflect generation cost, but also the hardest to keep track in small intervals and most complex for residential and small commercial consumers. An alternative could be an approximation of the average RTP over the next few hours (2 to 6); i.e. for a given y period, based on the RES generation forecast, a price $x \in [x_{min}, x_{max}]$ will occur, which is provided on a day (or few hours) ahead basis. For most consumers, average RTP or similar structures make more sense and are easier to follow [68, 69]

2.2.2 Direct Load Control

Dispatchable DR can provide various balancing services, mainly for grid reliability [65, 70]. Reserve services provided by DR can be utilized either for positive or negative regulation, [71], to respond to unexpected RES generation and/or avert RES curtailment or price following in general [72, 73, 74, 75, 76]. Yet, most of the Dispatchable DR services focus on frequency control [36, 77, 78, 79], and short term reserves, since frequency regulation is imperative for real time stability and reserves for short term stability of power systems.

National Grid's current DR services include reserve, frequency control and RES following, all of which belong to Dispatchable DR services (Table 2.1) [4, 70, 71, 80].

Firm Frequency Response (FFR) is the firm provision of dynamic or non-dynamic response to changes in frequency. FFR was created to give service

providers and National Grid a degree of stability against price uncertainty under the mandatory service arrangements.

Enhanced Frequency Response (EFR) is a dynamic service aimed at improving the management of system frequency pre-fault to maintain system stability. Active power changes proportionally in response to changes in system frequency.

Fast Reserve (FR) provides the rapid and reliable delivery of active power through an increased output from generation or a reduction in consumption from demand sources, following receipt of an electronic dispatch instruction from National Grid.

Short Term Operating Reserve (STOR) is a service that provides additional active power from generation or demand reduction, when at certain times access to sources is needed. Usually due to unforeseen generation unavailability or when actual demand on the system is greater than forecasted.

Demand Turn Up encourages large energy consumers (or generators) to increase demand (or reduce generation) at times of excess renewable generation and low demand. This typically occurs overnight and during weekend afternoons in the summer.

The main difference between the frequency services is time scale at which providers need to respond. These services can be provided by generators, spinning reserves or non-spinning reserves (e.g. batteries) and aggregators. Dispatchable DR services are an attractive alternative for aggregators due to the fact that there is no standby cost, the response time is short and no extra emissions are caused. Though barriers arise for residential and small commercial end users, due to their small individual demand, load profile deviation (residential users alter their consumption more than commercial or industrial users) and constrains of loads

based on their type, mainly delivery power and availability periods. Thus, aggregation of large numbers is essential, combined with other distributed sources in some cases.

On a local level, given the fact that loads are ubiquitous, there is potential for unique DR services, which central generation cannot provide. Those include voltage control, overload relief (transmission and distribution), congestion management; usually during peak times or in cases of faults [60, 74, 78].

Table 2.1: National Grid's balancing and reserve services where DR participates

Service	Response	Duration	Power
FFR	<10/30 sec	≥20 sec/30 min	≥1 MW
EFR	<1 sec	≥15 min	≥1 MW, ≤ 50 MW
FR	<2 min	≥15 min	≥50 MW (>25 MW/min)
STOR	<20* min	≥120 min	≥3 MW
Demand Turn Up	-***	260** min	≥1 MW

This table contains the main balancing services in which DR providers can participate in UK and their main technical requirements [4, 70, 71, 80].

** longer times are also acceptable (240) but not preferable, ** variable, 260 min is the average for 2016 *** variable, providers are given notice as contracted*

The services in Table 2.1 are not to be confused with the Balancing Mechanism (BM) market in UK. The Balancing Mechanism market is a special market which is used after the 4 main markets (two Day-Ahead and two Intra-Day) in order to match energy and supply in normal operation. Larger Generators participate in that 30min period market, with bidding offers to increase or decrease their output (4 bands each direction). These services are utilized in contingency cases (e.g. when BM is not enough to match supply and demand due to a fault or a power surge) and can be provided by either BM or non-BM units, where DR can participate.

2.3 Residential Sector for DR

There are three main consumer sectors in power systems: the industrial, commercial and residential sectors, with combined consumption of 91.56% of total consumption in the case of UK (Figure 2.1). Currently, the focus of DR is on the industrial sector due to the inherently large loads and existing AMI (sensors and metering technologies). The relatively low AMI cost to power rating per unit means aggregating units to minimum power ratings for DR (Table 2.1) is a simple and cost effective approach. The commercial sector is next in line, mostly from bigger energy consumers and/or groups of them. Residential loads are gaining more attention lately, but have not been largely used since the loads are small, distributed, and not automated [60, 61, 62, 73, 75, 78].

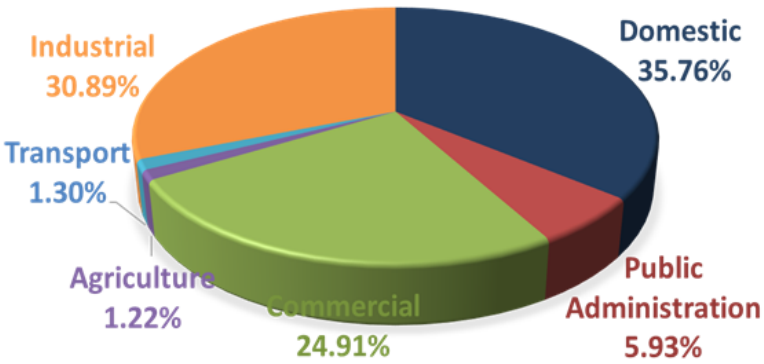


Figure 2.1: Breakdown of UK's electrical energy consumption per sector [52]

The major challenge in utilizing DR's full potential lies in the domestic sector, having the highest consumption of the three (35.76% [81]). Mainly because new sophisticated aggregation and estimation methods are required, in order to control simultaneously a large number of small units to achieve the same results as large commercial or industrial units (aggregation in the order of thousands of units to participate in Dispatchable DR). In addition, issues arise from deviation in load profiles, limited knowledge of their load composition, time availability and limited

knowledge of their potential for DR, including the end users' awareness and thus willingness to participate.

The introduction of smart meters, as well as smart loads, in residential and small commercial consumers makes such schemes possible [15]. A lot research has been also focused on disaggregation techniques via smart meters, giving load composition insight unseen before [41, 42]. Signals can be sent to smart meters for non-intrusive control of specific loads and prompt others to delay their start time, or next operation [19]. Consumers participating in DR schemes can benefit from discounts, lower pricing or rebates. They should also have the option to opt-out if needed for a certain period of time (a simple switch on/off option).

2.3.1 Residential Sector Analysis

Knowledge of the availability of the controllable loads is essential, which means knowledge of the residential sector's demand. Table 2.2 shows ownership statistics of selected residential appliances for the case of UK and EU [49, 68]. These were selected as potential loads for DR and have been studied in literature [51, 82]. The reason being that they have the higher power rate among residential ones and potential to shift their demand or alter their duty cycle as investigated in [51].

Domestic loads are not consistent, as habitual patterns of users and weather conditions drive demand profiles, which is the main problem faced when trying to cluster them based on demand profiles. For instance, activities like use of washing machines (and sometimes tumble dryers), are not done on a daily basis. A household, thus, has different daily profiles, causing high daily demand profile deviation, unlike commercial and industrial consumers. Clustering domestic demand profiles on a similar manner to commercial/industrial ones, would give

inaccurate results. This is apparent in Figure 2.2, where the difference of base load (orange) with non-daily load (blue) can be seen. Some activities occur on a weekly basis (e.g. laundry) and not a daily one and/or depend on weather conditions from day to day (e.g. heating). For instance, washing machines are reported to have on average 5 cycles/week and dishwashers 4.5 cycles/per week [49].

Table 2.2: Ownership statistics for selected appliances

Appliance	UK Ownership	EU Ownership
Fridge-freezer	69.4%	106%
Refrigerator	37.7%	
Chest freezer	15.5%	52%
Upright freezer	31.4%	
Washing machines	97%	95%
Tumble dryers	56%*	34.4%
Dishwashers	42%	42%
Heating Circulation pumps	88.8%**	70%
Electric space heating (storage)	6.13%	-
Electric space heating (direct)	/0.74%	-
Electric water heating	4.8%	-
Electric oven	65.5%	77%
Electric hob	44.8%	77%
Microwave	93%	-
Kettle	98%	-

Ownership statistics according to [81] & [51] **Includes washer dryers as given by UK statistics, ** based on number of dwellings with central heating/boilers [81]*

Even though the demand profiles of individuals cannot be predicted and vary daily, on a larger scale, aggregated demand profiles of "similar" users, do have

consistency. Due to habitual patterns, the probability of using specific appliances is predictable on a large "relatively homogeneous" cluster, based on historical data and conditions (such as working days, holidays, weather etc.) [82, 83].

Thus, a large number of end users (thousands) can be grouped in a few clusters based on their similarities, simplifying their management, supervision and forecasting. Moreover, this may allow unmonitored areas to be matched based on their characteristics to the closest template with a relatively low error (on an aggregated level) [84].

Load clustering can simplify data processing for various applications. The main problem with this, when it comes down to DR, is the fact that load profiles do not give information about the availability of controllable loads (volume, time, etc.) but only the overall shape of the profile. Aggregated load data provides even less information per load type and is harder to disaggregate. These is apparent by looking at Figures 2.2 - 2.6. Typical weekdays are shown in these figures since they constitute the majority but weekends can be simulated as well.

These Figures (2.2 - 2.6), as well as other household profiles in this chapter, are synthetic load profiles created via a bottom-up modelling approach via Markov chain Monte Carlo method. The same methodology is used as in [48] and in fact updated appliance data according to DECC has been used [52]. For more details the readers are prompted to read [48].

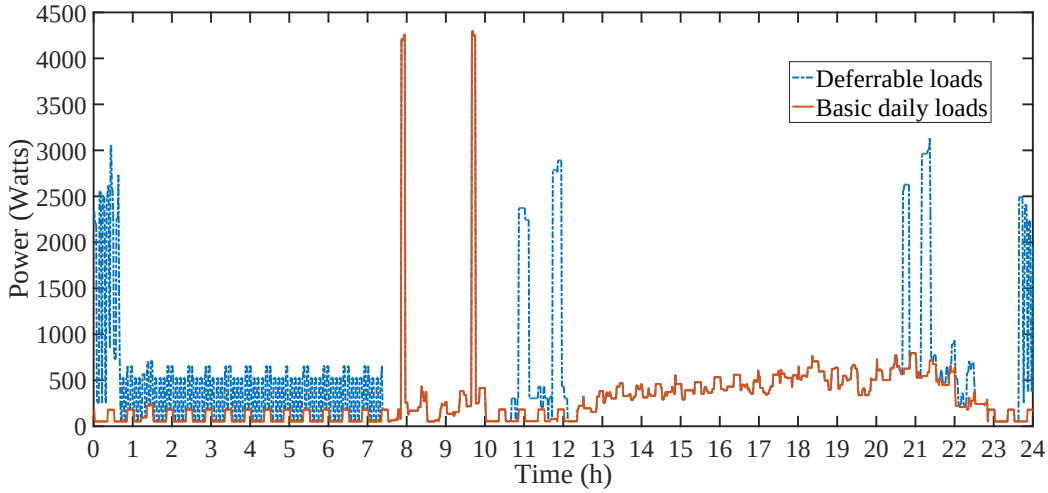


Figure 2.2: Household simulated daily basic demand and deferrable demand (dotted). Deferrable demand for selected household includes dishwasher, washing machine, tumble dryer, electric storage water & space heating. Simulation was for a typical winter weekday in UK.

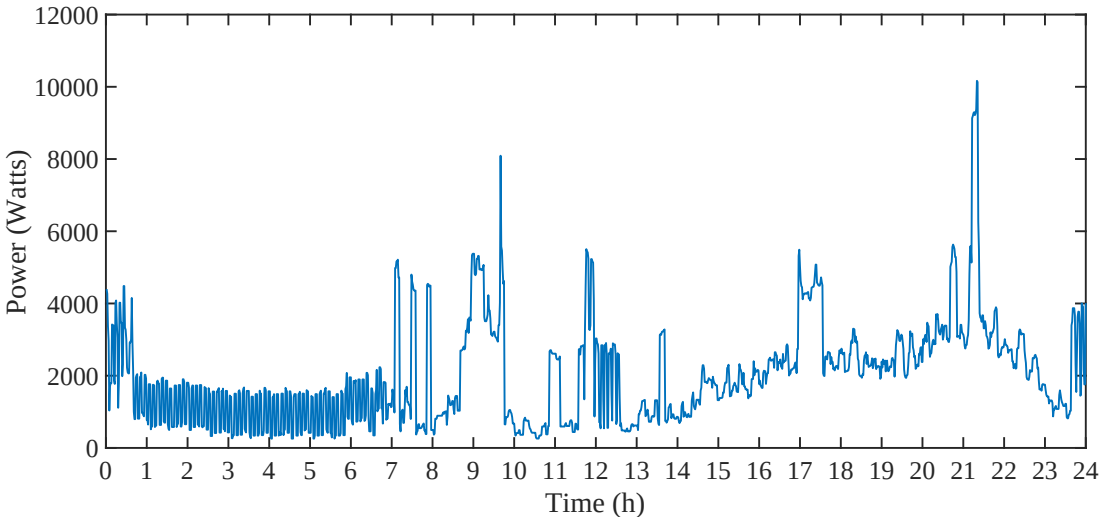


Figure 2.3: Aggregation of 5 simulated households, typical winter weekday in UK.

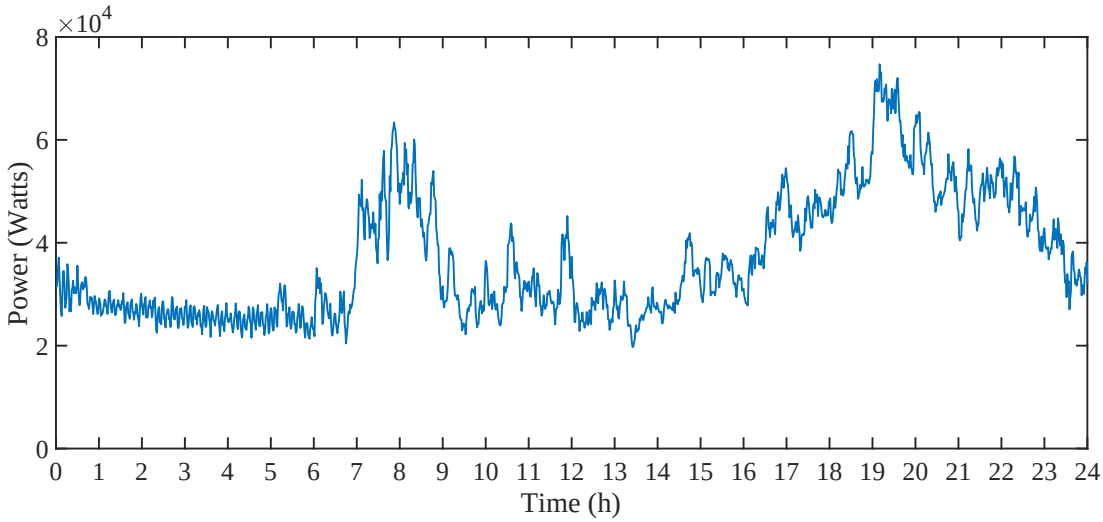


Figure 2.4: Aggregation of 100 simulated households, typical winter weekday in UK.

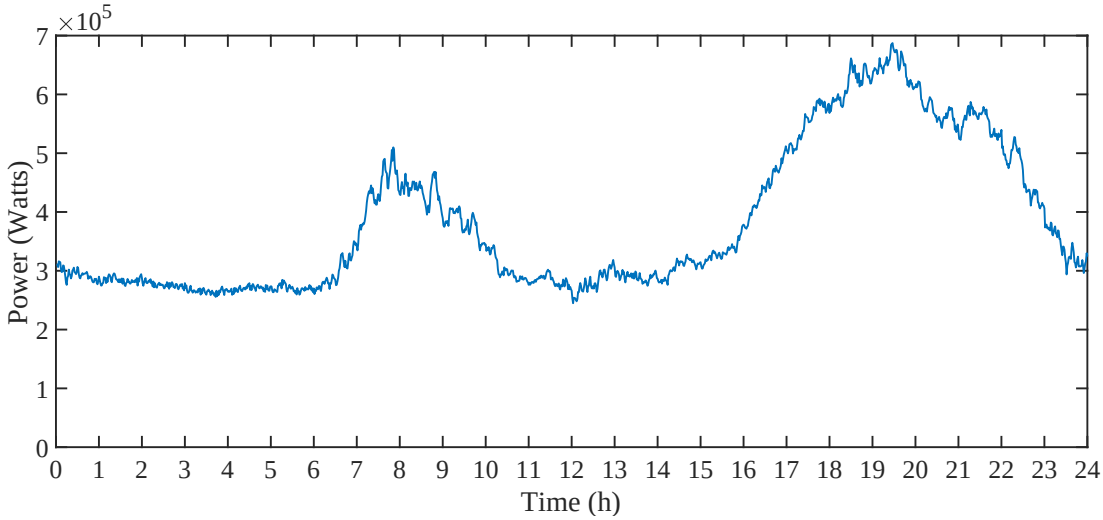


Figure 2.5: Aggregation of 1000 simulated households, typical winter weekday in UK.

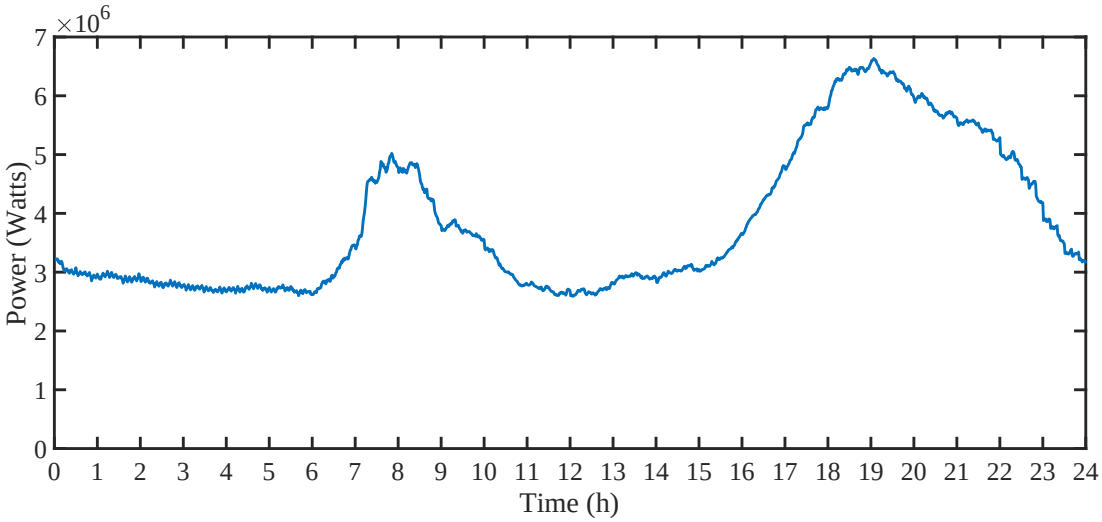


Figure 2.6: Aggregation of 10000 simulated households, typical winter weekday in UK.

2.3.2 Classification Based Aggregation Analysis

Classification takes into consideration occupation characteristics; total number and employment status, demand characteristics; overall demand and time of use.

Table 2.3: Occupancy mixture

Number of total occupants	Number of employed occupants				
	0	1	2	3	4
1	1210	2316	-	-	-
2	289	790	2290	-	-
3	105	395	1000	210	-
4	0	290	895	105	105

Number of households per type, based on UK population statistics which cover 98% of the population [85]. Total number of households: 10000

Table 2.3 is representation of the UK households (covers 95% of the population).

The importance of occupation in determining load demand in volume and time is investigated. The correlation between occupation and demand can be seen in Tables 2.4 and 2.5. Figures 2.7 & 2.9 are additionally presented to visualize some of these results.

Table 2.4: Daily mean demand (Power - Watts)

Number of total occupants	Number of employed occupants				
	0	1	2	3	4
1	350.673	273.045	-	-	
2	463.943	410.429	373.442	-	-
3	524.967	460.975	444.101	420.233	-
4	-	524.536	490.053	474.480	484.068

Household average consumption per type (as per Table 2.3).

Table 2.5: Relative standard deviation

Number of total occupants	Number of employed occupants				
	0	1	2	3	4
1	32.9%	36.7%	-	-	
2	30.4%	32.1%	36.2%	-	-
3	53.3%	31.6%	34.9%	29.9%	-
4	-	27.1%	28.7%	28.2%	28.4%

Relative standard deviation from average consumption per type.

Households of the same size, consisting of non-working occupants tend to have higher consumption, since the time spend in the house increases. Though that does not necessarily mean that the opposite in some cases does not occur. Cases such as A) work based from home or B) students (classified as non-working) being absent during working hours are just a few to name.

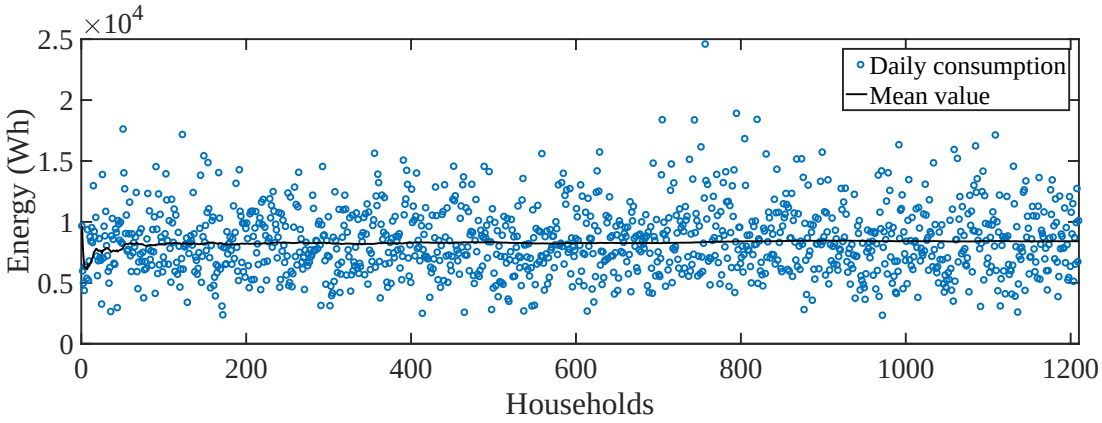


Figure 2.7: Daily energy consumption of households consisting of 1 unemployed occupant. Number of households: 1210 out of 10,000 as per Table 2.3

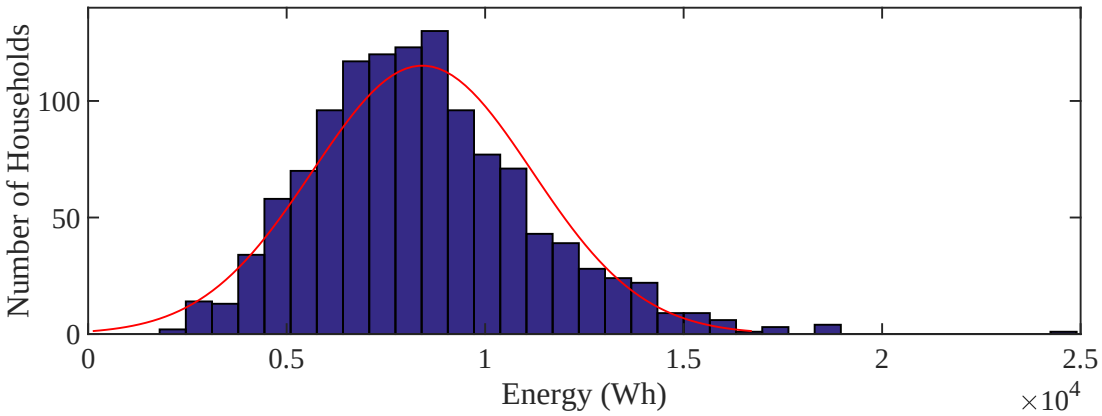


Figure 2.8: PDF of Figure 2.7. Mean energy consumption $8416.152Wh$. Relative standard deviation 32.9%

Thus a better approach is a combination of household characteristics and overall historical data. Probability Mass Function (PMF) based approaches have been suggested to tackle this issue [82]. For example, this allows case A to be placed in a "non-working occupants" dominant group and case B in a "working occupants" dominant group, assuming similar consumption characteristics.

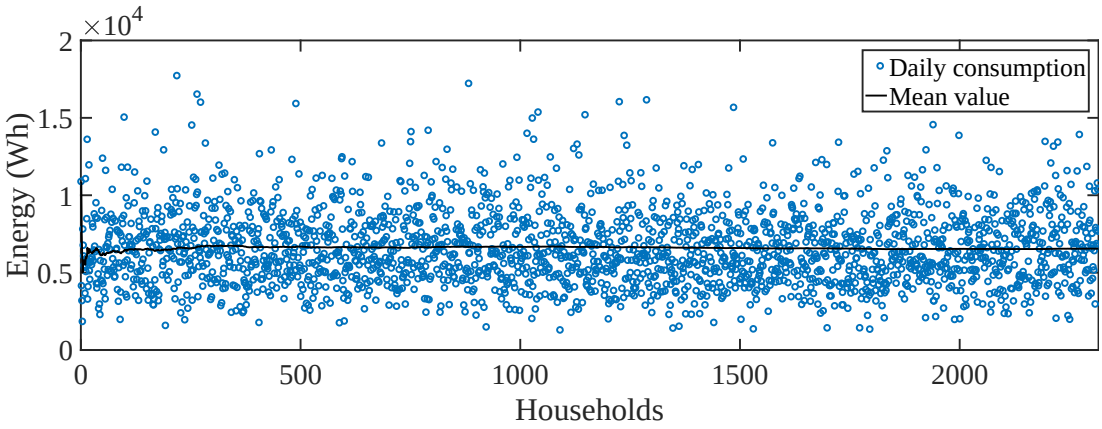


Figure 2.9: Daily energy consumption of households consisting of 1 unemployed occupant. Number of households: 2316 out of 10,000 as per Table 2.3

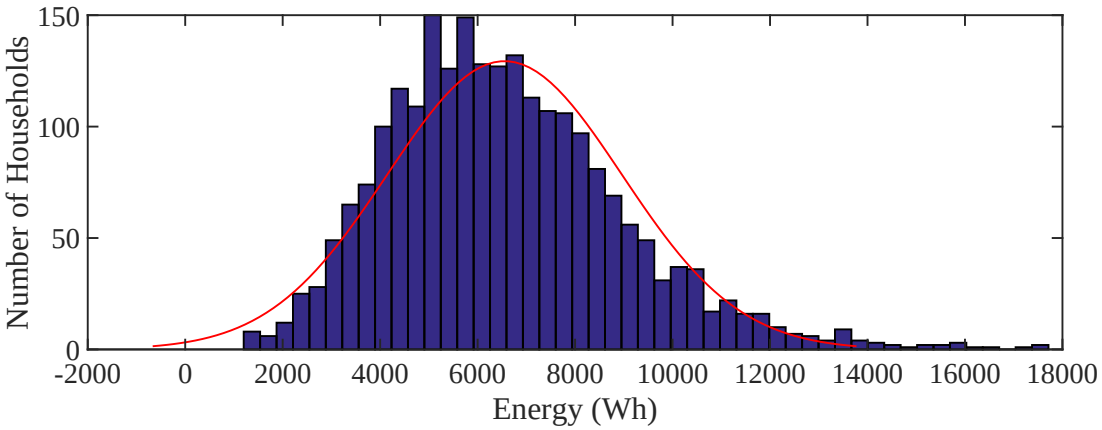


Figure 2.10: PDF of Figure 2.9. *Mean energy consumption* $6553.08Wh$. *Relative standard deviation 36.7%*

Figures 2.7 - 2.10 show the distribution and convergence of daily demand consumption. The mean converges within a sample of a few hundreds for households of similar type, which is positive for aggregation per household type. Another metric to look at is the relative standard deviation. In Table 2.3, it can be seen that some households have high RSD values, largest been the household of 3 non-working occupants, which also has a small number of samples. In which case,

RSD values are examined with caution to avoid wrong conclusions. Nonetheless, a decrease in RSD values is observed as the household number increases and as the number of working occupants decreases. The first one can be attributed to more consistent use of appliances, e.g. more frequent use of washing machine within a week for a bigger household thus less demand deviation. The second one can be attributed to occupants sharing more activities (habitual patterns) due to higher time flexibility as opposed to working occupants, especially in cases where their working hours do not align.

Comparisons between individual and aggregated demand from households with 1 employed and 1 unemployed occupant can be seen in Figures 2.11 and 2.15. The "base" consumption observed is mainly due to cold loads. Differences in peaks and minimum demand can be seen both in value and in time. Overnight, demand is almost the same, as expected.

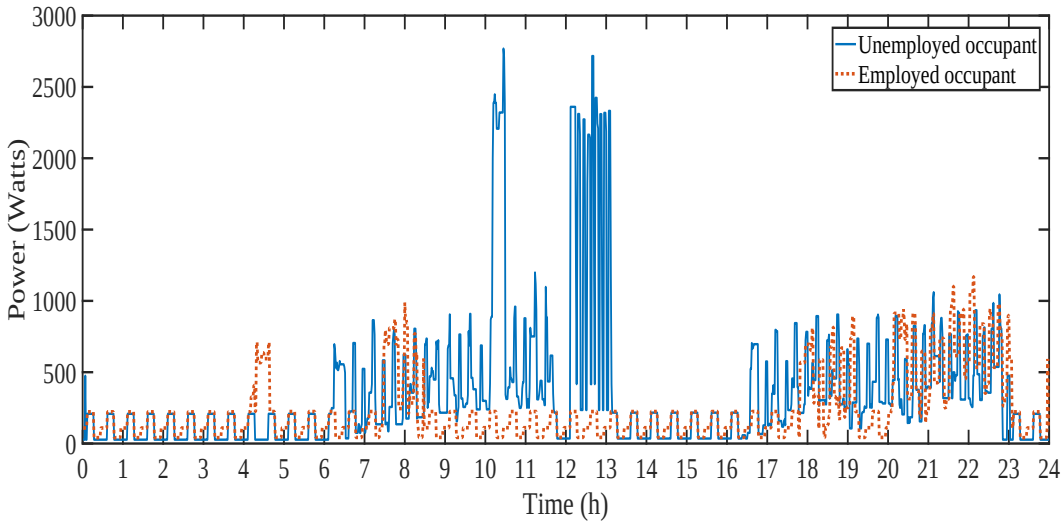


Figure 2.11: Typical simulated household demand profiles, one from a single unemployed occupant and one from an employed one.

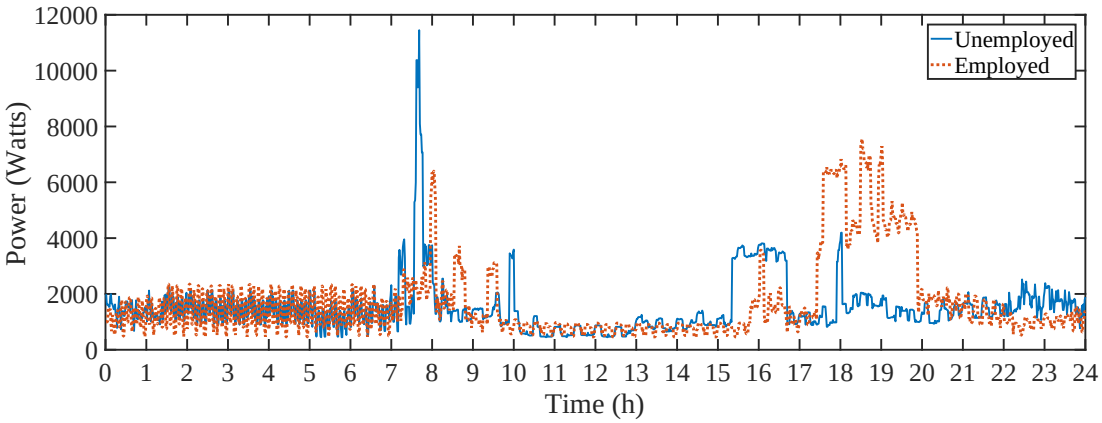


Figure 2.12: Aggregation of 5 simulated households with one employed occupant and 5 simulated households with one unemployed occupant, typical winter weekday in UK.

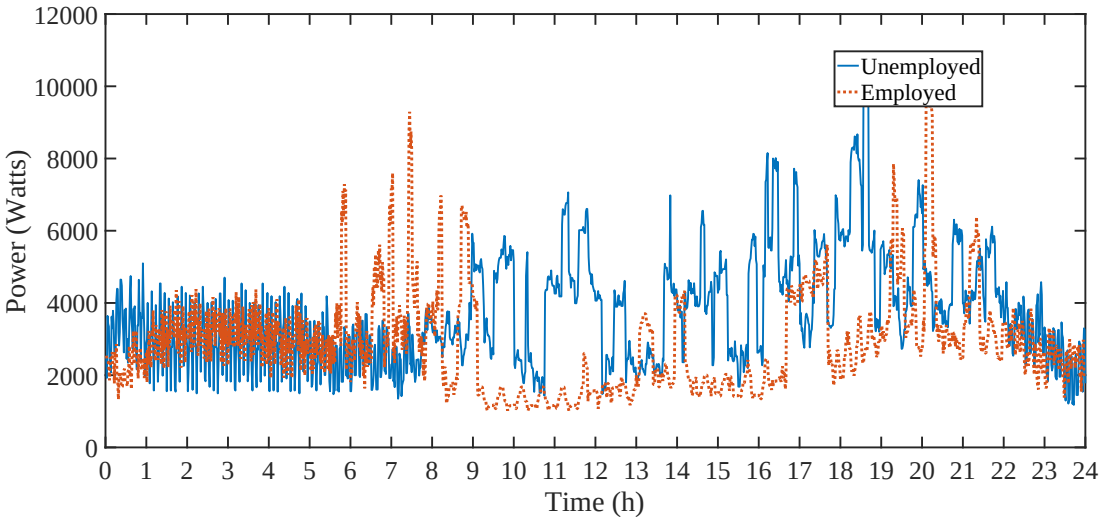


Figure 2.13: Aggregation of 10 simulated households with one employed occupant and 10 simulated households with one unemployed occupant, typical winter weekday in UK.

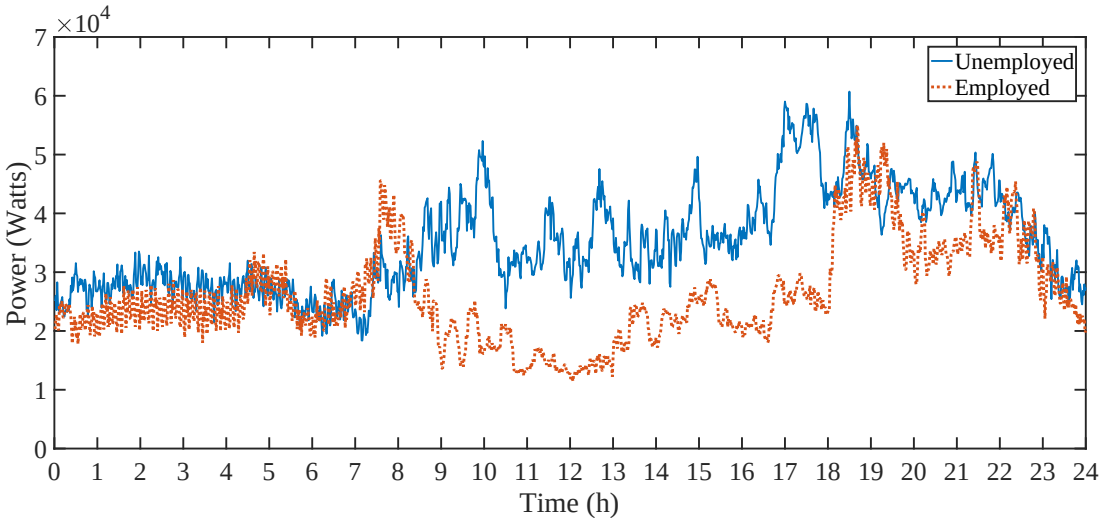


Figure 2.14: Aggregation of 100 simulated households with one employed occupant and 100 simulated households with one unemployed occupant, typical winter weekday in UK.

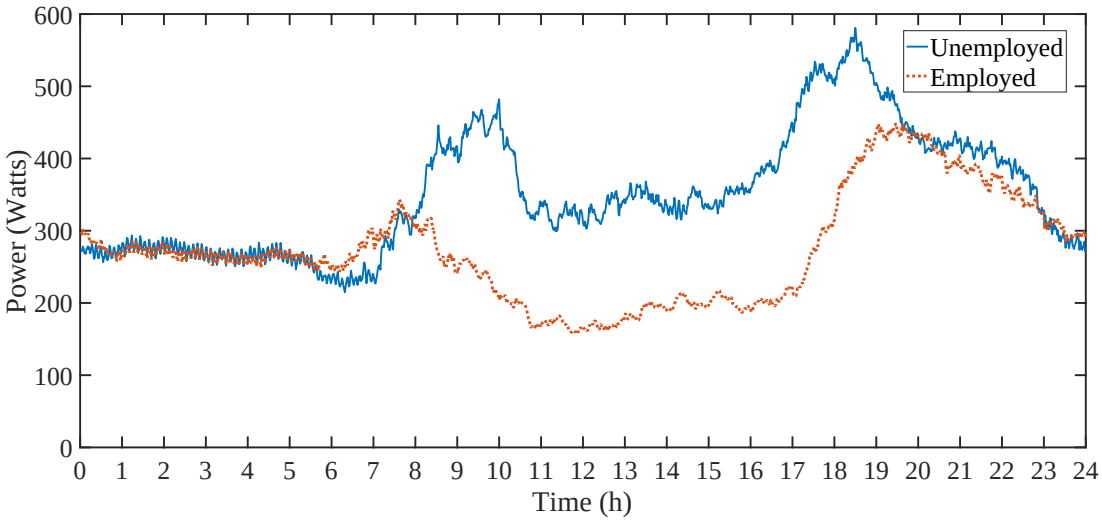


Figure 2.15: Averaged aggregation of all (1210) simulated households with one employed occupant and all (2316) simulated households per Table 2.3, typical winter weekday in UK.

A note at this point is that part of those peaks and consumption during hours of high demand is attributed to wet loads (washing machines, tumble dryers, washer-dryers, dishwasher). Wet loads, have a wide window to shift their operation [51]. As such, they have one of the highest potential for Non-Dispatchable DR (dynamic tariffs).

Finally, clusters of households consisting of only employed occupants and households with unemployed occupants can be seen in Figures 2.16 & 2.17). This shows the similarities between profiles, associated with consumption in mid-day and is accompanied by an expected increase in average demand as household size increases.

Another important aspect of household types between employed and unemployed occupants are the activities throughout the day, such as use of thermal loads. During the day, the probability of household activity differs significantly between them (as seen in Figures 2.11 - 2.14) which is taken into account in Chapter 3 where probability of interaction with refrigerators, freezers and fridges is considered, as well as room temperature (house heating in combination with heating profiles, Figure 3.2). The analysis of average demand does not give an insight in household activity but only total demand, which has some correlation with appliance ownership and use (e.g. a larger household is more likely to have a dishwasher and use the washing machine more often). Larger household are more likely to have loads available for DR.

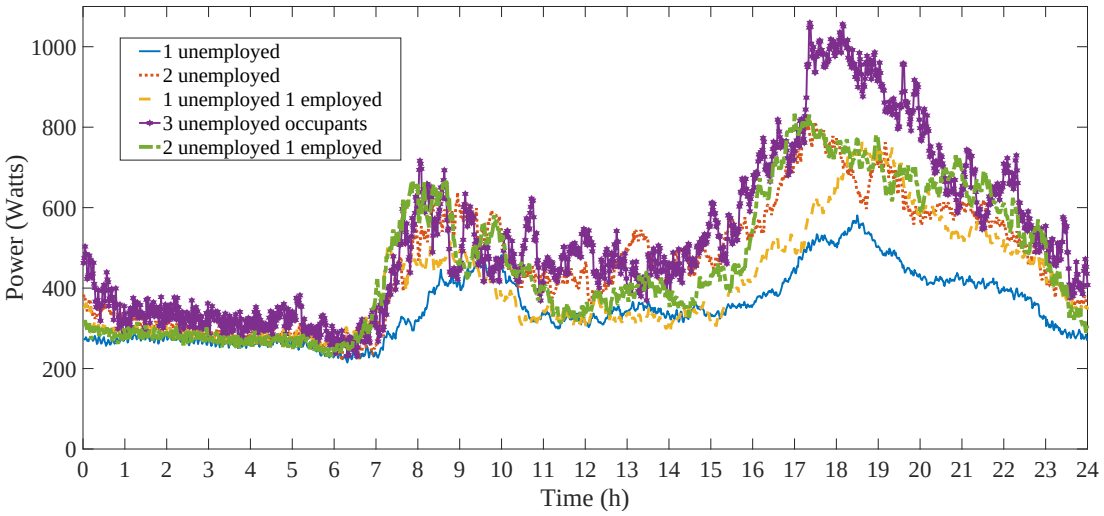


Figure 2.16: Average demand profiles, household clusters with unemployed occupants

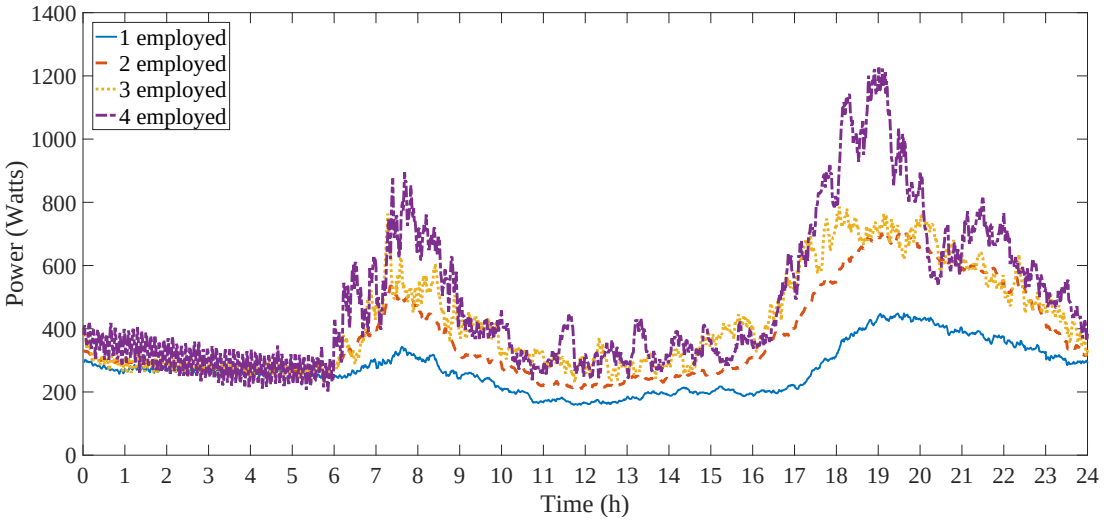


Figure 2.17: Average demand profiles, household clusters consisting of only employed occupants

2.4 Loads Analysis for DR

Knowledge of the composition of residential demand is imperative. This effectively means analysing the loads and their potential for DR, their total volume (aggregated power), and availability during the day, week, season (thus essentially main consumption driving factors). As seen from Table 2.2, cold loads are the most common, followed by wet loads. The main residential loads include the following:

Thermostatically controlled loads (TCLs). Which include cold loads (refrigerators, fridge-freezers, upright-freezers, and chest-freezers), electric space heating (direct and indirect), electric water heating (direct and indirect) and air conditions.

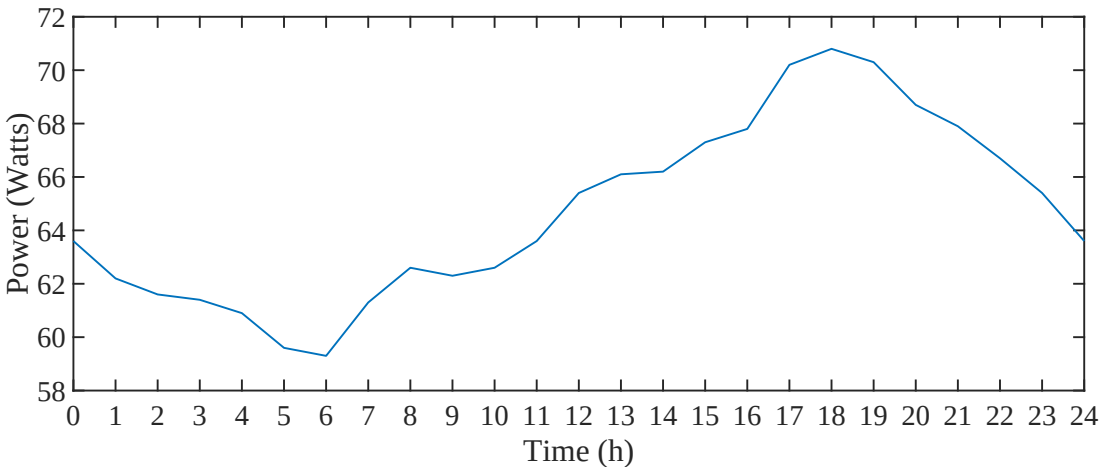


Figure 2.18: Average daily residential cold load consumption in UK, as given by DECC in half hourly intervals [52]. A significant change, around 20% between minimum and maximum demand is observed.

Cold loads have a ratio on/off around 1:2 to 1:3 (newer technologies), with a total cycle from 50 to 90 minutes (newer technologies) [25]. Their demand and cycle are mainly affected by human behaviour, directly and indirectly. Additionally,

cold loads have a defrost operation, accounting for around 5% of operation times (3-4 times per day on average, yet a relatively high demand for that short period [31, 86]. Defrost operation has potential for load shifting. The average UK cold load demand can be seen in Figure 2.18 [25]. Their 24/7 availability gives high potential for DR, with defrost cycling being suitable for RES following.

The load analysis here is used later in Chapter 3 to compare the average UK residential cold load with simulated models, such as in Figures 3.23 and 3.25 of Chapter 3, where they are used for validation of simulated aggregate cold load demand.

Storage space/water heating operate mainly overnight, usually operated during cheap overnight tariffs (such as Economy 7), store thermal energy, which is released the next day. Some reheating during the day might be required, for short periods, to maintain temperature. Suitable for balancing services [32] overnight (mainly) and also absorbing excess cheap RES power. Affected by weather mainly and human behaviour. Average UK electric storage heating and water heating load demand can be seen in Figures 2.19 and 2.20.

In the UK, Kemna et al. [87], state that 48% of houses use electricity to heat water for daily domestic use, of which the vast majority (90%) is used in electric shower systems.

HVAC (Heating, ventilation and air conditioning). HVAC is common in warmer countries, where indoors air conditioning during summer is more essential. In the UK, HVAC systems can be met mostly in commercial buildings. In places like California they are quite common, have high demand and are considered among the top loads for both Dispatchable [29, 44] and Non-Dispatchable DR [28, 31].

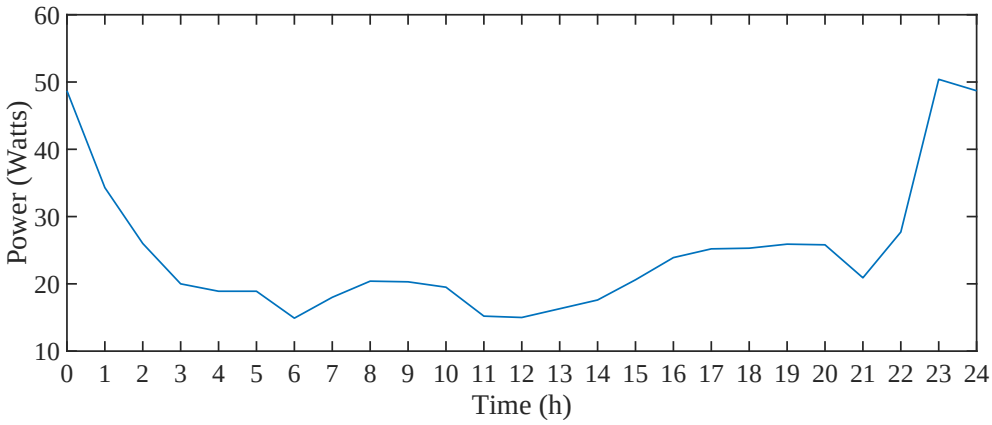


Figure 2.19: Average daily residential space heating load consumption in UK, as given by DECC in half hourly intervals [52].

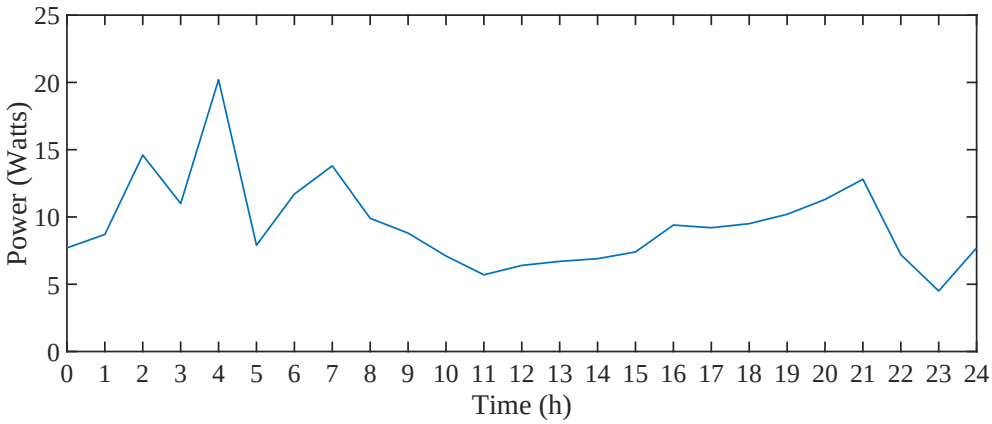


Figure 2.20: Average daily residential water heating load consumption in UK, as given by DECC in half hourly intervals [52].

Circulation pumps (gas or electric). Every central heating system, whether it is electric or gas, uses circulation pumps to circulate the warm water through the radiators. They can be used for both DR types [88]. Since they are used in conjunction with heaters, weather is the basic affecting factor.

Other TCLs: Direct SH (except for portable) are similar to cold loads, though

not so commonly used in UK, since it is not a feasible option compared to alternatives (electric storage or gas central heating). Direct water heating (showers mostly), has essentially no usable thermal storage capacity and thus cannot be used for DR.

Most TCLs have inherent thermal inertia (except for direct water heating like showers). Their operational temperature is not constant but fluctuates within a set range, with the help of a thermostat, who controls the on/off states of the appliance. This thermal inertia allows TCLs to alter their consumption pattern for short periods without affecting the quality of service [38, 40, 62]. With the use of an external signal, TCLs can switch state and switch back to the original upon reaching the normal temperature limit (Figure 2.21). This gives TCLs thermal storage capabilities [13], making them excellent candidates for short term Dispatchable DR, one of the most important forms of DR, if not the most important [3].

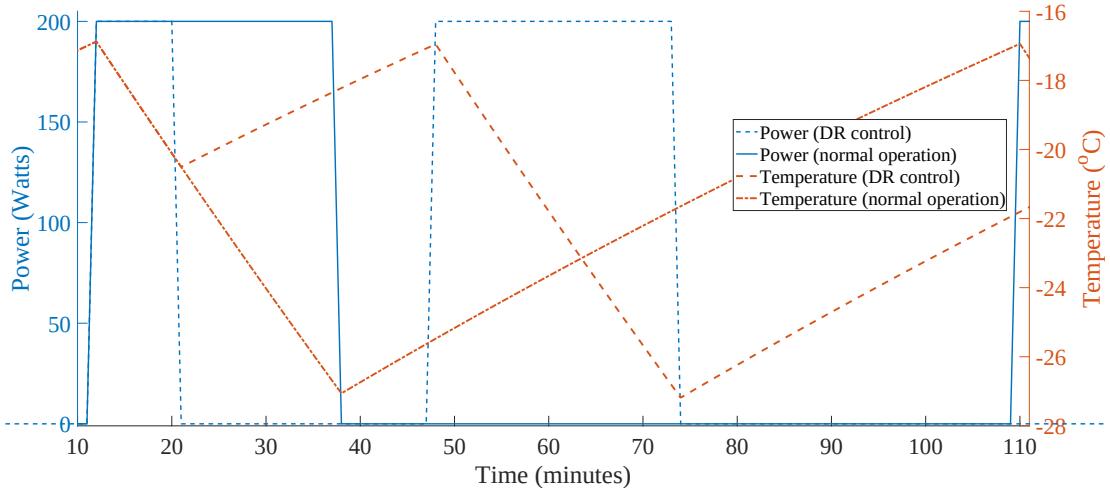


Figure 2.21: Cold load (freezer) with and without use of DR. The normal cycle, shown with solid line, starts at 10' where temperature is -17 ^oC and ends at 38' once the temperature has reached -27 ^oC. Use of DR actions can modify the cycle, e.g. in this example with a signal to switch off at 21', but without exceeding temperature dead-band limits.

Wet loads: Washing machines, tumble dryers, washer-dryers and dishwashers. The loads with potentially the highest level of freedom to shift their start time. Most new models include start delay options and they are already widely utilized by the consumer base, while also having the highest acceptable potential to be deferred from users [89]. Economy 7 users with relatively silent wet appliances

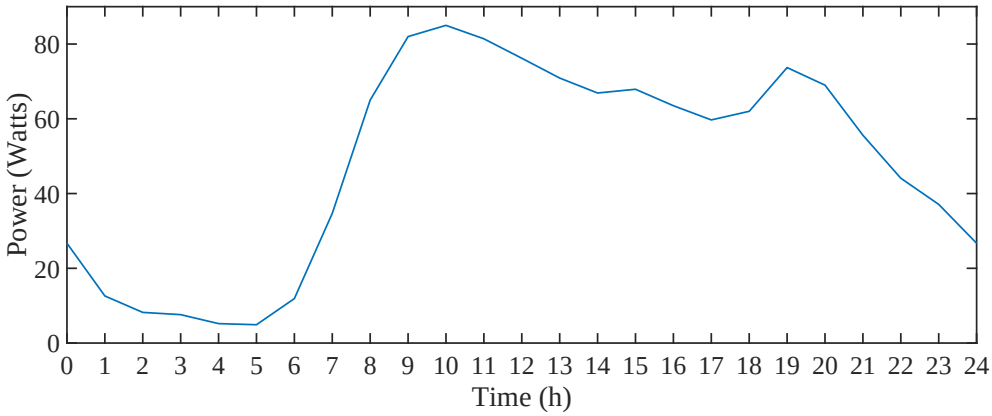


Figure 2.22: Average daily residential wet load consumption in UK, as given by DECC in half hourly intervals [52].

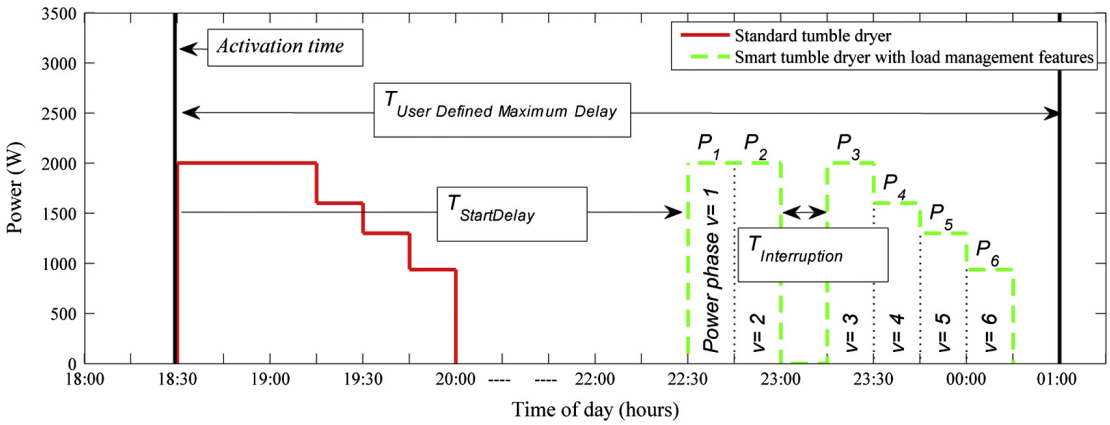


Figure 2.23: Example of a smart tumble dryer being used for DR [19].

make use of the cheaper price overnight. Suitable for dynamic pricing, but might also have potential for reserve services through new smart appliances. As investigated by Nistor et al. [19], during operation, an external command can be given to delay the next part of the operation, as seen in Figure 2.23.

Cooking: Including electric ovens, electric hobs, microwaves and kettles. Cooking loads cannot realistically be interrupted without affecting "quality of service", even though most new ovens have start delay options, this is essentially for the convenience of having the food ready at a certain time. Hence, it cannot realistically be utilized for dynamic pricing (time shift is very limited [51] based on human behaviour).

Lighting: Various types of lighting technologies exist, most common ones include gas incandescent lamps, halogen incandescent lamps, compact fluorescent lamps, high intensity discharge, light emitting diodes and more. They cannot realistically be used for DR.

Consumer electronics and home computing: This term includes, but not limited to TVs, set top boxes, power supply units, games consoles and desktops, laptops, mobile phones, monitors, printers, multi-function devices. Similarly to lighting, they are not suitable for DR.

2.4.1 Controllable Loads: Flexible and Deferrable

The main concept behind Demand Response (DR) derives from the potential of some loads called "controllable loads", thus making use of already existing components of the grid. DR services can be procured by electricity system operators through monitoring, aggregation and control of loads and distributed generation to maintain reliability of electric power systems. As described by S. Kawachi et al. [72], an ideal controllable load has:

1. minor loss of convenience by control of power consumption,
2. large enough power consumption and
3. response speed to signals fast enough to compensate to fluctuation in power systems

However, there are no ideal controllable loads, yet some loads satisfy partially these requirements and are treated as controllable loads in practice.

Based on the service that is provided, different loads or groups of loads are utilized. For instance, T. Masuta & A. Yokoyama [77] simulate frequency control with water heaters and EVs, which can be switched on/off for short intervals (in case of high frequency fluctuation) without affecting the quality of service. In [90], I. Hernando-Gil et al., make use of wet loads, but in this case shifting the load's operating time to achieve peak demand reduction. Thus, it is important to identify which services can be provided, by which controllable loads, when and in what volume.

Controllable loads fall mainly into two categories. The first type includes those who can provide balancing services, through altering or interrupting their cycle for a short amount of time without affecting the quality of service [38, 40, 62]. Most TCLs are such loads, EV and potentially wet loads. They can be switched off (or even reduce their consumption in the case of EVs) upon a dispatch signal, for a few minutes, as long as the battery gets fully charged or the temperature is within the thermostat's limits [77]. These are known as flexible loads.

Flexible load: *A controllable load, which can fulfil its nominal operational objectives under different cycle profiles. Alternation of its profile occurs upon an external command/signal.* Flexible loads are suited for Dispatcahble DR, thus in case of an emergency a central point can coordinate them properly (power volume, ramp up/down rates, duration) [68].

The second type, referred here as deferrable load (also found in literature as load shifting), can shift its operation in time [60, 88]. For instance, a washing machine or a dishwasher can be programmed to postpone (or advance) its start time for a

more favourable tariff (i.e. lower price due to excess RES generation or off-peak use) [38, 61, 91]. In general, some TCLs, EVs and wet loads can be deferrable.

Deferrable load: *A controllable load, whose operation can be moved in time under some limits without affecting "quality of service". Changes in time of operation are expected mainly due to dynamic pricing, received through smart meters.* Deferrable loads are suitable for Non-Dispatchable DR (indirect). Response to price changes depend on demand-price elasticity. Because of its nature (human behaviour), it can vary greatly between load types but also in time (even when assisted by automated systems [67]. This can be seen in [51], with mixed responses from users regarding acceptable start delays.

Cold loads are plugged in 24/7 and their thermal storage means they could be used for both Dispatchable and Non-Dispatchable DR, yet their thermal time constant (form of thermal inertia) is not as high as space heating/cooling or water heating. On the other hand, storage space/water heating has been used in basic forms of Non-Dispatchable DR such as overnight tariffs for a long time, storing thermal energy overnight and using it during the day. Cold loads, due to their 24/7 availability and high aggregated demand are great candidates for short term response (Dispatchable DR, e.g. FFR).

An estimated breakdown of UK's residential demand in flexible and deferrable demand according to Table 2.6 and DECC's average annual residential demand can be seen in Figure 2.24. The main flexible loads are considered to be cold loads, electric heating (space and water), whilst main deferrable loads are considered to be wet loads.

Table 2.6: DR potential of basic load types [68].

Load type	DR Potential	Main affecting factors
Cold Appliances	Dispatchable*	Human behaviour
Electric space heating	Both	Weather
Electric water heating	Both	Human behaviour
Circulation pumps	Both	Weather
HVAC	Both	Weather
Heat pumps	Both	Weather
Electric Vehicles	Both	Human behaviour
Wet appliances	Both**	Human behaviour

** Can be used for both but better fit for specified DR type*

*** Preferably not for fast Balancing Services where a spinning cycle would be interrupted to avoid QoS issues*

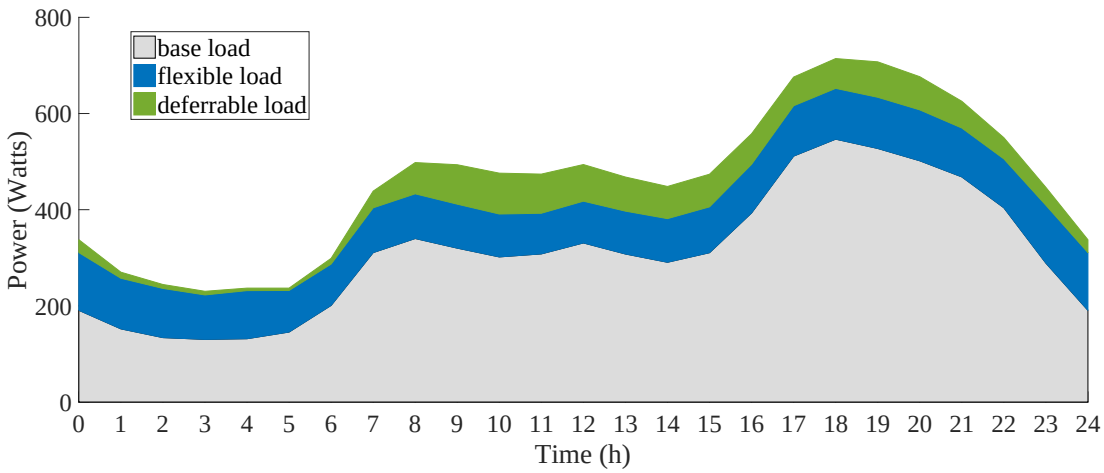


Figure 2.24: UK's average residential demand; breakdown to base, flexible and deferrable load [68]

.

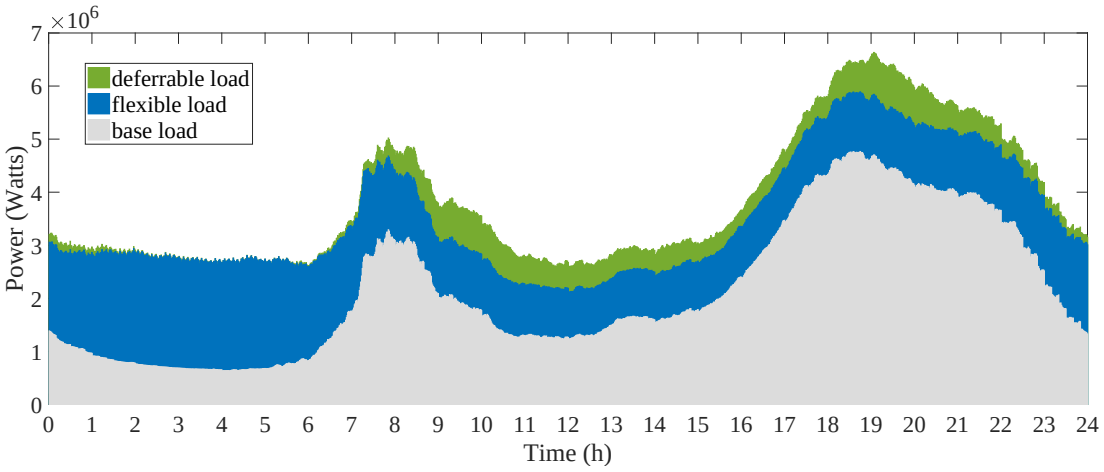


Figure 2.25: UK synthetic residential demand for 10,000 households, winter; breakdown to base, flexible and deferrable load.

.

Note that HVAC, Circulation Pumps and Heat pumps are not part of this figure, as DECC's data did not include them. A similar estimated breakdown for winter for 10,000 households' synthetic demand is given in Figure 2.25. The main difference are seen overnight, due to the higher use of electric storage water and space heating loads, as well as considering Circulation Pumps. Throughout the rest of the day both flexible and deferrable loads are very similar. The total demand between Figures 2.24 and 2.25 is very similar.

2.5 Virtual Power Plants

Clusters of units/sources can be seen as virtual micro-sources combined in an aggregated one (e.g. Figure 2.26). A virtual power plant (VPP) or aggregator is a cloud-based distributed power plant, which aggregates such heterogeneous micro-sources for the purposes of operating in power systems as single entity

and provide various services. Such examples of aggregators/VPPs exist in the United Kingdom [80], most of which provide balancing services, mainly frequency control, reserves and demand turn up, whilst also providing services to their clients (micro-sources), including but not limited to triad management, energy saving and monitoring of equipment.

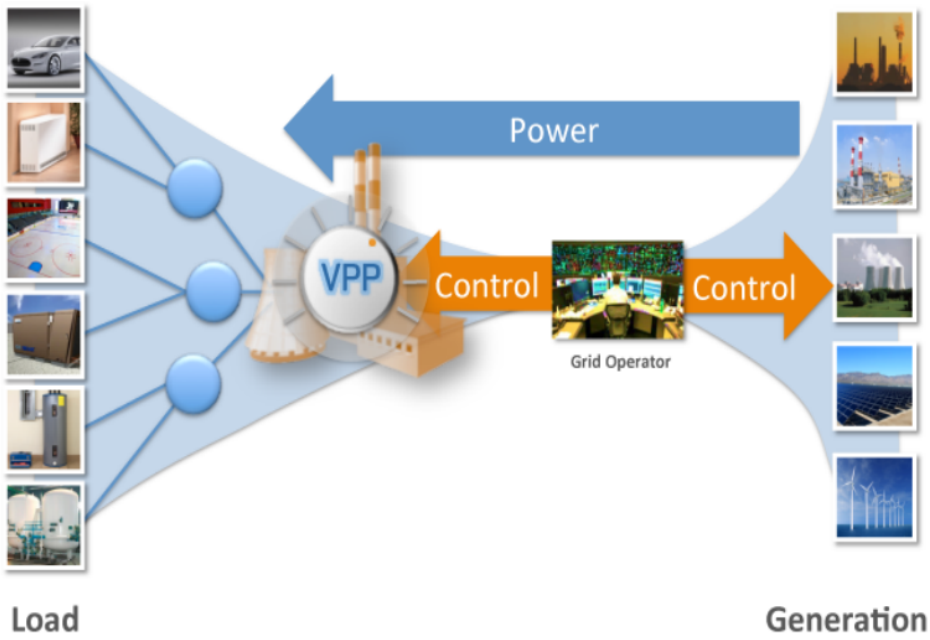


Figure 2.26: Virtual Power Plant example: EV, heater, cooling, water heating, cold load. Source:[92]

A VPP/aggregator may include types of micro-sources to give a reliable overall power supply. These can be either for dispatchable and non-dispatchable DR, depending on deferrable or flexible load. Usual DG systems include microCHPs, diesel-fired generators, natural gas-fired reciprocating engines, small-scale RES, hydroelectricity plants, small hydro, biomass, back-up gensets, and energy storage systems (ESS).

Given the nature of balancing services provided, accuracy and response within the specified limits is important. As such knowledge of available aggregated

power/energy in real time is important. For DR in specific, information of controllable loads is required, their current state and demand profile (volume and time). The next step is using appropriate aggregation models, to represent the entire cluster (based on the acquired information), and most importantly model the response of the aggregated model to DR actions.

VPPs and DR potential:

1. High amount of balancing capacity already in place (loads)
2. Loads are ubiquitous in the grid, this gives the possibility for spatial control
3. Reliable due to practically a 24/7 availability of loads and statistically safe
4. Instantaneous response (depends on communication only)
5. Emission free
6. No fuel requirements, thus cost "free"

Aggregators usually incorporate various sources in their portfolio, such as a mix of back-up units, ESS and DR. Flexitricity and GridBeyond are such examples of aggregators in UK, with heating loads, cooling loads, cold loads, fans, pumps, air conditioners, compressors, crushers and more. These loads can be turned off for short durations without cost or emissions and provide the same effect as a back-up unit or ESS for balancing. Back-up units or ESS might become unavailable, but DR is available as long as there is demand, which means always.

There are also various system benefits, such as extra balancing capabilities at low cost, the ability to deliver peak load reduction, load-following power generation or intermittent RES following demand on short notice. To some extent VPPs/aggregators can replace some of the conventional power plants operations while providing higher efficiency and more flexibility.

In order for a VPP/aggregator to participate in power systems using DR, there are some key challenges:

1. Minimum total capacity requirements to participate in balancing services (each type of service has its own)
2. Coordination and dispatch of large numbers accurately (scalability, aggregation methods)
3. Heterogeneity (in loads and in operation)
4. Monitoring large aggregated populations in real time

The availability of loads for dispatchable and non-dispatchable DR can be conflicting. In such case Dispatchable DR is most likely to take priority. Dispatchable DR is becoming increasingly crucial for power systems and is also commercially more attractive. Expected annual revenues, in California, from participation varies between \$10 and \$220 per TCL per annum, while load following and energy arbitrage are modest between \$2 and \$35 per TCL per annum [93]. Therefore, Dispatchable DR is expected to be the first DR of choice in most cases. At the same time, it has the highest precision requirements, and thus modelling, aggregation, estimation and control challenges. Models developed for Dispatchable DR are usually also accurate enough for Non-Dispatchable. Price-demand elasticity is the only additional tool needed in that regard.

As it is obvious from Table 2.6, TCLs are currently the most important type of load for DR [68]. Additionally, when both Dispatchable and Non-Dispatchable DR are possible, it is important to know that using both simultaneously is feasible, but with limitations. For example, electric heating as previously mentioned [32, 38, 72, 74, 73] can be used for DR, but its availability depends on weather conditions and human behaviour. If a low price signal caused the heating to operate at a given period (t_1, t_2), it should be anticipated (or able to be tracked) that there is extra load available for balancing services for that period and vice versa.

Approaches for the identification of appliances have been made, such as in [94], where H. Niska uses load clustering to extract information, whether electrical heating (a controllable load) is installed, by checking users' load profiles. A different method is proposed by Y. Lin & M. Tsai [95], using appliance signature recognition, where specific electrical signatures matching certain appliances can be spotted from a user's profile. Ledva et al. used disaggregation specifically to identify loads for DR on substation and feeder level in real time [41, 42].

2.6 Conclusions

This chapter discusses the potential of the residential sector to participate in DR services. These can categorized in DR for Balancing Services, such as FFR used by many aggregators in UK or DR for dynamic pricing, such as TOU tariffs which are more common with commercial consumers. An analysis of the domestic sector, common household types and their demand as well as explore the DR potential of common loads.

In that analysis, it is identified which type of DR those loads are compatible with Ancillary Services or Dynamic Pricing, and are categorized as flexible and deferrable loads accordingly. Specific loads can be perceived as both flexible and deferrable, in which case, one of DR types will probably out-weight the other. It is also important to check how much of the domestic demand can be expected to be able to participate in DR (as seen in Figures 2.24, 2.25 for UK). Low DR capability for a load type, or high uncertainty in availability for services such as FFR, makes further efforts in DR schemes questionable, especially when considering the cost and complexity of implementation.

Based on this analysis, TCLs are considered the best candidates for DR among residential loads, especially for Balancing Services which are crucial for Power

Systems with high RES penetration and will be even more important in the future. Thus the outcome of this analysis in Chapter 2 prompts to further investigate TCLs for DR in the later chapters.

An interesting point that comes from looking at households based on employment status is the (expected) activity during working hours, as seen in Figures 2.11 - 2.15, which shows not only the demand potentially available for DR but also probability of interaction. This for TCLs means heating (which is more like to be on during office hours for unemployed occupants) and interaction with cold loads, as well as changing the ambient temperature that TCLs see. This probability is later considered in Section 3.3 for ambient temperature together with heating preferences (settings) in different house types (Figure 3.2) and for human interaction with statistics from Smart-A project (Figure 3.3 [51]). These are inputs for simulation results (Section 3.8), as explained in Section 3.7. Also, the cold load data from DECC (UK) of Section 2.4 is compared in Figures 3.23 and 3.25 of Chapter 3, where it used for validation of simulated aggregated cold load demand.

NB: Dispatchable DR (Balancing Services) requires higher accuracy and fast response requirements than non-Dispatchable DR (Dynamic Pricing). In general different models are used for each, yet models of the former can be used for the latter if price-demand elasticity is combined and then price signal act as "external" commands.

Chapter 3

TCLs' Dynamic Behaviour in Time: Analysis and Modelling

3.1 Introduction

This Chapter concerns the non-static aggregated behaviour of TCLs, when no DR actions are in place. The steady state aggregated consumption of thermal loads is assumed almost static (especially in the case of cold loads), with the addition of white noise (Wiener process) [26, 29, 32, 69]. A Wiener process is introduced to model human behaviour and other external factors such as ambient temperature; "the noise term $W(t)$ aggregates the effect of external disturbances, e.g., door openings, changes in food content, and variations of θ_a"[1] as stated by Vrettos et al [96] for cold loads.

NB: A Wiener process is a continuous-time stochastic process described by a

[1] where $W(t)$ Wiener process and θ_a ambient temperature

Gaussian distribution with zero mean value and increment. As such, when a large number of units are aggregated the total sum tends towards the mean, 0.

In reality though, as shown in this Chapter (but also expected when looking at Figure 2.18 of Chapter 2), the sum of the effects of external factors is not 0, but rather it is driving the trend of TCLs' aggregated consumption. This has various impacts in provision of DR services; the aggregation models, the control actions and expected (or estimated) consumption, thus demand and availability of DR. Additionally, rebound effects after DR actions are not captured accurately. Though the extend of those has to be examined and this is one of the main purposes of this Chapter. For instance this can be seen in Figure 2.18 for Cold Loads, which most would initially assume to have flat consumption due to been connected 24/7. Note that domestic cold load contributes at least 36% of total domestic load in summer and at least 25% in winter, but there are large variations between households [59], which highlights their importance for DR.

In this Chapter, external factors and their effect on TCLs' demand are analysed. Those are separated in direct human interaction and external temperature changes (or indirect human interaction which causes them). A direct interaction can be a change on temperature settings, turning on/off the TCLs or interacting with them (e.g. opening the door of a refrigerator). Indirect interaction can be one that causes a change in ambient temperature or the thermal losses of a TCL without using it. Examples are opening windows or cooking or turning on the heating which changes the ambient temperature of a freezer. These are modelled as an extra factor, using TOU surveys and empirical or experimental data on the effect of those factors. Equivalent models of first order ordinary differential equations (ODE) are introduced where cold loads behaviour is a function of more than one thermal equations (i.e. multi-compartment colds loads or heating loads with more than one ambient temperatures, resulting to second order ODE). A state-of-the-art realistic model is developed to incorporate the above. Lastly,

bottom-up MC simulations of the model are compared to real world data of cold loads for validation.

The remainder of this Chapter is organised as follows, in Section 3.2 real world data of cold loads is presented and the divergence of existing TCL models is shown. In Section 3.3, non-linear consumption factors are analysed based on experimental data (mostly on cold loads). In Section 3.4 the effect of stochasticity due to population size and heterogeneity is investigated. Afterwards, in Sections 3.5 a more realistic TCL methodology is proposed, taking the above into consideration. This is done by modelling the direct and indirect external effects (human interaction and weather/temperature). TCLs' operation, especially when multiple of them are considered or multi-compartment (such as in case of Cold Loads), the accurate representation is normally given in second order ODE. This though requires a different computation model to be used; instead a virtual "equivalent" first order ODE is created for such Cold Loads in Section 3.6. Section 3.7, using Sections 3.5 and 3.6, describes the methodology of creating realistic bottom-up aggregated TCL populations, via Monte Carlo. Simulation results are shown in Section 3.8 and are compared to real world data.

3.2 Real World Data and Aggregated TCLs Analysis

The TCL model used in state of the art, for instance by Callaway [28] who focuses on finding the exact continuous solution of CPFE (homogeneous TCLs) or by Koch et al.[29] where an approach based on discretization is used for transition probabilities, was introduced by Mortensen and Haggerty in 1988 [97]. It is a computational model, discrete-time discrete-state Markov chain, created to study

the synchronization of TCLs which caused increased demand in short term and oscillation.

The magnitude and duration of the overload caused by TCLs, mainly space conditioning, heating and cooling loads, after an outage were the main focus by Ihara & Schweppe in 1981 [8] and by Mortensen & Haggerty in 1988 [97] and not the normal variation in TCL demand throughout operation. Under normal operation a large population of TCLs operates in an unsynchronized manner, but after an outage, many of them will start at the same time, causing an initial partial synchronization. The transient load and oscillatory behaviour after the outage is important to determine the total overload post outage. The purpose of this model was to model the thermodynamics of individual TCLs in a large population through data collection and computation of a reasonable effort. An effect similar to that is created when DR commands are given to TCLs to switch off for short term, since they also cause partial synchronization.

The model assumes a "snap shot" of the TCLs' demand in a given moment and explains the rebound effect after brown outs and/or black outs. The demand at that period may be assumed as the demand likely to occur in a similar given moment (date, day period, weather conditions), with a small "noise" factor for stochasticity. The model's equations are:

$$\theta_i(n+1) = a_i \cdot \theta_i(n) + (1 - a_i) \cdot [\theta_{a,i} - \mu_i(n) \cdot \theta_{g,i}] + \Phi_i(n) \tag{3.1a}$$

$$a_i = e^{-\frac{\tau}{C_i \cdot R_i}} \tag{3.1b}$$

$$\theta_{g,i} = R_i \cdot P_i \tag{3.1c}$$

$$\mu_i(n+1) = \begin{cases} 0, & \text{if} \quad \theta_i(n) < \theta_{\text{set},i} - \delta_i/2 \\ 1, & \text{if} \quad \theta_i(n) > \theta_{\text{set},i} + \delta_i/2 \\ \mu_i(n), & \text{otherwise,} \end{cases} \tag{3.2}$$

$$P_{tot}(n) = \Sigma_1^{N_{TCL}} P_i \cdot \mu_i(n) \tag{3.3}$$

where θ the temperature of the TCL, n the integer-valued time step, θ_a the ambient temperature (lower than θ, θ_{set} for heating loads and higher for cold and cooling loads), μ (a dimensionless discrete variable equal to 0 (off) or 1 (on), θ_g the temperature generated by the TCL (positive for heating loads and negative for cold and cooling loads), $\Phi(n) \sim N(mean, \Sigma)$ denotes multivariate normal (or Gaussian) distribution with mean vector $mean$ and covariance matrix Σ, C thermal capacitance, R thermal resistance, θ_{set} temperature set-point, δ temperature dead-band. TCLs are indexed by i, which takes on a unique value for each TCL in the population. The n-dimensional (discrete-time) stochastic processes $\Phi(n)$ denotes (Gaussian) process noise assumed to be an i.i.d process.

Even though this model is suitable to represent a "snap shot" of TCLs' population for a given moment (with knowledge of similar conditions, based on historic data), it does not reflect the dynamic nature of TCLs' population throughout the day, or the main driving factors behind this behaviour. As seen in some of these studies, the aggregated profile has significant differences to actual demand profiles. The best showcase being cold loads, which are plugged in 24 hours a day, where their aggregated demand is displayed as an almost straight line with a small fluctuation (noise). In reality, throughout the day, cold loads' aggregated demand follows the trend of the rest demand, lower overnight and higher during the day. Peaks occur around the same time as the total demand, in both residential and commercial cases [51, 52] (Figure 3.1), which can be attributed to human interaction (door opening) and heating when residents are present. El-Férik & Malhamé [98] proposed an identification algorithm to calculate and update the parameters of the model, thus be able to cope with real world changes; it was concluded that sampling intervals should be around 15 minutes or less. Between each update, assumptions of relatively steady ambient temperature (θ_a), Wiener noise with

small σ (practically no significant external disturbance) and homogeneity are required, which might not be the case in reality.

The changes in consumption can be attributed mainly to ambient temperature changes and human interaction [99]. It is important to note here that only part of the population causes this increase, thus that part has on average a higher increase than the mean displayed, it also varies between individuals. This is a good indication of the cold loads' population dynamics and how even identical appliances (homogeneous population characteristics) will behave differently in time due to external factors (heterogeneous population in operation, duty cycle).

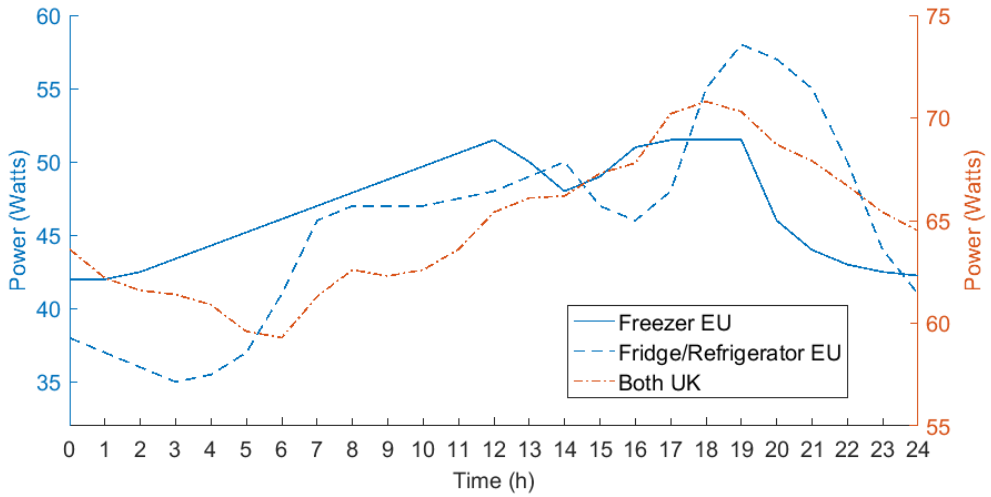


Figure 3.1: Average measured daily demand of domestic cold loads, EU and UK. The difference between maximum and minimum value for UK is around 20% and for EU around 60% and 20% respectively. Sources: Smart-A, DECC. [51, 52]

This is mainly due to human behaviour and the effect of human interaction on cold loads' power consumption as discussed in [99, 100]. In both studies, it was concluded that it was the most important factor in energy consumption. The overall deviation in domestic demand (and behaviour) was studied in [68], which was based on a detailed model of domestic demand using real data from a Time of

Use Survey (TUS) [50], as shown in Figure 2.15 earlier. The lowest demand for the second cluster (employed occupant) is attributed mainly to passive consumption of cold loads, with minimal to no interaction, whilst the first cluster shows human presence; interaction with appliances and most likely different household temperature (heating). As such cold loads' operation of even identical appliances (parameters R, C, P of (3.1a)) will differ between these 2 groups due to ambient temperature and human interaction.

The dimensionless discrete parameter $\mu(t)$ defines whether a TCL is in *on* or off state. The evolution of the discrete state $\mu(t)$ is governed by the dead-band δ (thermostat setting) and switches from 1 to 0 when $\theta(t)$ reaches θ_{off} and vice versa for θ_{on}. An important note here is that for discrete models, this transition is assumed to occur **on discrete time steps only**. As such, the evolution of the continuous state $\theta(t)$ depends on $\mu(t)$ and the discrete state $\mu(t)$ transition depends on $\theta(t)$. This means that the model consists of two interconnected subsystems, a linear continuous one and a non-linear discrete one. Also, μ has a stochastic nature as θ and δ are stochastic. For a large enough number of loads N_{TCL} (Kolmogorov's law) and assuming that $\mu_i(t)$ is an i.i.d process with (common) mean value $E(\mu(t))$:

$$\bar{\mu(t)} = \frac{1}{N_{TCL}} \Sigma_1^{N_{TCL}} \mu_i(t) \approx E(\mu(t)) \tag{3.4}$$

This can be assumed for both homogeneous and heterogeneous populations, albeit with some error. The interesting point is that it is connected to an operational characteristic of TCLs, the duty cycle (D), defined as the runtime ratio within a cycle, $D = t_{on}/T$, where t_{on} the *on* state duration and T the cycle duration. The probability of a TCL in a random moment within a period to be *on* will be equal to its duty cycle D and the probability to be off is equal to $1 - D$. At the same time, the probability of a unit to be *on* in a random moment within a period is actually the probability of $\mu_i(t)$ to be 1, so $E(\mu_i(t)) = D_i(t)$.

$\Sigma_1^{N_{TCL}} E(\mu_i(t)) = \Sigma_1^{N_{TCL}} D_i(t) \Rightarrow \frac{1}{N_{TCL}} \Sigma_1^{N_{TCL}} E(\mu_i(t)) = \frac{1}{N_{TCL}} \Sigma_1^{N_{TCL}} D_i(t)$, thus,

$$E(\bar{\mu}(t)) = \bar{D}(t) = \frac{1}{N_{TCL}} \Sigma_1^{N_{TCL}} D_i(t) \tag{3.5}$$

Which means that for a population without partial synchronization between TCLs, the probability of TCLs to be *on* is equal to the mean duty cycle of the population, $\bar{D}(t)$. Given no external interactions or changes, this quasi-equilibrium condition is called the natural diversity of a cycling load.

Now if we take (3.5) and multiple by N_{TCL}, we can tell that the expected number of TCLs to be *on* ($N_{on}(t)$) at any given time is equal to

$$E(N_{on}(t)) = N_{TCL} \cdot \bar{D}(t) \tag{3.6}$$

Consider also the mean power rating, $\bar{P}$, of the population ($\bar{P} = \frac{1}{N_{TCL}} \Sigma_1^{N_{TCL}} P_i$). We can then get the following approximation:

$$P_{tot}(t) \approx N_{TCL} \cdot \bar{D}(t) \cdot \bar{P} \tag{3.7}$$

This is an approximate equation (not to be confused with (3.3)) and can be used for (almost) homogeneous population, whilst there will be some error for heterogeneous ones, as one could easily observe in an example where TCLs with the highest duty cycle D also have the highest power ratings P. In general, the larger the population the smaller the expected error. As we see from (3.6) and (3.7), the duty cycle has an important link to the number of units in *on* state, subsequently to the consumption and provides valuable information. This is something that will be investigated further in the following Chapters.

Finally, a similar approach can be taken for a highly heterogeneous population

of TCLs. In which case depending on how high the level of heterogeneity is, clustering might be the best approach. The first step is to create clusters of TCLs based on their duty cycle, which changes in time. The ith cluster with population $N_{TCL,i}(t)$ and mean duty cycle $\bar{D}_i(t)$, has a total power $P_{tot,i}(t)$, which are functions of time. The total Power is approximated by (3.8). Larger cluster population $N_{TCL,i}(t)$ and smaller time steps improve accuracy of estimating total demand. An interesting note here is that some level of heterogeneity is sought-after for control dynamics, in particular desynchronizing the population after control actions or other causes of partial synchronization [28].

$$P_{tot}(t) = \Sigma[P_{tot,i}(t)] \approx \Sigma[N_{TCL,i}(t) \cdot \bar{D}_i(t) \cdot \bar{P}_i(t)] \tag{3.8}$$

Evidently, the population's demand is approximately linearly dependent on the duty cycle. Upon synchronization (e.g. a DR signal or power outage) this is no longer the case, synchronization causes units to operate simultaneously in phase and natural diversity will be re-established slowly in time due to heterogeneity and random factors or through corrective control actions [28]. This synchronization causes the rebound effect and its magnitude is based on the number of synchronized units.

Another note here regarding real cold load demand is that the aggregated power in (3.3) is not completely accurate, since it assumes that units in off state consume no power. In reality most TCLs maintain a small idle consumption for electronics operation and thermostat control. For the purpose of this thesis though it can be neglected, which is the norm in literature as well.

3.2.1 DECC and Smart-A Project Data

As this Chapter is looking into how TCLs' demand (and particularly focuses on Cold Loads) changes throughout the day and examines the reason behind this behaviour as well as trying to model it, it is deemed important to use validated official data for comparison. These come from the Department of Energy and Climate Change (DECC) and Smart-A project.

DECC was created on October 2008, by then Prime Minister Gordon Brown to take over some of the functions related to energy of the Department for Business, Enterprise and Regulatory Reform, and those relating to climate change of the Department for Environment, Food and Rural Affairs. It released a major White Paper in July 2009, setting out its purpose and plans. The majority of DECC's budget was spent on managing the historic nuclear sites in the United Kingdom, in 2012/13 this being 69% of its budget spent through the Nuclear Decommissioning Authority. On July 2016 DECC became part of the Department for Business, Energy & Industrial Strategy.

DECC contains a variety of data for UK energy consumption [52]. 'Energy Consumption in the United Kingdom' is an annual statistical publication that provides a comprehensive review of energy consumption and changes in intensity and output since the 1970s, with a particular focus on trends since 2000. It covers the following key Chapters, Overall energy consumption in the UK, Energy intensity by sector, Primary energy consumption, End uses and Electrical products consumption and stock.

The data of importance here is the 'Household average daily electricity consumption by appliance type' which includes average daily (24h) consumption (Wh) for the following loads:

1. Cold Appliances
2. Cooking
3. Lighting
4. Audio-visual
5. ICT
6. Washing/ drying/ dishwasher
7. Water heating (electric)
8. Heating (electric)
9. Showers (electric)
10. Other
11. Unknown

This data shows average electricity use profiles from 250 households in UK, monitored over 12 months using meters on total electricity use and main appliances as per above. Unfortunately, it does not report the data in detail for weekdays and weekends or per season. Yet it is official data from a UK government body, thus is used for validation. The data is given in average consumption per load type per year. A more preferable official data would have been per load type, per season, separately for weekdays and weekends for UK. Unfortunately such official data was not found at the time of writing from any source.

The project "Smart Domestic Appliances in Sustainable Energy Systems (Smart-A)" is an EU project led by the University of Bonn, with partners from Austria, Belgium, Germany and the United Kingdom. It is also co-funded by the Federal Ministry for the Environment, Nature Conservation and Nuclear Safety.

Smart-A aims to develop strategies in which smart domestic appliances can contribute to load management in future energy systems. In order to do this, the project assesses the options for load-shifting (for dynamic tariff schemes) by a variety of appliances across Europe and compares these with the requirements

from energy systems both on the local and regional level. The technical aspects of the assessment include an analysis of potential changes to appliances operation, of characteristics of local energy generation (from renewable energies and also co-generation) and of load management requirements in the larger electricity networks. The project also features a detailed assessment of the acceptance of smart appliances operation by users, and an evaluation of the usability of available control technologies and communication standards.

Data is reported for EU countries (United Kingdom, Denmark, Italy, France, Spain, Sweden, Poland, Hungary, Finland, Czech). It includes domestic load type ownership statistics per country, behavioural statistics, willingness for shifting the time of operation per load type per country and more. The data of importance here is the demand per load type which is given in average daily demand (24 hours) for the following loads:

1. Washing machines
2. Tumble dryer
3. Dishwasher
4. Oven and stove
5. Refrigerator
6. Freezer
7. Air conditioner
8. Water heating (electric)
9. Heating (electric)
10. Showers (electric)
11. Heating circulation pump

An interesting note is the high similarity in demand profiles per load type in different EU countries, with the exception of air conditioners and electric heating, where countries could be grouped in southern and northern in terms of demand profiles.

The data in Smart-A project is also given as average consumption per load type per year, without any more detail, as in per season or separately for weekdays and weekends. Fortunately for comparison purposes, amongst residential TCLs, Cold Loads have the least seasonality since they are connected 24/7 despite weather conditions and room temperature indoors varies less than outdoors temperate due to human preference.

3.3 External Consumption Factors

Since external factors affect the consumption of loads and thus their load profile and availability for DR, they have to be analysed and modelled. In order to do so, the following three step were taken (for cold loads in this particular instance, without loss of generality).

First, various experimental data on TCL parameters were analysed, such as in [28] for cooling/heating loads and specific data on cold loads ([30, 86, 69, 101]). Then, thermal properties of such loads ([102, 103]) and the effect of human interaction ([99, 100]) were studied. Lastly, using the above as a basis, simulations were carried out and validated against real-world demand profiles. In specific, real-world data was taken from UK's government DECC [52] and the EU Project Smart-A [51].

The model was created in such a manner that it may be used for either cold/cooling or heating loads. The governing Stochastic PDE (Partial Differential Equations) can be applied to all TCL types by changing the input only. The reason for choosing cold loads as case studies is the fact that they exhibit the highest DR potential; highest aggregated demand among TCLs, ubiquitous, available 24/7. Important TCLs consumption factors are:

1. Ambient temperature, a function of human behaviour and weather
2. Appliance characteristics and operation settings
3. Human interaction, preferences and socio-economic factors

3.3.1 Ambient Temperature

Thermal loads used by the industrial, commercial and residential sector make use of heat transfer mechanisms; advection, conduction, convection and radiation. For relatively small temperature changes, such as in the case of TCLs (domestic and commercial), Newton's law of cooling applies. Therefore, heat transfer and consecutively the thermal load is practically linearly dependent on the temperature difference,

$$\frac{d(\theta(t))}{dt} = -\lambda \cdot [\theta(t) - \theta_a(t) + \mu(t) \cdot \theta_g] \tag{3.9}$$

where λ thermal time constant ($\lambda = 1/(RC)$). As expected TCLs' energy consumption depends mainly on the ambient temperature and any change affects them directly. Refrigeration demand during winter is about 2/3 of the one in summer, while there is also a deviation between daytime and night time [30, 52]. Previous work on TCL modelling and simulations has assumed steady or relatively steady ambient temperature around a set value (a plethora of such can be found in literature, including [28, 31, 40, 69, 86, 104, 105]). In [106] variations of 10 ^oC in all house types during heating periods were found, which highlights the significance of θ_a in heterogeneity and deviation of heating practices. Zehir et al. [86], used real data to study cold loads' demand side management, where room temperature was also monitored, showing changes of 3 ^oC in a few minutes. Average temperature profiles for 292 dwellings in UK during February are shown in Figure 3.2 [106]. This clearly indicates the different temperature set-points between different dwellings and show the heterogeneity in human behaviour which

has a direct effect on TCL operation. This is taken into account in the model, with varying ambient set-points (θ_a) within the day and different set-points between TCLs to reflect dwelling's expected maximum, minimum and average temperature within the day. The way this is modelled can be seen in Section 3.7.3 and results of the simulated mean ambient can be seen in Figure 3.18.

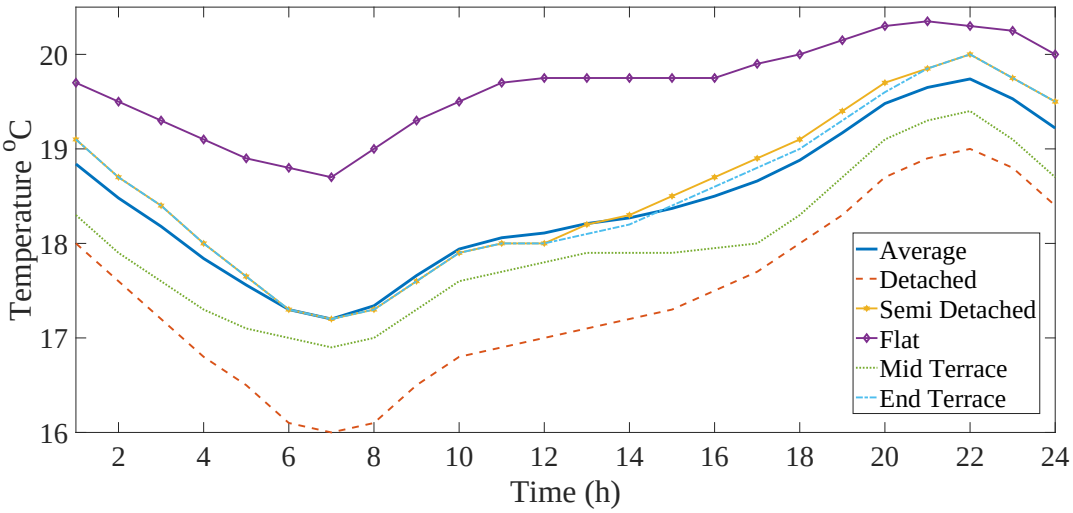


Figure 3.2: Temperature profiles per dwelling type, measured for 292 dwellings in UK, averaged daily profiles over February 2010. [106].

In Masjuki et al.'s measurements [99], energy consumption increased from 0.56 kWh/day to 1.12 kWh/day when the ambient temperature was raised from 16^oC to 31^oC, a 100% overall increase or equivalently around 37.3 Wh/day for a 1^oC increase in temperature. Hasanuzzaman et al. [100] has reported an increase in consumption from 1.2 kWh/day to 1.7 kWh/day, when ambient temperature was raised from 18^oC to 30^oC, 41.66% increase or 46 Wh/day per 1^oC. Therefore, a change in ambient temperature in short periods will have a direct effect on TCL's demand and therefore on duty cycle ($D(t)$). According to Newton's law of cooling, the thermal flow, for a given heat transfer coefficient, is a function of temperature difference with the ambient. Thus, a TCL's thermal demand depends on temperature difference to ambient. Therefore, if that difference increased

two-fold, according to Newton's law of cooling, the thermal flow (thermal load for TCLs) would also increase two-fold, which would increase the electrical consumption by the same amount given that electrical to thermal conversion remains the same. We could say, that since the set-point temperature is the mean temperature of the TCL during operation (mean of dead-band), that the change in thermal and thus electrical load is approximately equal to the change of the difference between TCL set-point and ambient. A TCL's demand is also linked to its duty cycle as described earlier. Thus the following approximation can be assumed for a TCL:

$$\frac{\theta_a(t) - \theta_{set}(t)}{\theta_a(t_0) - \theta_{set}(t_0)} \approx \frac{D(t)}{D(t_0)} \approx \frac{P(t)}{P(t_0)} \tag{3.10}$$

Note that the above equation doesn't not account for other factors that affect demand, the delay in TCLs response to ambient changes (thus delay in duty cycle change) and small changes to heat transfer coefficients under different conditions. Yet, it can still be used approximately and an example can be seen in Table 3.4.

3.3.2 TCL Characteristics and Operation

A set-point value θ_{set} is targeted by users and the temperature fluctuates between θ_{on} and θ_{off} values, which are the switching (on/off) points. Excluding external interference, thermal losses are defined by the thermal characteristics of the appliance, namely thermal capacitance C, resistance R, efficiency η and also the difference between ambient θ_a and set-point temperature θ_{set} ($\Delta\theta$ as discussed above). Consequently, set-point temperature is of equal importance to ambient temperature. Set point actuation control algorithms are based on this fact (thermal load $\propto \Delta T$), [28, 105, 107]. The effect, as expected, is essentially the same as ambient temperature θ_a changes [99].

3.3.3 Human Behaviour, Preferences and Socio-economic Factors

Human behaviour, preferences and socio-economic factors vary across consumers and TCL end function. They are multi-variable dependent and stochastic, posing the biggest challenge to model, especially because they introduce heterogeneity in operation even among identical appliances. Thus, making a homogeneous population behave heterogeneously in time, which directly affects the accuracy of aggregation models and control actions. Yet, for large populations, statistical approaches are fit for such tasks, especially because humans are "creatures of habit", thus proper examination can lead to appropriate aggregation and control frameworks.

Space heating/cooling loads operation is essentially a combination of weather and human behaviour. Human behaviour is a function of 3 connecting variables; time, comfort zones (conditions, preferences) and socio-economic factors. For instance, a household of employed individuals, is statistically less likely to use heating/cooling during office hours (individuals not present) on a given working day [68], Figure 2.15 shows the difference in demand. Once individuals are present, heating/cooling might be used; it depends primarily on weather conditions and individuals' preference of "comfort zones" but also socio-economic factors. De Cian et al. [108] examined the interaction between income, temperature and energy demand, where an income interaction model was created, examining the income/temperature elasticity of electricity demand. Additionally, in Kane T. et al's [106] work (real world measurements), the economic rebound effect was greater than expected, which is attributed to the above socio-economic behaviour. In particular, the energy savings from energy efficiency improvements has been lower than expected and that is because blanket heating practice behaviour was assumed. Thus, geographical clustering for thermal loads, such as in [105],

where relative homogeneity is assumed (air conditioners) and similar operation characteristics leads to error, θ_a alone is inadequate, even more so for TCLs with incremental human interaction. Water heating is similar, but less reliant on weather conditions or comfort zones, rather based on preferences and habits instead. In this thesis heterogeneity in operation is a key factor, and instead of blanket behaviour, distribution of expected behaviours (Section 3.7.2) and heating practices (dwelling ambient as mentioned in previous section and Section 3.7.3) have been used.

The major electric TCLs consumers, cold loads, have one external affecting factor, human behaviour. It can be broken down to door opening and loading of compartments. Experimental tests, using ISO standards [99], with door opening of 12s at a 90^oC angle, report an increase in consumption from 0.85 kWh/day to 1.42 kWh/day, for 75 such events; 7.6 Wh (or 0.894%) increase per event. In practice, during door opening the insulation alters drastically but for a short period (thermal resistance and therefore overall heat transfer coefficient), warm air mixes with cool air inside and heat transfer occurs through convection. A note at this point is that such events are dependent on $\Delta\theta$; it is the same $\Delta\theta$ during normal operation and during the event, with the change of overall heat transfer coefficient (h), thus the proportional increase (0.894%) can be assumed relatively constant. This can be expressed as:

$$Q_{normal} = h_{normal} \cdot E \cdot (-\Delta\theta) \tag{3.11}$$

$$Q_{event} = h_{event} \cdot E \cdot (-\Delta\theta) \tag{3.12}$$

$$\frac{Q_{event}}{Q_{normal}} = \frac{h_{event}}{h_{normal}} \approx const \tag{3.13}$$

where Q is the rate of heat transfer, h is the heat transfer coefficient (assumed independent of temperature and averaged over the surface), E is the heat transfer surface area. During interaction with a cold load (e.g. door opening) the

heat transfer between room and compartment increases, whilst the temperature difference (room and internal of compartment) remain practically the same. It can actually be seen as if the total thermal resistance (insulation) has decreased, since one side is not there temporarily (open door). In which sense that particular TCL's R has decreased and subsequently h has increased, increasing the rate of heat transfer for the event.

When an event occurs, only the heat transfer coefficient changes and the increase in heat transfer is proportional to this change (3.13). This means that the same event, under different ambient temperature conditions will have proportionally almost the same increase in thermal load, and consecutively in electric load.

Other experimental studies on refrigeration have actually showed that new load can have the greatest impact in consumption, especially in short term [99, 100, 102]. Zehir et al [86] mention that these effects need to be considered during simulations, yet they are hard to model. In Masjuki et al.'s measurements [99], energy consumption increased from 0.96 kWh/day to almost 2.3 kWh/day with 18kg of water added (room temperature), though not linearly. Until about 9kg the increase was linear, with a rate of about 37.5 Wh/kg; an increase of 3.9% per kg. Hasanuzzaman et al. [100] has reported an increase in consumption from 1.2 kWh/day to 1.9 kWh/day, with 12kg of water added (room temperature), about 58.3 Wh/kg, an increase of 4.83% per kg.

Taking these findings into account, as well as the fact that those are not randomly distributed during the day (Figure 3.3), it is obvious that a Gaussian process with mean equal to 0 does not properly model such effects or their impact in consumption and the TCLs populations' dynamic nature in time.

In aggregated models, such as those based on CFPE ([26, 28, 27]), Markov Chains ([29, 32, 40]) etc., none of the above (θ_a, human interaction) is included; ambient

is assumed constant or almost constant (quasi-static) and a Wiener process has 0 mean, thus inadequate to model them. For example, in the case of cold loads the effect of human behaviour can only be a positive value that increases duty cycle, such as when opening the fridge or when putting shopping goods/cooked food etc., but cannot be negative as it would require to put a cooling source in the compartment with lower temperature. Similarly, the same argument can be used for water heating (use of hot water). In the case of space heating and cooling it could be either, but the main activities will tend towards increased duty cycle. Ambient temperature's effect is almost proportional to demand and duty cycle as described in (3.10). An illustration of the error can be seen in Figure 3.1 where cold load demand varies considerably during the day (20%+). Results from the model developed in this Chapter, in comparison with no heterogeneity can be seen in Figures 3.22, 3.23 and 3.24.

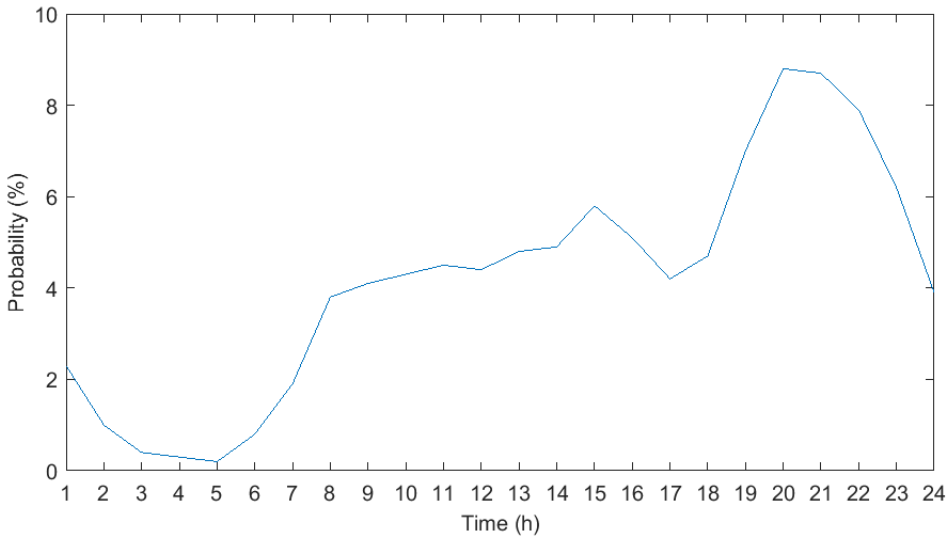


Figure 3.3: Probability distribution of door-openings per day [51].

3.4 Population Size and Stochasticity

When dealing with stochastic approaches, a limited population size is detrimental for statistical accuracy. Especially for realistic studies with heterogeneous populations, the number should be large enough to be able to model real world dynamics and maintain a degree of statistical validity. Relatively small population sizes, such as commercial units (i.e. supermarkets refrigerators), might be more appropriate to model and study using bottom up models (such as MC). These also have higher demand than residential loads and if CFPE or Euler-Maruyama approximation are used the statistical errors will be significantly higher. As Tindemans et al. [37] stated, due to TCL units being statistically independent of each other, the relative deviations from the expected mean value of TCL units *on* will decrease approximately to $1/\sqrt{N_{TCL}}$. For large N_{TCL} (3.7) holds true, but with a small error. The point is to quantify that error and select a large enough N_{TCL} for simulations.

For highly heterogeneous populations, subject to external factors, it is hard to define a minimum acceptable population size. Yet, in the case of homogeneous, free of external factors TCL population, minimum population size can be calculated simply and efficiently through Confidence Intervals. As previously mentioned, estimating available power in time is directly linked to estimation of the state of the population in that given time. It is thus imperative to define such a metric, as well a metric for estimation of its validity. The parameter that describes the state of an individual TCL (on/off probability) is its duty cycle $D \in (0, 1)$. Based on the above assumptions (homogeneous population, no external factors), this falls under a Bernoulli process. The Confidence Interval can be linked directly to duty cycle and the size of the population.

The probability of each value x of a Binomial distributed random variable X is

defined through its probability mass function:

$$X \sim Bin(n,p) \leftrightarrow Pr(X=\kappa|n,D) = f(\kappa;n,D) = \binom{n}{\kappa} D^{\kappa}(1-D)^{n-\kappa} \tag{3.14}$$

where n number of trials, κ observed outcome. Expected value and variance respectively:

$$E(X) = n \cdot D \tag{3.15}$$

$$Var(X) = n \cdot D \cdot (1-D) \tag{3.16}$$

The above also holds true for heterogeneous populations when clustered in relatively homogeneous clusters, where D is replaced by the average duty cycle for the ith cluster in a time period ($\bar{D_i}(\tau)$). Taking into consideration the operational heterogeneity (D changes in time), then clusters themselves change in time, as TCLs might be shifting to other clusters as their duty cycle changes. For instance, assuming clustering TCLs every 15 or 30 minutes according to their duty cycle, the above becomes:

$$Pr_i(X=\kappa|n_i,\bar{D_i}(\tau)) = f_i(\kappa;n_i,\bar{D_i}(\tau)) = \binom{n_i}{\kappa} \bar{D_i}(\tau)^{\kappa}(1-\bar{D_i}(\tau))^{n_i-\kappa} \tag{3.17}$$

The larger the population, the closer to the expected value in a random moment for a relatively homogeneous population of TCLs, thus a "random/uniform" distribution can be assumed with a small error. Confidence Intervals can be used as a metric; adjusted Wald, Wilson-Score, and exact Clopper-Pearson methods were considered. For large populations ($=>$ 1000) and $0.1 < p < 0.9$, any of those 3 methods yield practically the same results. For smaller populations the exact method is preferable. Such calculations of Confidence Intervals can be seen in Tables 3.1 and 3.2 for different duty cycle values.

Table 3.1: Confidence Intervals (0.95 & 0.98) for 1/3 duty cycle

D=0.3$\bar{3}$3 (p)	CI 95%			CI 98%		
Population size	Low	High	MoE(%)	Low	High	MoE(%)
1,000 (exact)	303.8	363.2	8.77%	298.6	368.8	10.39%
10,000 (adj. wald)	3241	3426	2.76%	3224	3444	3.30%
100,000 (any)	33041	33630	0.87%	32988	33677	1.03%

MoE stands for Margin of Error

Table 3.2: Confidence Intervals (0.95 & 0.98) for 1/2 duty cycle

D=0.50 (p)	CI 95%			CI 98%		
Population size	Low	High	MoE(%)	Low	High	MoE(%)
1,000 (exact)	468.5	531.5	6.30%	462.8	537.2	7.44%
10,000 (adj. wald)	4902	5098	1.96%	4884	5116	2.32%
100,000 (any)	49690	50310	0.62%	49635	50365	0.73%

MoE stands for Margin of Error

The above values of duty cycle were selected since in reality cold loads' duty cycle fluctuates usually within this region. It can be observed that for populations (clusters) of around 1,000, stochasticity is significant and their expected state would not be as accurate (statistically). The fluctuation introduced to the model from the binomial distribution will affect the accuracy of assessing the impact of the factors described above. For a population size of 100,000 it is well within limits and preferable, yet computationally slow. A model with size of 10,000 has acceptable accuracy to study TCLs' dynamics, whilst remaining computationally efficient. This is shown in Figures 3.4 & 3.5 where a small improvement ($2std/mean$) is observed from 2.97% to 2.16% whilst the computation time increases to almost 4 times.

For population size up to 100,000 (Figures 3.6 & 3.7) the stochasticity is still just above 1% (1,28% for 100,000), thus simulations with populations >100,000 are required to drop the noise introduced due to stochasticity below 1%. These figures were based on data from Tables 3.5 & 3.6 of Section 3.7. Where such data is not available, for different expected duty cycles (D) and population size, Confidence Intervals as described can be used for the expected stochasticity (error) due to population size. In reality, for a heterogeneous population, these values are expected to be slightly higher. This can also serve as a metric for clusters' minimum size for statistical accuracy.

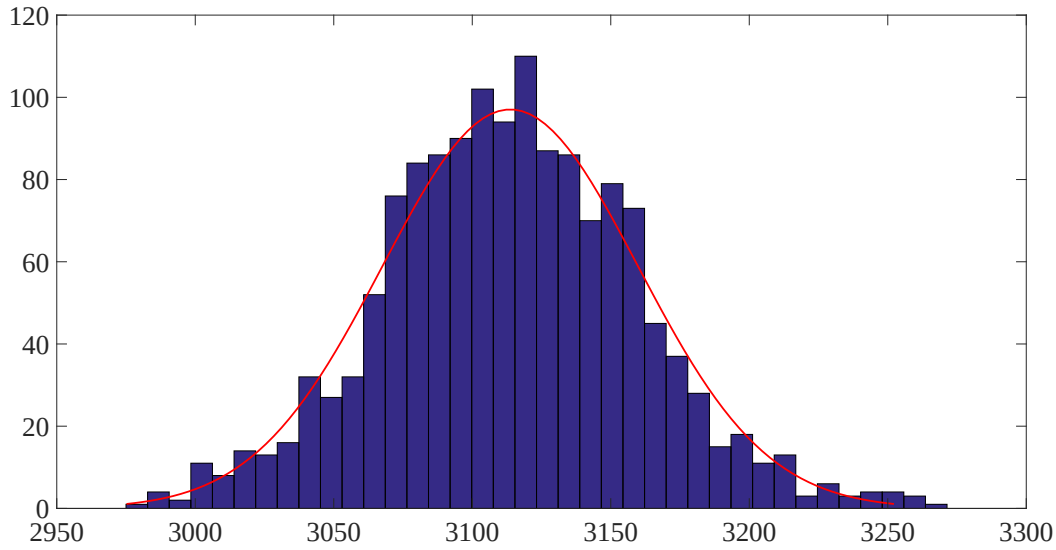


Figure 3.4: Distribution of TCL units on with a population of 10,000, TCL data is based on Tables 3.5 & 3.6, initialization (on or off) is based on each unit's duty cycle (D). Mean value 3113.7 and 2 standard deviations 92.346, percentage of 2 standard deviations to mean value 2.97%

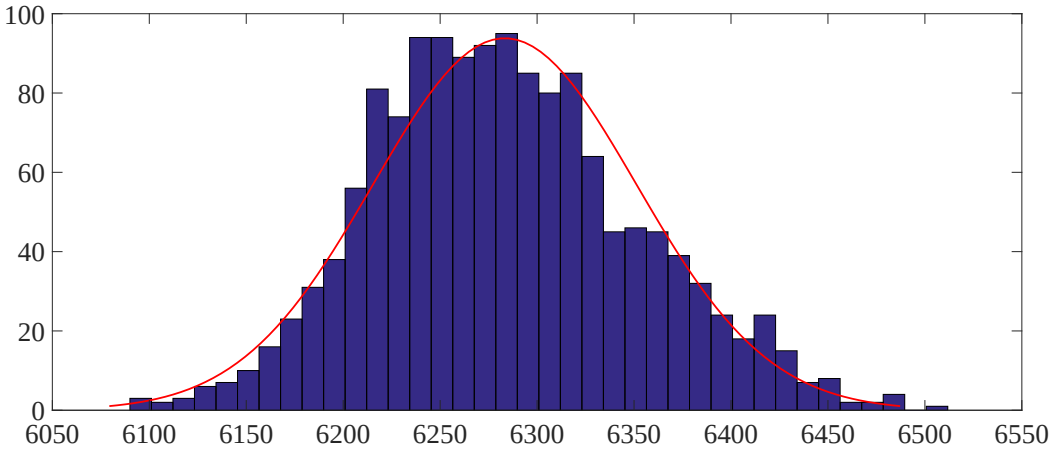


Figure 3.5: Distribution of TCL units on with a population of 20,000, TCL data is based on Tables 3.5 & 3.6, initialization (on or off) is based on each unit's duty cycle (D). Mean value 6283.3 and 2 standard deviations 135.945, percentage of 2 standard deviations to mean value 2.16%

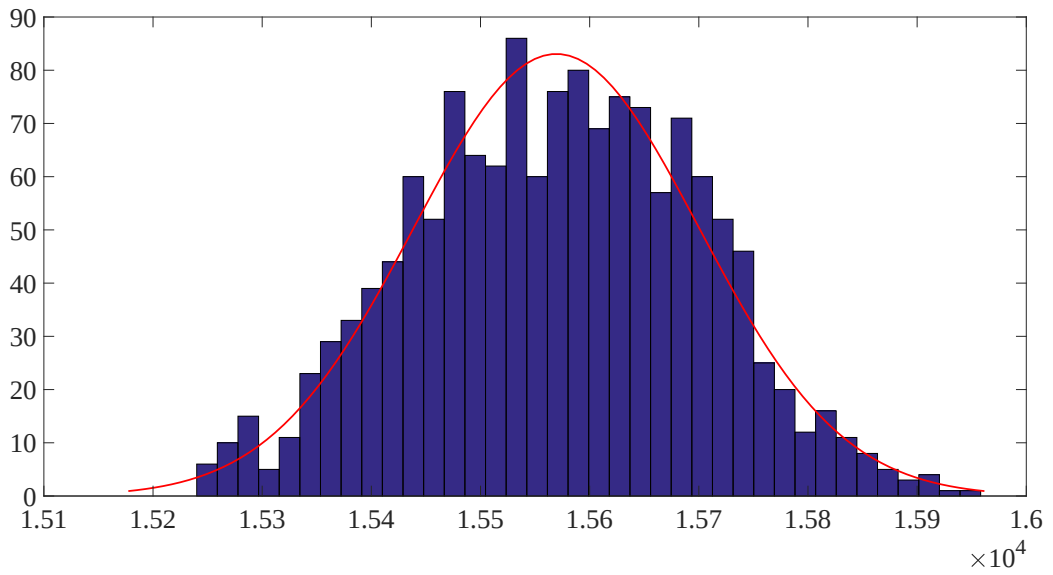


Figure 3.6: Distribution of TCL units on with a population of 50,000, TCL data is based on Tables 3.5 & 3.6, initialization (on or off) is based on each unit's duty cycle (D). Mean value 15,570 and 2 standard deviations 261.426, percentage of 2 standard deviations to mean value 1.68%

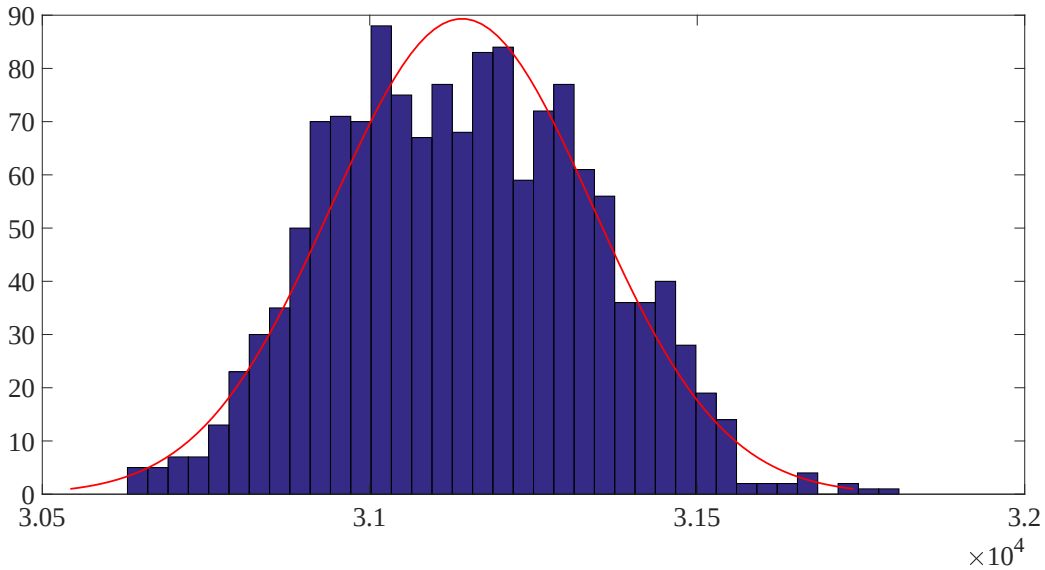


Figure 3.7: Distribution of TCL units on with a population of 100,000, TCL data is based on Tables 3.5 & 3.6, initialization (on or off) is based on each unit's duty cycle (D). Mean value 31,141 and 2 standard deviations 398.678, percentage of 2 standard deviations to mean value 1.28%

3.5 Realistic TCL Modelling and Human Factor

3.5.1 Physically-Based Model of a Single TCL

Given the previous analysis, a TCL can be described by:

$$\dot{\theta}(t) = -\lambda[\theta(t) - \theta_a(t) - \mu(t) \cdot \theta_g - \upsilon(t) \cdot \theta_e(t)] \tag{3.18}$$

Where θ_e temperature gain due to human interaction and υ a non-dimensional variable which takes values 0,1. Note that (3.18) is different to the classic one, as $\theta_a(t)$ is not constant, and $u(t)$, q_e represent interactions due to human behaviour, which in theory can be either positive or negative. In practice though, it is positive

for cold loads, negative for water heating (use of water), either of them for space heating and positive for space cooling.

Note: θ_g can be perceived as a temperature to which the system tends to, during the *on* state (but cannot reach due to switching off). If it does not switch off (meaning $\mu(t) = 1$ always), after time $t >> t_{on}$, where $losses \cong gain$, then $\frac{d(\theta(t))}{dt} \sim 0 \Rightarrow \theta(t) \cong \theta_g + \theta_a$.

General solution of first order ODE: Assume ODE $y' + P(x)y = Q(x)$, the solution to this is given by $y(x) = e^{-\int P(x)dx}[\int e^{\int P(x)dx}Q(x)dx + Const]$. Equation (3.18) can be written as $\dot{\theta}(t) + \lambda \cdot \theta(t) = \lambda \cdot \theta_a(t) + \lambda[\mu(t) \cdot \theta_g + \upsilon(t) \cdot \theta_e(t)]$.

TCL DE discrete solution: The above then has the following solution: $\theta(t) = e^{-\int \lambda dt}[Const + \int e^{\int \lambda ds}[\lambda \cdot \theta_a(s) + \lambda(\mu(s) \cdot \theta_g + \upsilon(s) \cdot \theta_e(s))]ds]$.

To get the discrete solution of (3.18) a sampling time τ will be used ($t_n = n\tau$, $t_{n+1} = (n+1)\tau$ etc.).

$\theta(t) = e^{-\lambda(t-t_n)}[Const + \int_{t_n}^{t} e^{\lambda(s-t_n)}[\lambda \cdot \theta_a(s) + \lambda(\mu(s) \cdot \theta_g + \upsilon(s) \cdot \theta_e(s))]ds]$. By placing $t = t_n$: $\theta(t_n) = e^0[Const + 0]$, thus $Const = \theta(t_n)$. The above is then written as:

$\theta(t) = e^{-\lambda(t-t_n)}\theta(t_n) + e^{-\lambda(t-t_n)} \int_{t_n}^{t}[e^{\lambda(s-t_n)}\lambda \cdot \theta_a(s)ds] + e^{-\lambda(t-t_n)} \int_{t_n}^{t}[e^{\lambda(s-t_n)}\lambda(\mu(s) \cdot \theta_g + \upsilon(s) \cdot \theta_e(s))]ds$

The same assumption as used to derive (3.1a) in the original works of Malhame [9] and Mortesen [97] will be used; μ switches state only on discrete time steps. The same will extend for υ and θ_e as well. Thus, by zero-order hold (ZOH), μ, υ and θ_e are equal to $\mu(t_n)$, $\upsilon(t_n)$ and $\theta_e(t_n)$ respectively on $t \in [t_n, t_{n+1})$:

$\theta(t) = e^{-\lambda(t-t_n)}\theta(t_n) + e^{-\lambda(t-t_n)}\int_{t_n}^{t}[e^{\lambda(s-t_n)}\lambda \cdot \theta_a(s)ds] + [\mu(t_n) \cdot \theta_g + \upsilon(t_n) \cdot \theta_e(t_n)]e^{-\lambda(t-t_n)}\int_{t_n}^{t}[\lambda e^{\lambda(s-t_n)}]ds$ whose solution is

$$\theta(t) = e^{-\lambda(t-t_n)}\theta(t_n) + [\mu(t_n)\theta_g + \upsilon(t_n)\theta_e(t_n)][1 - e^{-\lambda(t-t_n)}] + e^{-\lambda(t-t_n)}\int_{t_n}^{t}[e^{\lambda(s-t_n)}\lambda\theta_a(s)ds] \tag{3.19}$$

or equivalently by using the notation $\theta(t_n) \equiv \theta(n)$ and similarly for μ, υ and θ_e , we get:

$$\theta(n+1) = e^{-\lambda(\tau)}\theta(n) + [\mu(n)\theta_g + \upsilon(n)\theta_e(n)][1 - e^{-\lambda(\tau)}] + e^{-\lambda(\tau)}\int_{n\tau}^{(n+1)\tau}[e^{\lambda(s-t_n)}\lambda\theta_a(s)ds] \tag{3.20}$$

where now $\upsilon(n)$ will be assumed stochastic (Gaussian) in order to reflect the nature of human interaction.

3.5.2 Fit for $\theta_a(t)$ as a function of time

For space heating/cooling (heat pumps, air conditioners, electric space heating, gas heating), ambient temperature is perceived as the outdoors, which has relatively small hourly variations and even smaller on a minute scale (usual range of simulations time step). In reality the change in external temperature depends on many factors and the best fit between small time steps is a polynomial fit of small order or a linear one. For cold loads and water heaters, when heat is turned on, such as early morning or when returning home, ambient temperature can have more drastic changes. The temperature increase in this case, based on experimental data [106], seems to also follow a polynomial or linear trend (Figure 3.2). As such, a linear fit is preferred for simplicity, though in any case where $\Delta\theta_a$ is significant (maybe in an industrial environment), exponential fits might be

more accurate, yet probably still of small importance. A note here is that in the model developed, a simulation time step of 1 minute is used, with temperature readings every 1 hour or half hour. Obviously, if ambient temperature sampling is similar to simulation sampling, then $\theta_a(t)$ can be assumed constant for the equations and updated in each time step of the simulation.

Ambient temperature linear fit between hourly readings, $\theta_\mathbf{a}(\mathbf{t}) = \gamma\mathbf{t} + \beta$, discrete solutions

The solution of the integral in (3.20), given linear fit, is: $\int_{n\tau}^{(n+1)\tau}[e^{\lambda(s-t_n)}\lambda\theta_a(s)ds] = \int_{n\tau}^{(n+1)\tau}[e^{\lambda(s-t_n)}\lambda(\gamma s+\beta)ds] = [(\gamma t+\beta-\frac{\gamma}{\lambda})\cdot e^{\lambda(t-t_n}]_{(n)\tau}^{(n+1)\tau}$ or

$$(\gamma n\tau+\beta-\frac{\gamma}{\lambda})(e^{\lambda\tau}-1)+\gamma\tau\cdot e^{\lambda\tau} \tag{3.21}$$

Substituting (3.21) in (3.20):

$$\theta(n+1) = e^{-\lambda\tau}\theta(n) + [\mu(n)\theta_g + \upsilon(n)\theta_e(n) + (\gamma n\tau+\beta-\frac{\gamma}{\lambda})][1-e^{-\lambda\tau}] + \gamma\tau \tag{3.22}$$

Defining $a \equiv e^{-\lambda\tau}$, $H(n) \equiv [1-e^{-\lambda}][\upsilon(n)\cdot\theta_e(n)]$ and $\theta_a(n) \equiv \theta_a(n\tau) = \gamma n\tau+\beta$ (3.22) is simplified to:

$$\theta(n+1) = a\cdot\theta(n) + [1-a][\mu(n)\cdot\theta_g+\theta_a(n)] + \gamma\frac{a-1+\lambda\tau}{\lambda} + H(n) \tag{3.23}$$

Comparing the above with the steady θ_a case (3.1a), the dynamics of a non-constant ambient temperature are obvious. The term $\frac{a-1+\lambda\tau}{\lambda}$ defines the importance of γ, which is the rate of change of the ambient temperature. This rate also depends on the time step τ. $H(n)$ defines the effect of human behaviour, which as mentioned earlier exerts a considerable effect in cold loads' consumption. It can be modelled to include a Gaussian $\Phi(n)$ process as described earlier to induce

some noise, though not mandatory when MC simulations are used, since MC introduces stochasticity. In this case a $\Phi(n)$ process was deemed unnecessary (MC models already have noise) and it would only incur extra computational cost.

Constant $\theta_{\mathbf{a}}(\mathbf{t}) = \theta_{\mathbf{a}}(\mathbf{n}) =$ const, discrete time solutions

Practically this is special cases of the above for $\theta_a(t) = \theta_a = const$:

$$\theta(n+1) = a \cdot \theta(n) + [1-a][\mu(n) \cdot \theta_g + \theta_a(n)] + H(n) \tag{3.24}$$

This can be used for when θ_a is practically constant or updated in short intervals, where $\theta_a(n) \approx \theta_a(n+1)$. As expected, (3.1a) and (3.24) are very similar.

3.6 Equivalent Models for Multi-Compartment TCLs

The above models, as well as the ones used in the state of the art, are derived from first order ODE. In reality though, many thermal loads follow second order ODE [30, 31]. For instance, fridge-freezers, the most common cold load in households (about 69.7% ownership in 2014), with the largest average consumption per unit [52], are a case of multi-compartment thermal load and thus follow second order ODE. The latter requires different modelling to be accurate for DR and thus, difficult to form aggregation models from those. Therefore, equivalent first order ODE models are introduced.

ISO 8187, ISO 8561, and ISO 7371 are the relevant standards for testing the energy consumption of household refrigerator-freezers having two or more compartments. At least one compartment (the fresh food storage compartment) is suitable for storing unfrozen food, and at least one compartment (the food

freezer compartment) is suitable for freezing fresh food and for the storage of frozen food at -18^oC or lower [30, 99]. They may be equipped with one or two compressors. In the case of one compressor, the operating cycle is controlled by both the refrigerator's and freezer's air temperature, while commonly a damper or fan is used to assist heat transfer from the refrigerator to the freezer. This is the most common set-up due to cost. In the case of two compressors (or more), each compartment has its own independent compressor, yet there will be some synchronization as time progresses, as investigated by Leth et al. in [109, 110]. This is an expected behaviour due to the heat exchange between compartments.

In the case of one compressor (most common case), each compartment's operation can be described by (note the term $-\upsilon(t) \cdot \theta_e$ of (3.18) is excluded for simplicity, but can easily be added in the same manner as $-\mu(t) \cdot \theta_g$):

$$\dot{\theta}_f(t) = -\lambda_f \cdot [\theta_f(t) - \theta_a(t) - \mu(t) \cdot \theta_g] - \lambda_{fr} \cdot [\theta_f(t) - \theta_r(t)] \tag{3.25a}$$

$$\dot{\theta}_r(t) = -\lambda_r \cdot [\theta_f(t) - \theta_a(t)] + \lambda_{fr} \cdot [\theta_f(t) - \theta_r(t)] \tag{3.25b}$$

where λ_f heat transfer coefficient of freezer compartment to ambient, λ_r heat transfer coefficient of refrigerator compartment to ambient, λ_{fr} heat transfer coefficient of freezer compartment to refrigerator, θ_f freezer temperature and θ_r refrigerator temperature. Adding (3.25a) and (3.25b) results in:

$$\dot{\theta}_f(t) + \dot{\theta}_r(t) = -\lambda_f \cdot [\theta_f(t) - \theta_a(t) - \mu(t) \cdot \theta_g] - \lambda_r \cdot [\theta_r(t) - \theta_a(t)] \tag{3.26}$$

Similarly for the case of two compressors:

$$\dot{\theta}_f(t) = -\lambda_f \cdot [\theta_f(t) - \theta_a(t) - \mu_f(t) \cdot \theta_{g,f}] - \lambda_{fr} \cdot [\theta_f(t) - \theta_r(t)] \tag{3.27a}$$

$$\dot{\theta}_r(t) = -\lambda_r \cdot [\theta_r(t) - \theta_a(t) - \mu_r(t) \cdot \theta_{g,r}] + \lambda_{fr} \cdot [\theta_f(t) - \theta_r(t)] \tag{3.27b}$$

where μ_f, $\theta_{g,f}$ the respective parameters of the freezer compartment and μ_r, $\theta_{g,r}$

of the refrigerator compartment. Adding (3.27a), (3.27b) gives the formula which describes the operation of the appliance:

$$\dot{\theta}_f(t) + \dot{\theta}_r(t) = -\lambda_f \cdot [\theta_f(t) - \theta_a(t) - \mu_f(t) \cdot \theta_{g,f}] - \lambda_r \cdot [\theta_r(t) - \theta_a(t) - \mu_r(t) \cdot \theta_{g,r}] \tag{3.28}$$

A couple of things become apparent by observing (3.26) and (3.28). Firstly, the heat exchange between compartments does not matter in the unit's total heat exchange, though it is the cause of synchronization [109, 110]. Secondly, when compared to (3.18) and its solutions (3.23), (3.24) (or the state of the art discrete model (3.1a)), it is not possible to simulate such a unit with them. There are 2 different temperatures (θ_f and θ_r), 2 different set points, 2 different heat transfer coefficient etc. If either are used, the simulated model would be incorrect.

For which purpose, Keep et al [31], used (3.29), instead of (3.1a) or variations of it. The case of one compressor only with 2 compartments was modelled, via discretization using Euler's method and simulated for a time step of 1 minute, smaller time step simulations had no significant changes in accuracy.

$$\dot{\theta}_f(t) = \frac{1}{m_f \cdot c_f}[q_{o,f} + (q_c + q_a) - p \cdot \eta \cdot s] \tag{3.29a}$$

$$\dot{\theta}_r(t) = \frac{1}{m_r \cdot c_r}[q_{o,r} - (q_c + q_a)] \tag{3.29b}$$

where mass of contents m_f, m_r (kg), content thermal capacitance c_f, c_r $(kJ/kg - K)$, heat transfer to ambient $q_{o,f}$, $q_{o,r}$ (kW), inter-compartment heat transfer: conduction q_c (kW), inter-compartment heat transfer: air flow q_a (kW), compressor binary on/off state s, compressor real power consumption p (kW), compressor coefficient of performance η. For more details on Keep et al.'s model the reader is encouraged to read [31] (derivation of equations, assumptions, limitations etc.).

The main problem with Keep et al.'s approach, as mentioned by the authors [31], is

that this is a model specifically designed for one compressor with 2 compartments. Which means that cold loads (and thermal loads in generally) would have to be modelled separately for single compartment (freezer or refrigerator only), multi-compartment with single compressor and multi-compartment with multiple compressor (for each combination) and aggregated and controlled separately. Moreover, when more complex models than (3.1a) (or similar ones) are used, their aggregation models (e.g. CFPE) are more complex and harder to derive and subsequently their control algorithms. It would be thus beneficial to have an "equivalent" simplified model which follows (approximately) the same duty cycle and thermal dynamics (e.g. duty cycle changes to ambient) and thus enable aggregation of units following the classic model (3.1a) and the ones developed (3.23), (3.24).

Assumption: since for a single compressor (two compartments) there is one duty cycle, there should be an "equivalent" unit with approximately the same duty cycle and thermal dynamics that can be represented by the state of the art discrete model (3.1a) or (3.23) and (3.24). Also, in the case of two compressors, since they tend to synchronize, it means that their duty cycles synchronize, thus operate as "one unit".

The total thermal load losses of a refrigerator-freezer in steady state is the sum of the freezer's and refrigerator's load:

$$Q_{tot} = (Q_f + Q_{fr}) + (Q_r - Q_{fr}) = Q_{eq} \tag{3.30}$$

where Q_f heat transfer from room to freezer, Q_{fr} heat transfer from refrigerator to freezer and Q_r heat transfer from room to refrigerator. According to Newton's law of cooling, $Q = h \cdot A \cdot (-\Delta\theta)$, where $\Delta\theta$ the temperature difference to the ambient. A freezer compartment will have a temperature difference to ambient, $\Delta\theta_f$, with typical freezer temperatures between -11^oC and -17^oC

for θ_{on}, -21^oC and -31^oC for θ_{off} [30, 31, 69, 86, 100, 101, 102, 103]. A refrigerator compartment will have a temperature difference to ambient, $\Delta\theta_r$, with typical freezer temperatures between 0^oC and 4^oC for θ_{off}, $4oC$ and 8^oC for θ_{on} [30, 31, 69, 86, 100, 101, 102, 103]. A virtual "equivalent" single compartment unit then would be defined as $Q_{eq} \equiv h_{eq} \cdot E_{tot}(-\Delta\theta_{eq})$, where $E_{tot} = E_f + E_r$, E_f, E_r the surface of heat flux between f, r to ambient and $\Delta\theta_{eq}$ the temperature difference to ambient which reflects the same equivalent total thermal load of all compartments (two in this example). The aim is to calculate those temperature points (θ_{on} and θ_{off}) of the equivalent "virtual" TCL.

$$\frac{E_{tot}}{E_{tot}} = \frac{E_f}{E_{tot}} + \frac{E_r}{E_{tot}} \Rightarrow 1 = w_f + w_r \tag{3.31}$$

The heat transfer coefficients of the two compartments are not equal but similar (as expected due to similar material). The heat transfer coefficient of the "equivalent" single unit could be assumed as either of them, or the average of them or better a weighted average based on surface: $h_{eq} \equiv w_f \cdot h_f + w_r \cdot h_r$. Setting the equivalent "virtual" TCL to reflect the same total thermal load, $Q_{tot} = Q_{eq}$:

$$h_f \cdot E_f(-\Delta\theta_f) + h_r \cdot E_r(-\Delta\theta_r) = h_{eq}E_{tot}(-\Delta\theta_{eq}) \tag{3.32}$$

Dividing by E_{tot}, using (3.31) and replacing h_{eq}, the above is written:

$$w_f \cdot h_f \cdot \Delta\theta_f + w_r \cdot h_r \cdot \Delta\theta_r = h_{eq} \cdot \Delta\theta_{eq} = (w_f \cdot h_f + w_r \cdot h_r)\Delta\theta_{eq} \tag{3.33}$$

Heat transfer coefficients are the sum of the internal cabinet coefficient and the appliance's outside surface coefficient with the room. Studies have shown a difference less than 9% between freezer's and refrigerator's internal cabinet heat transfer coefficients (convective plus radiative). By adding the outside surface's coefficient, the sums' difference is less than 4% (3.7%) [100]. Due to the small difference between heat transfer coefficients of the compartments, either of those

can assumed for the virtual "equivalent" or some weighted average of those:

$$h_f = (1-e)h_r \text{ or } h_r = (1+f)h_f \tag{3.34}$$

where $e, f \in (0,1)$ (typical values of e, f are expected between $(0, 0.04)$ [100]). Thus, (3.33) can be written as:

$w_f \cdot (1-e) \cdot h_r \cdot \Delta\theta_f + w_r \cdot h_r \cdot \Delta\theta_r = (w_f \cdot (1-e) \cdot h_r + w_r \cdot h_r) \cdot \Delta\theta_{eq} \Rightarrow$

$w_f \cdot (1-e) \cdot \Delta\theta_f + w_r \cdot \Delta\theta_r = (w_f \cdot (1-e) + w_r) \cdot \Delta\theta_{eq} \iff$

$w_f \cdot (1-e) \cdot \theta_f - w_f \cdot (1-e) \cdot \theta_a + w_r \cdot \theta_r - w_r \cdot \theta_a = (w_f \cdot (1-e) + w_r) \cdot \theta_{eq} - (w_f \cdot (1-e) + w_r) \cdot \theta_a \Rightarrow$

$w_f \cdot (1-e) \cdot \theta_f + w_r \cdot \theta_r = (w_f \cdot (1-e) + w_r) \cdot \theta_{eq} \Rightarrow$

$\frac{w_f \cdot (1-e) \cdot \theta_f + w_r \cdot \theta_r}{w_f + w_r - e \cdot w_r} = \theta_{eq}$,

similarly for $h_r = (1+f)h_f$, $\frac{w_f \cdot \theta_f + w_r \cdot (1+f) \cdot \theta_r}{w_f + w_r + f \cdot w_r} = \theta_{eq}$ or

$$\theta_{eq} = \frac{\theta_f \cdot w_f + \theta_r \cdot w_r - e \cdot w_f \cdot \theta_f}{1 - e \cdot w_f} \tag{3.35a}$$

$$\theta_{eq} = \frac{\theta_f \cdot w_f + \theta_r \cdot w_r + f \cdot w_r \cdot \theta_r}{1 + f \cdot w_r} \tag{3.35b}$$

It is important to note that this virtual "equivalent" model does **not** have a physical meaning or does not try to model some physical parameter of the unit. It is a way of simulating a multi-compartment thermal load (2nd ODE and above) as if it was a single-compartment one (1st ODE) so the "state of the art" model (3.1a) or the ones developed in this Chapter (3.23), (3.24).

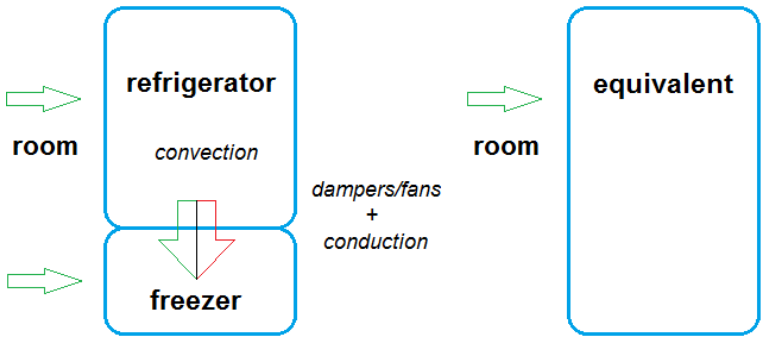


Figure 3.8: Thermal model of a common fridge with 2 compartments (on the left) & the equivalent thermal model to be used for aggregation and DR (on the right)

Heat transfer is essentially always from the room to the freezer, directly or indirectly (through refrigerator Figure 3.8), thus for normal operation the thermal behaviour is practically closer to the freezer's thermal properties, as if though the freezer was larger (larger surface for heat loss) and it's set temperature was closer to the ambient. Yet human interaction with the refrigerator, i.e. door opening, extra load etc., will have an effect (heat transfer to refrigerator's compartment) closer to the refrigerator's thermal properties. The above are relevant to each compartment's relative size, surface typical weight factors can be seen in Table 3.3 [25].

Table 3.3: Typical w_f and w_r values

Freezer's compartment (w_f)	Refrigerator's compartment (w_r)
0.33	0.67
0.27	0.73
0.36	0.64
0.35	0.65

Based on experimental data from [31, 99, 100, 102]

An examination of the above can be done with the experimental measurements from [100] in Table 3.4, where the virtual "equivalent" single compartment unit's

parameters have being calculated based on (3.35a). The proportional duty cycle change (from 0.303 to 0.429) for a 2-compartment unit is practically the same as the proportional demand change $\frac{(D(t)-D(t_0))}{D(t_0)} = 0.4158 \approx \frac{(P(t)-P(t_0))}{P(t_0)} = 0.4166$, an error $< 0.2\%$, as expected from (3.10). Yet, the first part of (3.10) **does not** hold true when looking at the refrigerator's or freezer's parameters separately. Thus when modelling such a unit, using the characteristics of the refrigerator (e.g. θ_{off}, θ_{off}) would be inaccurate, using those of the freezer would also be inaccurate and in both cases neglect that those two compartments work in synch due to the thermal flow between them [109, 110].

Table 3.4: Comparison of relative Power increase to relative Temperature increase

Specifications		Operation	
Power rating	165 W	Ambient 18 ^oC	1.2 kWh/day
Freezer factor (w_f)	0.33	Ambient 30 ^oC	1.7 kWh/day
Refrigerator factor (w_r)	0.67	ΔP (%)	41.66%
Freezer θ_{min}, θ_{set}, θ_{max}	-21.5, -16.75, -12	$\Delta(\theta_{eq,set} - \theta_a)(\%)$	55.53%
Refrigerator θ_{min}, θ_{set}, θ_{max}	1.5, 3.5, 5.5	$\Delta(\theta_{r,set} - \theta_a)(\%)$	82.75%
"Equivalent" θ_{min}, θ_{set}, θ_{max}	-6.573, -3.608, -0.643	$\Delta(\theta_{f,set} - \theta_a)(\%)$	34.53%

Based on experimental data from [100]. Where $\Delta(\theta_{set} - \theta_a) = \frac{\theta_{set}-\theta a}{\theta_{set}-\theta a_0} - 1$

3.7 Bottom-Up Heterogeneous Modelling Methodology

Even though simulations of aggregated models are much faster than the MC ones, accuracy of MC ones is higher, especially when heterogeneity is considered, relatively small populations or mixed load populations. On the other hand, they are slower to compute and do not give a state space or something that can be used for aggregated control algorithms. Rather, MC models are useful (and should be

used) to validate aggregated models and control actions (as those are given by control algorithms). A flowchart of the methodology can be seen in Figure 3.9

NB: Defrost heater power and cycles are ignored for simplicity. Defrost heaters are typically operating for less than 5% of the time [86, 104], yet their demand during this time is usually considerably higher than normal operation (e.g. 480W [86]). They operate a few times per day, with daily consumption around 0.35 kWh [86] and are more suited as deferrable loads (non-dispatchable DR) rather than as flexible loads (dispatchable DR). They do not follow the models described in this Chapter, whose focus is a realistic model to represent the dynamics of TCLs (mainly cold loads as a test case) for dispatchable DR.

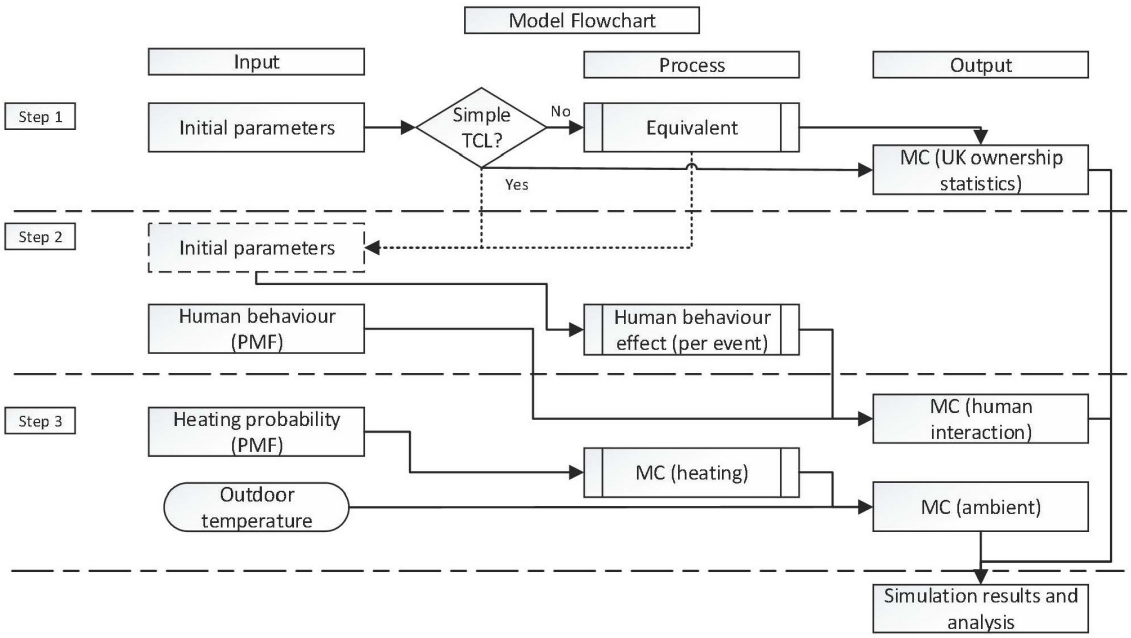


Figure 3.9: Model flowchart [25].

3.7.1 Step 1: Calculation of Basic TCLs' Parameters Using Experimental Data

It is important to mention that the values of [86] have significant variations compared to the rest; as stated by the authors, data was taken from a G class top mounted refrigerator, and was chosen as a good example for a low efficient refrigerator with high energy demand (relatively older model). According to Residential Energy Consumption Survey 2009, in USA, this was one of the most common (more than 60% percent) refrigerators in households, thus chosen by the authors for their study in 2009, also mentioning that customers mostly use the same appliance for more than 5 years and up to 14 years. Newer models (post 2009) are more efficient and similar to those used in [69]. Some fridges (refrigerator-freezer) from table 3.5 have significantly different conditions than what is commonly observed in the UK (i.e. 33^oC room temperature). In the case of UK, the values from [69, 100, 102, 103] are more suitable for a realistic model.

Table 3.5 is used to calculate the required parameters, λ, θ_g, which are constant, dependent on appliance characteristics and independent of ambient temperature and human interaction [25]. Simply knowing those is enough to populate (3.9), a representation of a TCL's physical model. This principle was also used in [98], where an algorithm was developed to identify λ, θ_g through readings of duty cycle (D). During a deterministic cycle (without external effects) and constant ambient temperature (θ_a), a TCL's off cycle is described by ($\mu = 0$)

$$\theta(t) = (1 - e^{-\lambda t}) \cdot \theta_a + \theta_0 \cdot e^{-\lambda t} \tag{3.36}$$

and a TCL's *on* cycle is described by ($\mu = 1$)

$$\theta(t) = (1 - e^{-\lambda t}) \cdot (\theta_a + \theta_g) + \theta_0 \cdot e^{-\lambda t} \tag{3.37}$$

Table 3.5: Cold load experimental data

t_{on}	t_{off}	cycle	θ_{min}	θ_{max}	θ_a	D	$P(W)$	Type
25'	75'	100'	3.5 ^oC	7 ^oC	20 ^oC	0.25	-	R
30'	65'	95'	2 ^oC	5 ^oC	20 ^oC	0.316	-	R
70'	110'	180'	4.5 ^oC	7.5 ^oC	21 ^oC	0.389	-	R
140'	110'	250'	4.5 ^oC	7.5 ^oC	21 ^oC	0.560	-	R
30'	60'	90'	4 ^oC	8 ^oC	20 ^oC	0.33	-	R
34'	66'	100'	0 ^oC	7 ^oC	33 ^oC	0.34	197	M
			-22 ^oC	-15 ^oC				
12'	36'	48'	3.5 ^oC	6.5 ^oC	20 ^oC	0.25	-	M
			-22 ^oC	-11 ^oC				
28'	65'	93'	1.5 ^oC	5.5 ^oC	18 ^oC	0.30	165	M
			-21.5 ^oC	-12 ^oC				
32'	64'	96'	2.1 ^oC	4.3 ^oC	25 ^oC	0.33	175	M
			-19.1 ^oC	-16.9 ^oC				
30'	60'	90'	4 ^oC	8.2 ^oC	20 ^oC	0.33	120-160	M
			-31 ^oC	-15 ^oC				
			-27 ^oC	-17 ^oC				F

R stands for refrigerator, M for multi-compartment, F for freezer.

Sources: [30, 31, 69, 86, 100, 101, 102, 103].

At the end of the off state, $t = t_{off}$:

$$\theta(t_{off}) = (1 - e^{-\lambda t_{off}}) \cdot \theta_a + \theta_0 \cdot e^{-\lambda t_{off}} \tag{3.38}$$

and at the end of the on state, $t = t_{on}$:

$$\theta(t_{on}) = (1 - e^{-\lambda t_{on}}) \cdot (\theta_a + \theta_g) + \theta_0 \cdot e^{-\lambda t_{on}} \tag{3.39}$$

For the off state, after time t_{off} the units switches on, thus $\theta(t_{off}) = \theta_{max} = \theta_{on}$ and $\theta_0 = \theta_{min} = \theta_{off}$, substituting on (3.38) and solving for λ results:

$$\lambda = -\frac{1}{t_{off}} ln(\frac{\theta_{on} - \theta_a}{\theta_{off} - \theta_a}) \tag{3.40}$$

Similarly, for on state, $\theta(t_{on}) = \theta_{min} = \theta_{off}$ and $\theta_0 = \theta_{max} = \theta_{on}$, substituting on (3.39) and solving for θ_g results:

$$\theta_g = \frac{\theta_{off} - \theta_{on} \cdot e^{-\lambda t_{on}}}{1 - e^{-\lambda t_{on}}} - \theta_a \tag{3.41}$$

Table 3.6: Parameters λ, θ_g (calculated at 20^oC)

Cold load type	λ (10^{-3})	θ_g (^oC)	θ_a (^oC)	D
Refrigerator	3.179	-58.815	20	0.250
Refrigerator	2.805	-52.172	20	0.316
Refrigerator	1.458	-38.541	21	0.389
Refrigerator	4.795	-41.856	20	0.333
Freezer	3.987	-125.701	20	0.333
Fridge (equiv.)	2.909	-107.401	33	0.340
Fridge (equiv.)	7.102	-87.323	20	0.250
Fridge (equiv.)	4.249	-71.512	18	0.301
Fridge (equiv.)	1.202	-85.819	25	0.333
Fridge (equiv.)	5.812	-71.123	20	0.333

Calculated using equations (3.40) *and* (3.41)

Sources: [30, 31, 69, 86, 100, 101, 102, 103].

These are used via MC to construct heterogeneous populations of TCLs (the 3 basic types of cold loads) with randomized variance in parameters θ_g, λ, θ_a, and initial conditions $\theta(t)$, $\mu(t)$; $\mu(t)$ based on *on/off* probability which is equal to

the duty cycle D of each. The population's mixture, as a percentage of each cold load type, is based on UK ownership statistics [52], which is about 107% per household, meaning that on average 7% of households have an extra unit.

3.7.2 Step 2. Human Interaction

The effect of human interaction as discussed in Section 3.3.3 is relative to $\Delta\theta$., since right before and during interaction the ambient temperature is the same. Two similar interactions (e.g. door opening of a freezer) at different ambient temperatures (thus different initial consumption) will have different impact in increased consumption in absolute terms, but similar as a percentage of their respective initial consumption. Therefore, it is best to map the experimental percentile increase in consumption to the term $H(n)$ of (3.23) and (3.24). As shown previously by (3.10) and experimental data, the increase in TCLs' consumption is proportional to increase in their Duty Cycle.

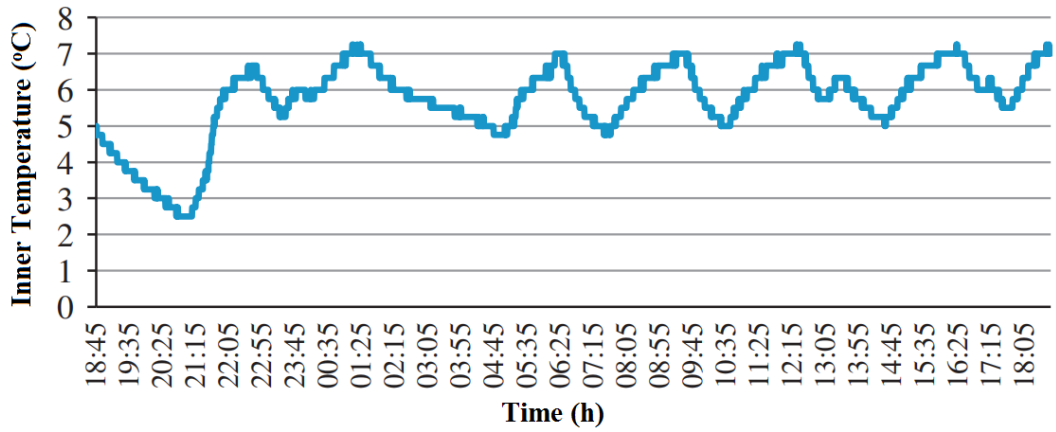


Figure 3.10: Real refrigerator measurements of inner temperature. [86]

An example of how temperature and t_{on}, t_{off} are affected can be seen in Figure 3.10, which has real measurements gained by a data logger for a refrigerator's inner temperature [86]. As it can be noticed, t_{on} and t_{off} vary, and a few sudden

temperature jumps (increase of inner temperature) are noticed whilst in *on* state, increasing the t_{on} required to reach 5 ^oC (θ_{off}).

It can be assumed that an interaction which causes x(%) increase in consumption, causes x(%) increase in Duty Cycle (D). This means an ϵ(%) increase in t_{on} or ζ(%) decrease in t_{off}, such that Duty Cycle (D) increases by x(%). In order to calculate the H(n) value which causes x(%), it is best to express H(n) as a function of ϵ and ζ. Thus we have the following two cases:

$$D^{'} = \frac{t_{on}^{'}}{t_{on}^{'} + t_{off}} = \frac{t_{on}(1+\epsilon)}{t_{on}(1+\epsilon) + t_{off}} = \frac{t_{on}(1+x)}{t_{on} + t_{off}} \tag{3.42}$$

$$D^{'} = \frac{t_{on}}{t_{on} + t_{off}^{'}} = \frac{t_{on}}{t_{on} + t_{off}(1-\zeta)} = \frac{t_{on}(1+x)}{t_{on} + t_{off}} \tag{3.43}$$

where $D^{'}$ denotes the new Duty Cycle ($D^{'} = (1+x)D$). The solutions to the above are:

$$\epsilon = x\frac{t_{on} + t_{off}}{t_{off} - x \cdot t_{on}} \tag{3.44}$$

$$\zeta = x\frac{t_{on} + t_{off}}{t_{off}(1+x)} \tag{3.45}$$

Now $H(n)$ has to be expressed as a function of ϵ and ζ. Equation (3.24) is used, where $\theta_a = const$ for the duration of the event, due to relatively short duration (e.g. opening of freezer door). When new loading is considered in addition to door opening, there is an extended time required to chill products, which can be estimated by using a simple model of natural convection around the product and conduction inside the product, as per [102]. Studies have reported that the heat removed from new food loadings accounted for the majority of cold load's increase in demand, compared to other variables (ambient temperature change, thermostat setting change, door opening) [99, 100, 102]. The quantity of heat to be removed can be calculated from knowledge of the product, including its state upon entering the refrigerating space, final state, mass, specific heat above and below freezing

temperature, and latent heat. Latent heat of fusion of a product is related to its water content and can be estimated by multiplying the percent of water in product by the latent heat of fusion of water. Once the product attains the desired cooling temperature, it does not affect the energy consumption significantly until fresh products are placed in the refrigerator again [99]. In this thesis there is no point to go into such detail, but rather use results of previous studies which quantify the impact in % increase in demand. In short, in this case $H(n)$ has a higher and more prolonged impact.

As such, $\bar{H}$ is defined as the cumulative $H(n)$ for each state (off, on). So the aim is to calculate the $\bar{H}$ which results in a x% increase in consumption (and duty cycle (3.10)), using (3.24), with $\theta_a = const.$ (even if θ_a changes later during the cycle, when new load was added, it had the θ_a at the time of event). In a sense, $\bar{H}$ can be seen as an "internal heat source" (the opposite for heating loads), which results in increased consumption by changing t_{on} (3.46) and t_{off} (3.48).

$$\theta(t'_{on}) = [1 - e^{-\lambda t'_{on}}][\theta_g + \theta_a] + \theta_{on} \cdot e^{-\lambda t'_{on}} + \bar{H}$$

$$\theta_{off} - (\theta_g + \theta_a + \bar{H}) = [\theta_{on} - (\theta_g + \theta_a)] \cdot e^{-\lambda t'_{on}}$$

$$t'_{on} = -\frac{1}{\lambda} \cdot ln(\frac{\theta_{off} - (\theta_g + \theta_a + \bar{H})}{\theta_{on} - (\theta_g + \theta_a)}) \tag{3.46}$$

using $t'_{on} = (1 + \epsilon)t_{on}$:

$$-\tfrac{1}{\lambda} \cdot ln(\tfrac{\theta_{off} - (\theta_g + \theta_a + \bar{H})}{\theta_{on} - (\theta_g + \theta_a)}) = (1 + \epsilon) \cdot [-\tfrac{1}{\lambda} \cdot ln(\tfrac{\theta_{off} - (\theta_g + \theta_a)}{\theta_{on} - (\theta_g + \theta_a)})]$$

$$\tfrac{\theta_{off} - (\theta_g + \theta_a + \bar{H})}{\theta_{on} - (\theta_g + \theta_a)} = [\tfrac{\theta_{off} - (\theta_g + \theta_a)}{\theta_{on} - (\theta_g + \theta_a)}]^{(1+\epsilon)}$$

$$\theta_{off} - (\theta_g + \theta_a + \bar{H}) = \tfrac{[\theta_{off} - (\theta_g + \theta_a)][\theta_{off} - (\theta_g + \theta_a)]^{(\epsilon)}}{[\theta_{on} - (\theta_g + \theta_a)]^{(\epsilon)}}$$

$$\bar{H} = \theta_{off} - (\theta_g + \theta_a) - [\theta_{off} - (\theta_g + \theta_a)][\frac{[\theta_{off} - (\theta_g + \theta_a)]}{[\theta_{on} - (\theta_g + \theta_a)]}]^{(\epsilon)}$$

$$\bar{H} = (\theta_{off} - \theta_a - \theta_g)[1 - (\frac{\theta_{off} - \theta_a - \theta_g}{\theta_{on} - \theta_a - \theta_g})^{\epsilon}] \tag{3.47}$$

Similarly for t'_{off}:

$$t'_{off} = -\frac{1}{\lambda} \cdot ln(\frac{\theta_{on} - (\theta_a + \bar{H})}{\theta_{off} - \theta_a}) \tag{3.48}$$

And for ζ during *off* state:

$$\bar{H} = (\theta_{on} - \theta_a)[1 - (\frac{\theta_{on} - \theta_a}{\theta_{off} - \theta_a})^{(-\zeta)}] \tag{3.49}$$

Using the above and experimental data described in Section 3.3, human behaviour effect per event for each cold load type is calculated, separately for door opening and new load. Some results of the impact of door opening and adding new load to cold loads can be seen in Figures 3.12, 3.14 and 3.15. These reflect the impact of human interaction ($\bar{H}$) and with the addition of non-constant ambient (θ_a), which alter the duty cycle of TCLs (Cold Loads in this case) during operation (as seen later in Figure 3.32 in Section 3.8) and result in the changes in demand throughout operation, as seen later in Figures 3.22 - 3.27. The procedure to obtain Figures 3.11 - 3.14 is:

- Monte Carlo Markov Chain on human interaction for every hour as per Figure 3.3 (interaction probability distribution) (alternatively any other interaction probability distribution if such data is available). This probability is converted from hours to minutes, since simulations later use a simulation step of 1 minute.
- Afterwards the impact of the event is given, in this case randomly chosen between a small event (door opening) and larger event (new load), but with different possibilities for each. Each of those potential events are represented

by Gaussian distributions. The impact of events is taken from experimental data as discussed in Section 3.3.3 ([99, 100, 102]).

- Lastly, that output is converted to $\bar{H}$ which expresses the thermal losses (or gain in the case of Cold Loads) caused, according to (3.47), (3.49). In short, that $\bar{H}$ causes a temperature gain which shortens t_{off} or extends t_{on} (or both) and alters D to $D^{'}$. For a metric, the impact of $1\bar{H}$ is assumed as a door opening for 12s at a $90^{o}C$ angle, which resulted in an increase in consumption by 0.894%. (*Section 3.3.3: Experimental tests, using ISO standards [99], demand change from 0.85 kWh/day to 1.42 kWh/day, for 75 such events*)

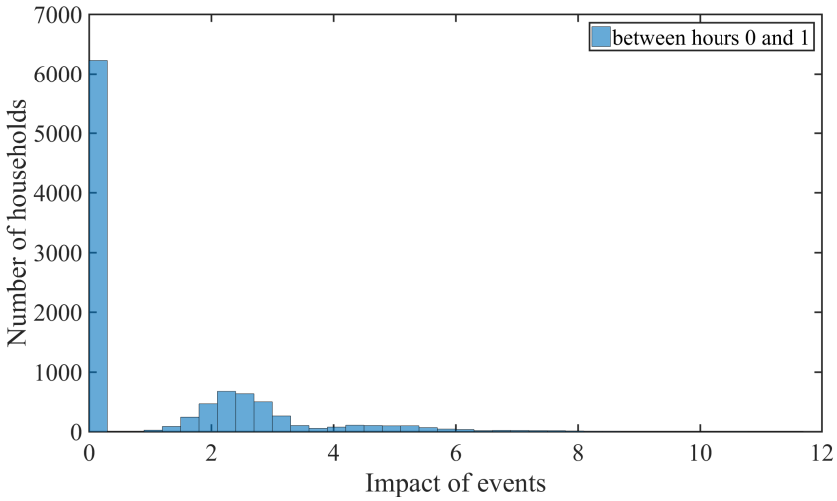


Figure 3.11: Human interaction impact: distribution between 0am and 1am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^{o}".

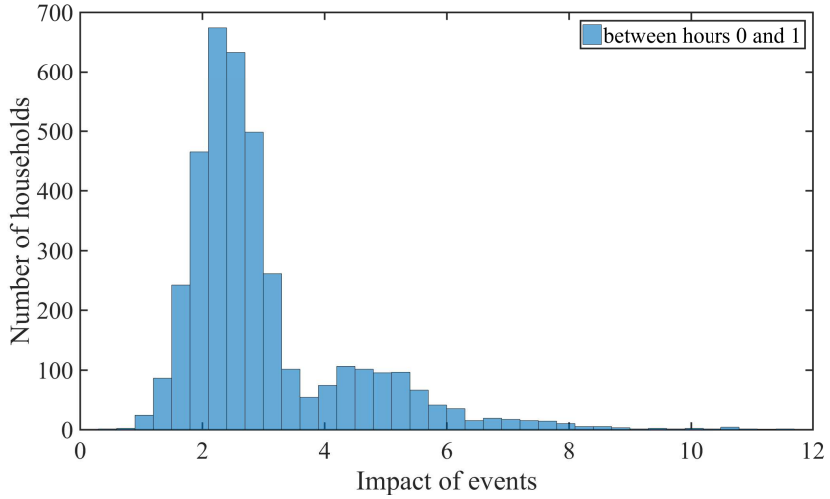


Figure 3.12: Human interaction events: distribution between 0am and 1am for 10,000 households, "zoom" to non-zero values.

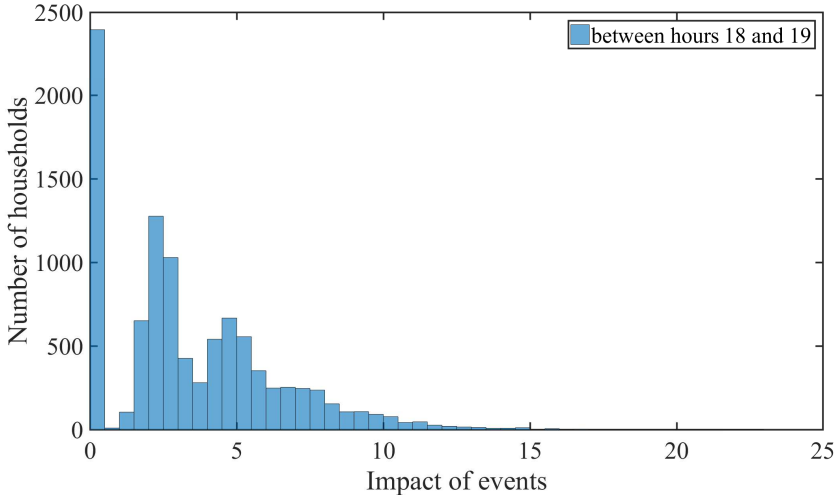


Figure 3.13: Human interaction impact: distribution between 6pm and 7pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

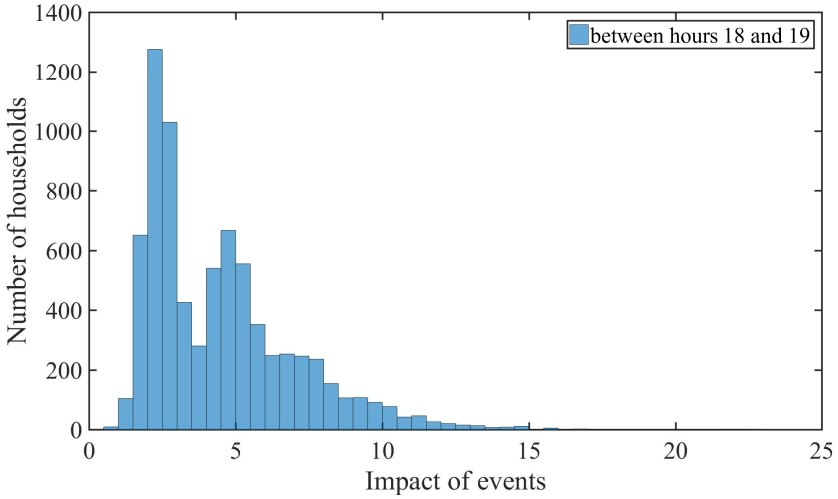


Figure 3.14: Human interaction impact: distribution between 6pm and 7pm for 10,000 households, "zoom" to non-zero values.

The rest of them can be found in the Appendix. Additionally CDF of those are given. Figure 3.15 shows the CDF between midnight and 4am, which shows how interaction overnight reduces (as expected) in the majority of households.

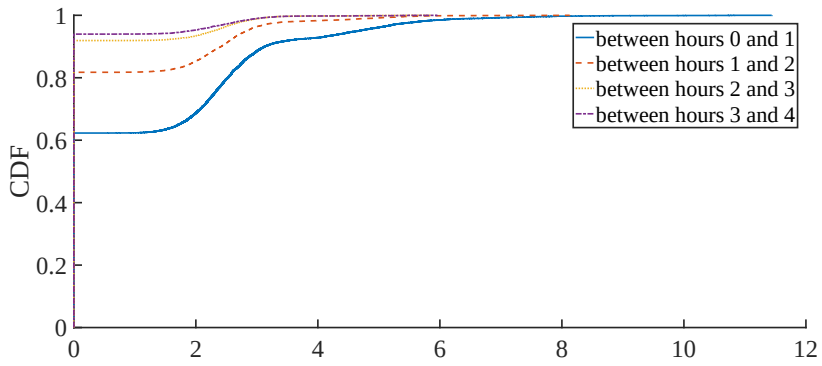


Figure 3.15: Cumulative distribution of human interaction events.

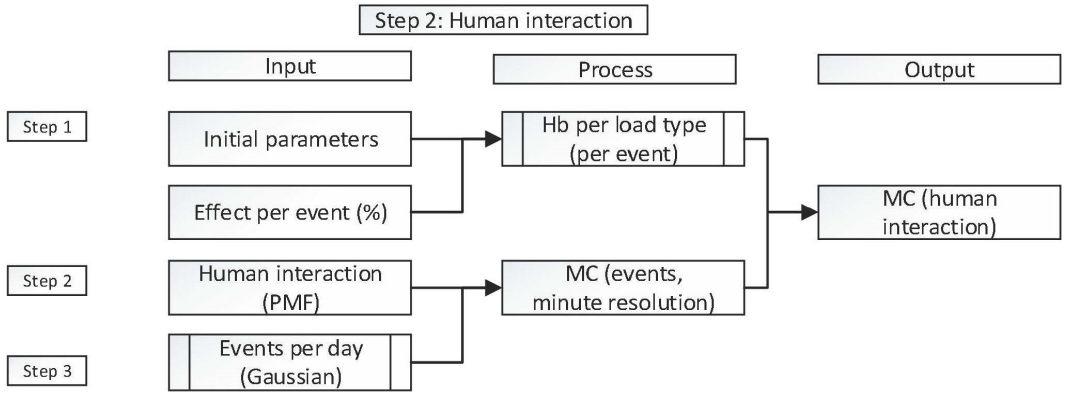


Figure 3.16: Detailed human interaction modelling, (Step 2 of flowchart 3.9) [25]

3.7.3 Step 3. Ambient (Room) Temperature

Kane T. et al [106] examined the variation of indoor temperatures and heating practices in UK dwellings and variations of up to 10 ^oC were reported. UK's heating "preferred" comfort zones and outdoors temperature were used to create a realistic MC model. The heating PMF is used to create an MC model of possible heating action, which is then compared to probable indoor temperature (based on outdoor temperature) and thus hourly temperature is defined per dwelling. This can be seen as a similar model to reality, where based on knowledge of weather conditions and historic data (population's behaviour in these conditions) the temperature can be probabilistically estimated. The resulting MC model is then converted to a time step resolution of 1 minute.

Temperature changes in dwellings due to use of heating can occur in shorter periods than an hour [86], but for this model more "mild" fluctuations were used to describe the general behaviour of the population. Simulated average household temperature and reported average temperature ([106]) can be seen in Figure 3.18 and as expected they are very close. Assuming that this report is a good representation of UK's heating practices overall, the cold load demand is

also expected relatively to be close to the actual report in [52]. The procedure Figures 3.17, 3.18 is generated is:

- Heating practices in UK according to [106]. PMF is taken from here for the following dwelling types: flats, detached, mid terrace and semi-detached/end terrace (since the last 2 were almost the same).
- Weather statistics of average winter UK temperature (since UK household data was available for February only, annual average or summer was not unfortunately [106]). If external temperature changes more than 2^oC (temperature change which humans notice) then in next hour (assuming a delay for room temperature to be affected ([106]) an MC is performed and compared to the PMF, resulting in success or failure. If the outcome was a success, then heating is assumed to be switched *on* or off accordingly, following heating preferences for dwelling type [106].
- For simulations later on (Section 3.8) the output is converted from hours to minutes (for each household).
- The average of the output is taken for each hour and compared with the average of Figure 3.2.

An important note here is that for any given simulation period, data on room temperature is needed. Thus, if someone wanted to model summer months instead, where little to no heating is needed, some hourly or sub-hourly data of temperature ranges as in Figure 3.2 would suffice or something similar. Despite the season, room temperatures in occupied dwellings will tend towards some temperature band, due to human preferences. For instance, in very hot climates/seasons, occupants are very likely to have room cooling, which will be set to comfortable temperatures and the aggregated data will tend towards that. The opposite will occur in cold climates/seasons and thus the ambient

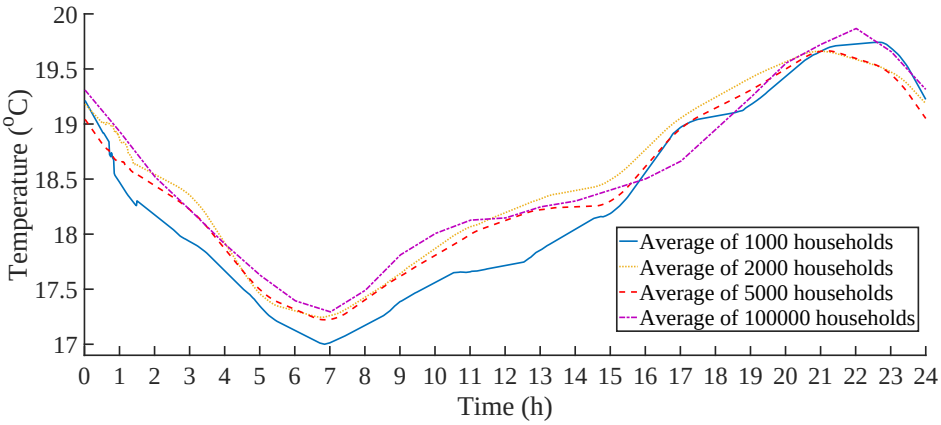


Figure 3.17: Mean indoors temperature of synthetic simulated households (1000, 2000, 5000, 100000 households).

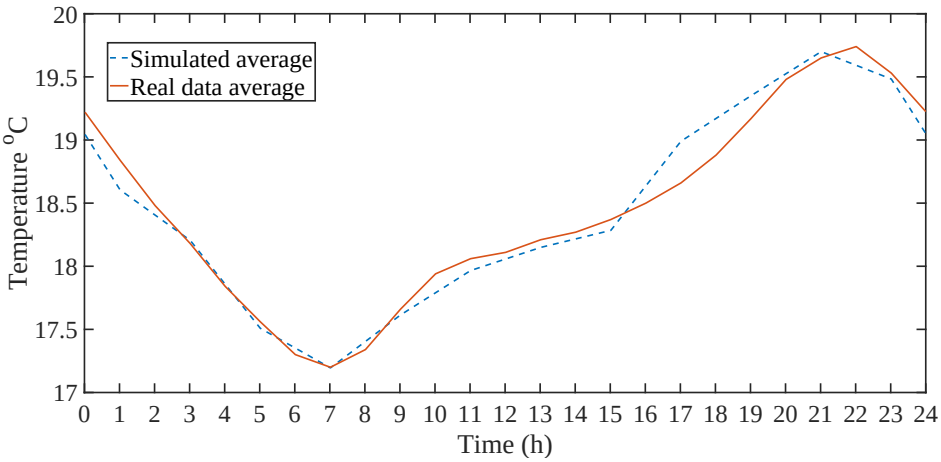


Figure 3.18: Mean indoors temperature of synthetic simulated households (10,000 households) compared to real average (Figure (3.2))

temperature of cold loads should not vary drastically. Still, higher consumption is expected during summer. Assuming that average temperatures for seasons are known, (3.10) can be used to calculate deviations/errors from the "average day of a whole annum". For the purposes of creating realistic aggregated cold load demand that follows real world data, the used data was deemed sufficient.

3.8 Simulation Results

One important aspect of a model described by (3.23) is that it can focus on the state of the TCLs specifically without including the power demand of each. Thus, it can give a clear picture of TCLs' state and how they respond to different commands as well as other factors. Demand can be added through (3.3) or (3.7) to determine the total power demand of the TCLs' population.

The model can be used in the same way as classic models (Malhamé and Chong [9]), by assuming constant θ_a and minimal external factors (Wiener process with mean 0 and small variance σ). In which case the heterogeneity of the population due to technical parameters and their effects on stochasticity (Section 3.4) can be shown through simulations. Figures 3.19-3.20 show that for 2 population sizes, 10,000 and 20,000 respectively. This is simply the "white noise" due to stochasticity, which is modelled via a Gaussian noise process in "state of the art" models similar to $\Phi(n)$ described here for discrete models and as a Wiener Process for continuous ones (e.g. [26, 28, 29, 32, 40, 69, 86, 104, 105]). In contrast, when human behaviour and ambient changes are considered and modelled (as mentioned in the previous section) the result is varying in addition to some "white noise", this can be seen in Figure 3.21. A direct comparison of the "state of the art" model (10,000 population - Figure 3.19) and the one developed (Figure 3.21) is seen in Figure 3.22.

Populations of smaller size (e.g. 1,000) will have significant elements of stochasticity and thus are not proposed for DR studies of TCLs. Unless the study is specifically aimed for small populations, then one of the most important elements to be studied is stochasticity and control actions need to take it into account. As expected the stochastic behaviour reduces as the population size increases, the values (fluctuation) are close but higher than those of Table 3.1 due to heterogeneity in parameters (different duty cycle D). Increasing the population size from 10,000

to 20,000 yields a small improvement, thus a size of 10,000 is deemed satisfactory for examining effects of human interaction and changing ambient temperature.

NB: To assume an almost uniform distribution (around 1% fluctuation or less) for heterogeneous populations, requires population sizes of the order of 100,000. Computational time would increase considerably and memory allocation might be an issue for some software.

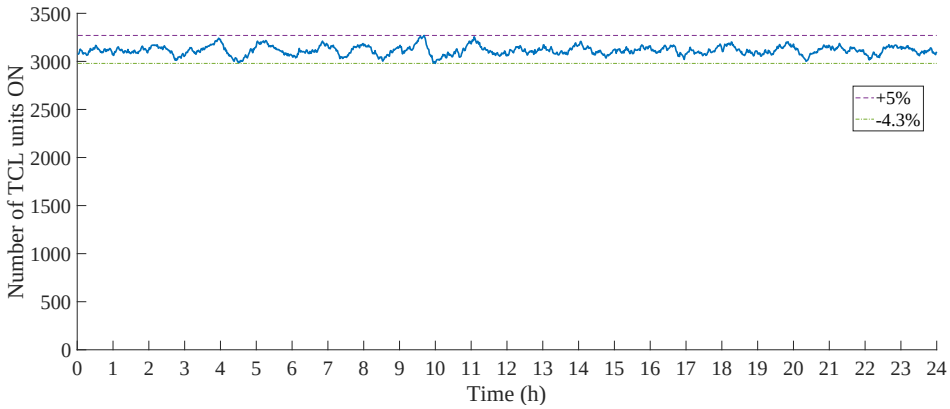


Figure 3.19: TCLs' population (10,000) state without external factors or ambient change. The level of noise due to population stochasticity and duty cycle is just above the expected range (Figure 3.4)

Arguably, one of the other important points of this model is the realistic representation of TCLs under different conditions during the day and their dynamic behaviour in time. Detailed bottom up models like this one give the highest fidelity for studying behavioural characteristics. Figure 3.21 displays simulation results for varying ambient (room) temperature and human behaviour, as described in steps 2 and 3 of the previous Section. The comparison to experimental data from DECC and Smart-A can be seen in Figure 3.23; comparison is done after converting the simulation from 1 minute resolution to hourly resolution since DECC and Smart-A have only 1hour resolution data.

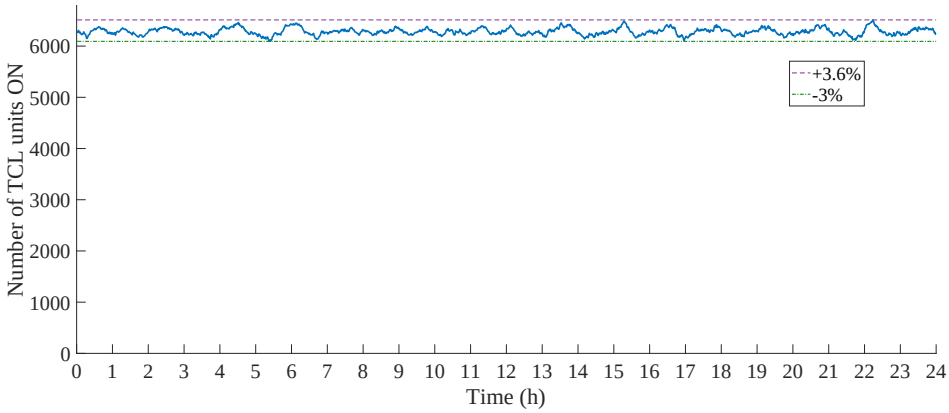


Figure 3.20: TCLs' population (20,000) state without external factors or ambient change. The level of noise due to population stochasticity and duty cycle is just above the expected range (Figure 3.5)

A direct comparison of the same data from DECC and Smart-A and the classic/"state of the art" used in most TCL studies (such as [26, 28, 29, 32, 69]) can be seen in Figure 3.24. This model is described by (3.1a) as described in the beginning of this Chapter. The result will be better with varying ambient temperature if simulation time step is small (up to a few minutes). Again in this case ambient should randomized for a heterogeneous TCL population, otherwise all TCLs' duty cycle will follow exact same trend. Yet, they should have a realistic change for the time of the day, i.e. randomization to be sampled on some distribution of real data, otherwise it will simply be "white noise" as in Figure 3.22.

Other important characteristics and phenomena that can be studied in detail with this model are the response to external commands during different conditions (periods of the day), rebound effects and oscillation damping after those, due to partial synchronization (Figures 3.26 and 3.27). The exact response of a DR

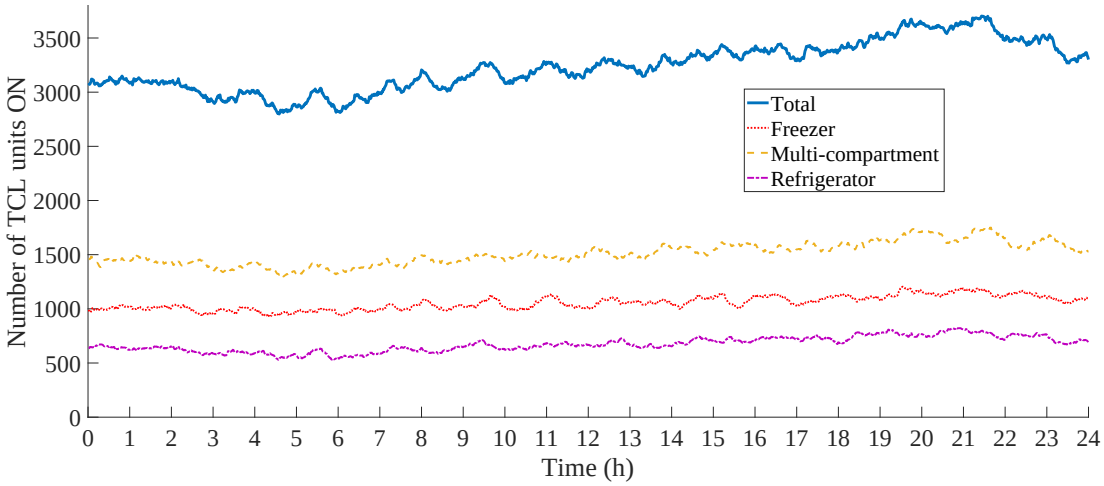


Figure 3.21: Realistic cold load demand simulation, total cold load and cold load per simulated type

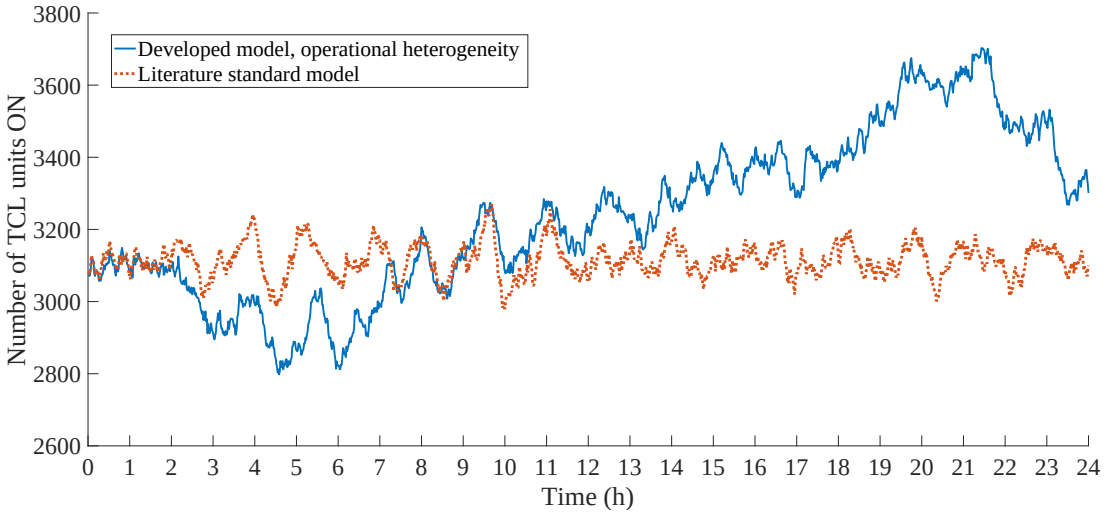


Figure 3.22: Comparison of "state of the art" model and developed model with varying heterogeneity during operation. Y axis shows number of TCL units in operation (on) out of a population of 10,000. TCL data is based on Tables 3.5 & 3.6, initialization (on or off) is randomized around each unit's duty cycle (D). The developed model follows a trend similar to Figure 3.1, as expected.

action (or other forms of interruptions such as brown-outs), are directly affected by TCLs aggregated state at that given period. In Figures 3.26 and 3.27 a short term switch off is applied (1 minutes for units in on state and forcing units in

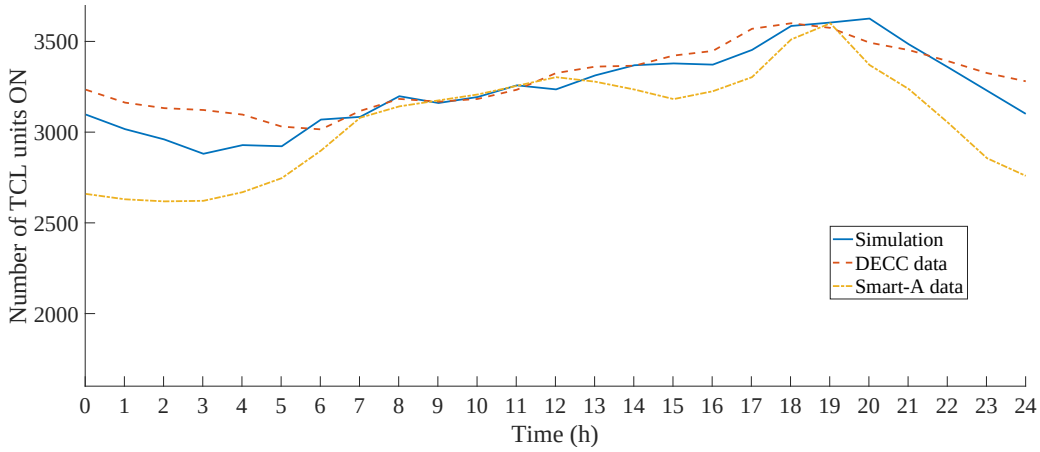


Figure 3.23: Comparison of simulated model and experimental data

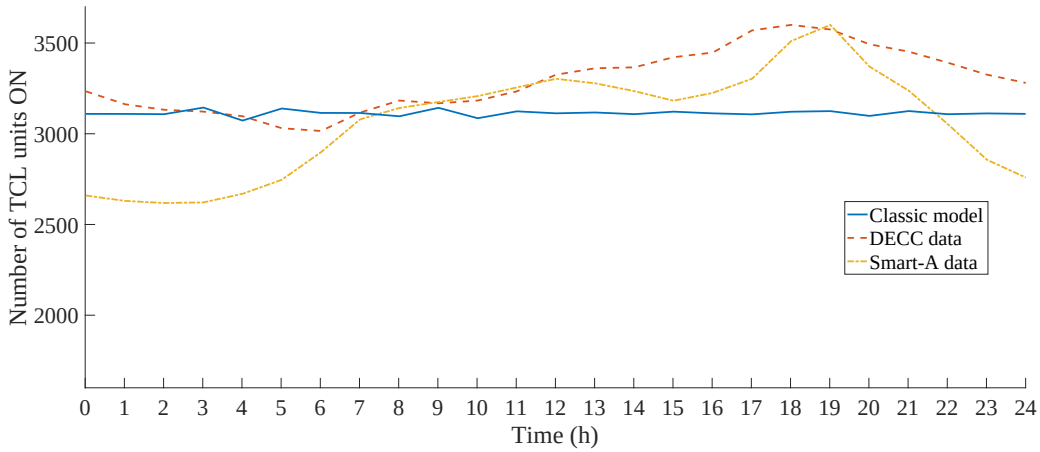


Figure 3.24: Comparison of simulated classic/"state of the art" model and experimental data

off state that would have switched on, not to), with an external command for all units. This does reduce demand for that short period as observed, but forces units in *on* state to change their normal duty cycle, whilst units in off state are not affected, causing a partial aggregated synchronization. Since part of the TCL population synchronizes, rebound effects occur, which are detrimental for DR actions.

The rebound effect and why it occurs is explained in the following Figures 3.28 - 3.31. For simplification, assume a relatively homogeneous population of 54 units

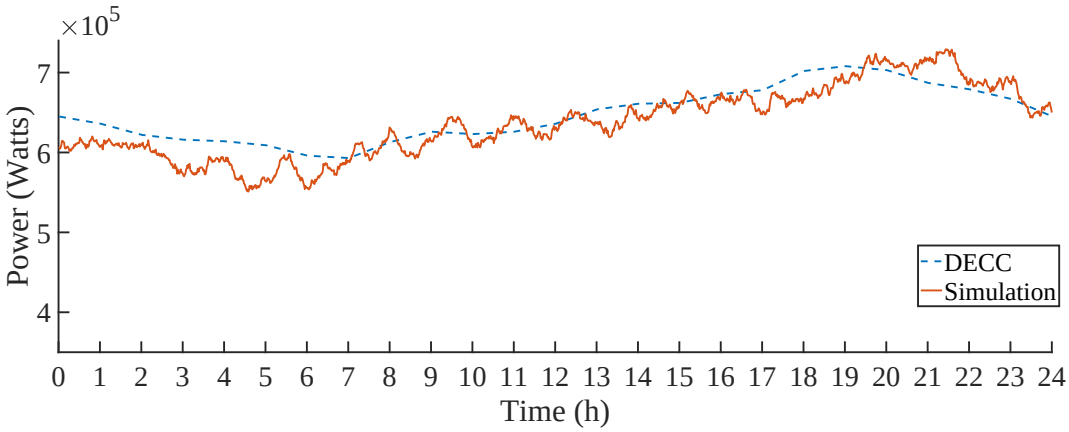


Figure 3.25: Comparison of simulated model and experimental data for 10,000 households. DECC data according to Section 2.4. Power rating of some TCLs can be seen in Table 3.5

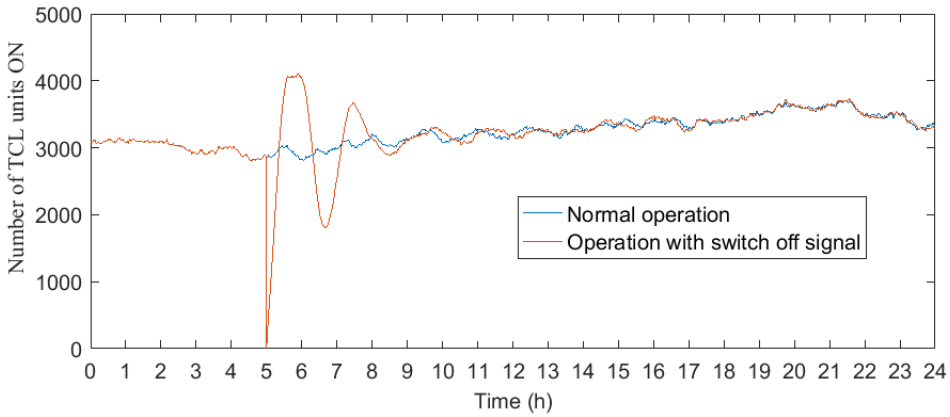


Figure 3.26: Switch off at t= 300' for 1 minute, partial synchronization of TCLs is caused as expected

(or any multiple of that), with a duty cycle D equal to 1/3, as represented in Figure 3.28. This means that, under normal operation and even distribution, 18 units are *on* (i.e. $D \cdot N_{TCL}$). For every time step, a few units will advance in their respective state, with units in the off state advancing at half the average rate compared to the units in *on* state, because of $D = 1/3$. If $D = 0.5$ then the average rates between states would be equal and half the units would be in each

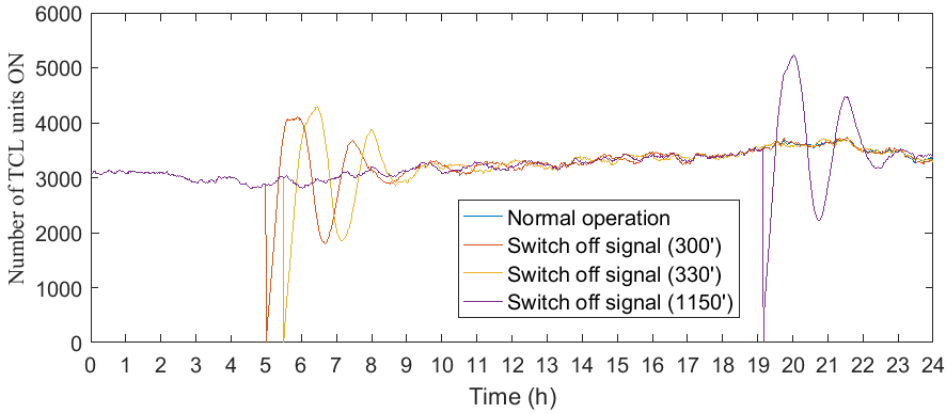


Figure 3.27: Comparison of switch off actions at different conditions (moments), 1 minute and the partial synchronizations caused

state. Note that the time step size is important in determining where units will advance (within their sub-state or next ones).

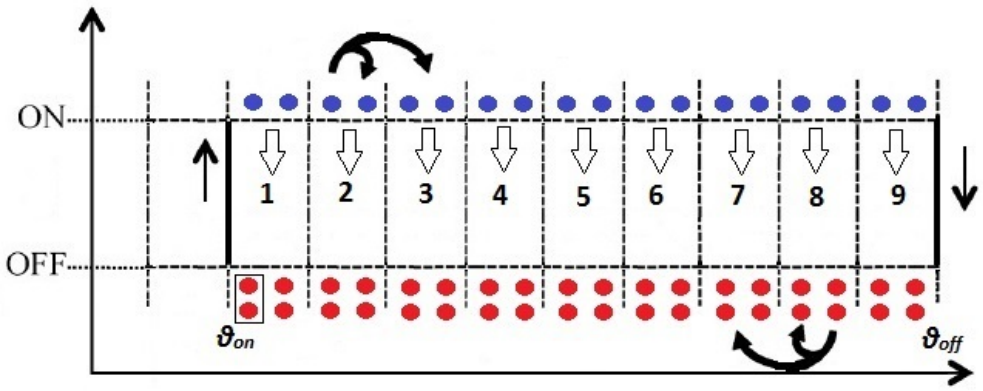


Figure 3.28: Relatively homogeneous population of 54 TCLs, evenly distributed within their temperature dead-band. The temperature dead-band has been separated in sub-temperature states for simplification. Depending on the time step, units might advance within their sub-temperature state or to next ones, as depicted by the arrows of sub-state 2 and 8.

When the units are forced to switch off, they change state, as marked by the arrows pointing downwards in the 9 sub-states of Figure 3.28 (instead of the

normal cycle), whilst retaining their temperature of course. This results in Figure 3.29, where partial-synchronization has been forced on the TCL population.

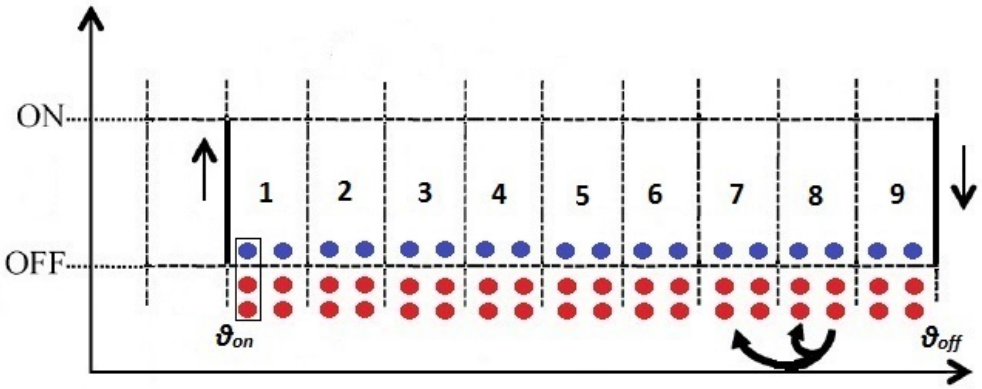


Figure 3.29: Relatively homogeneous population of 54 TCLs, when an external signal forces them to switch off. Partial synchronization occurs

In the next time step, where external control no longer applies, TCLs will resume normal operation, but now the population is under partial synchronization. Instead of 2 units switching on when θ_{on} has been reached (as marked in Figure 3.28), now 3 units will switch on, as marked in Figure 3.29. This results in Figure 3.30 in the next time step.

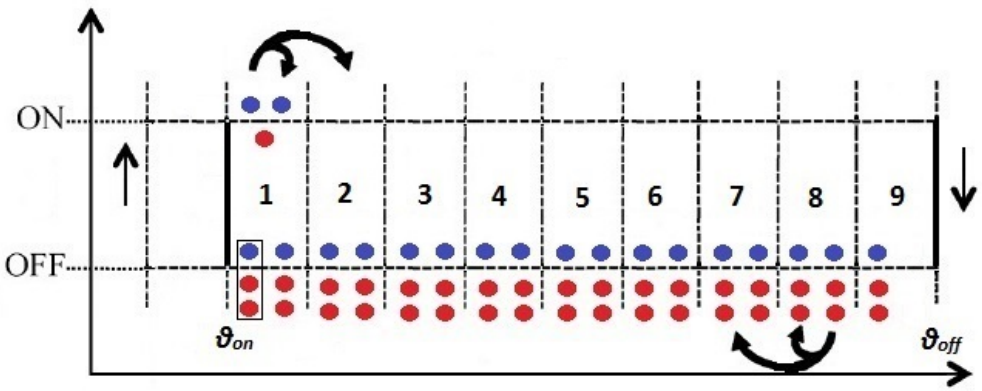


Figure 3.30: Relatively homogeneous population of 54 TCLs, step after partial synchronization due to external signal.

This process will repeat for the following steps, until TCLs reach the state described in Figure 3.31. At this point the rebound effect reaches its peak. Note that now 27 units are in *on* state compared to 18 units of the normal operation (Figure 3.28). This is an increase by 50%! In reality, due to TCL populations being heterogeneous, this partial synchronization decays over time. Figures 3.26 and 3.27 in comparison, show a rebound effect with a peak around 30-40% higher than the previous "normal" state, lower than the 50% of the homogeneous population ($D = 1/3$), due to their heterogeneity. If inrush currents were included (which are not), these peaks would be slightly higher.

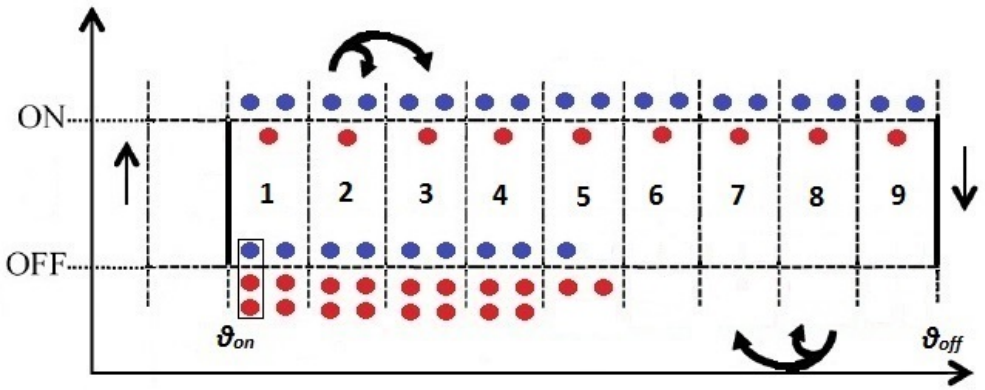


Figure 3.31: Relatively homogeneous population of 54 TCLs, peak of rebound effect is reached after a few time steps, following partial synchronization due to external signal.

These effects are significant for both Dispatchable and non-Dispatchable DR actions. For instance, when a certain amount of power is required for frequency control, TCLs can provide it by reducing demand output equal to that required amount, but after some time when the rebound effect is introduced an almost equal extra amount of energy is required, which if not countered properly by additional control actions could practically recreate the original problem [69].

The higher the current demand, the higher the rebound effect and its oscillations.

It is thus essential for control actions to take such effects into consideration and try to minimize/counter them or at least postpone them, as well as consider the natural damping occurring due to heterogeneity [28]. It is crucial to note that, in non-intrusive actions (thermal limits are maintained), the overall thermal load remains essentially the same. Consecutively, the overall electrical demand remains also the same, but spreads in time, unless the control framework is designed otherwise.

One of the starting arguments of this thesis and Chapter, but also motivation for this work, was whether small heterogeneity can be assumed for TCLs, such as cold loads, and the accuracy of models based on such assumptions. Figure 3.32 shows the variation of duty cycles during the day for 4 TCLs with relatively similar starting duty cycle (and thus operation), yet significant changes are observed during the day. This is true for the majority of the population, and the original point, that even in a (relatively) homogeneous population there will be significant heterogeneity in operation in time holds true, as well as that a Wiener process with small variance is inadequate to represent it. It is also important to note that response to control actions varies during the day (i.e. night-morning hours have relatively shorter cycles than evening), based on the state of the TCLs' population, which can also serve as a factor in identifying the available power for DR.

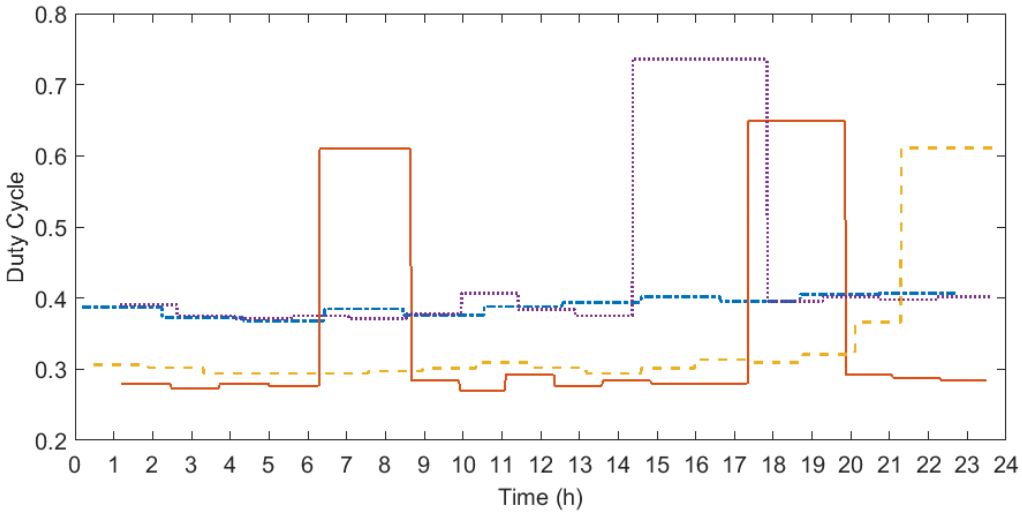


Figure 3.32: Duty cycles of 4 randomly selected TCLs during the day

3.9 Conclusions

TCLs modelling was examined and a detailed bottom up realistic model was developed for DR studies, including human interaction modelling and its effect. Controlling large amounts of loads is inevitably only possible through probabilistic models, but with proper approaches it can be accurate enough. One of the key issues for TCLs' applications is the high heterogeneity and the dynamic behaviour in time. This Chapter focused on data from experimental studies to determine the most important factors that drive TCLs' dynamic behaviour. Based on those and the developed MC model, simulations were carried out for the case of colds loads. The initial hypothesis of high heterogeneity in operation, even among relatively homogeneous loads was shown to hold true, thus the accuracy of various existing aggregation models is questionable.

The following Chapters will focus on aggregation methodologies, distributed state estimation of those and a new aggregation model. Finally, approaches

to dynamically update the developed aggregation model to follow the dynamic behaviour and duty cycle changes were described in this Chapter.

Important points and outcomes can be summarised in 6 points (with DR significance order):

1. Illustration the dynamic nature of TCLs and how their duty cycle (operation) changes during the day significantly, thus the problematic nature of using aggregation models such as CFPE by assumptions of relative heterogeneity.
2. The above occur due to external factors, essentially human behaviour, be it directly or indirectly, which were modelled and simulated using real world data to highlight their actual effects in TCLs' consumption and subsequently in TCLs clusters' state and thus DR actions.
3. A simplified equivalent thermal model of multi-compartment cold loads was developed, which converts the actual second order thermal model to a first order one, thus useable for the S-PDE models developed so far for DR without loss of TCLs characteristics.
4. The bottom up approach used, can be computationally demanding due to its detail, yet that detail allows for high flexibility and the highest possible accuracy to study DR actions, rebound effects and the performance of control algorithms.
5. Additionally, it can be used to evaluate the accuracy of more computationally efficient aggregation models and simulate smaller populations taking into consideration the magnitude of stochasticity in such cases.
6. Finally, due to the main equations including external factors, minimum and maximum limits of "normal" operation can be deduced, thus abnormal (faulty) TCL consumption can be tracked, something useful for commercial loads mostly. So far this has mainly been done through regression models;

yet it can be used in conjunction, to improve accuracy or to estimate consumption under different operational conditions for energy efficiency.

Chapter 4

State Estimation via Belief Propagation

4.1 Introduction

Various approaches are used for control algorithms and aggregation models, though the probably the most common assumption is accurate knowledge of thermal state in real time, meaning integrated costly thermal sensors, which is not feasible for individual light commercial and residential units [54]. In addition, most assume that TCL parameters are known (e.g. [26, 28, 29, 32, 40, 69, 86, 104, 105]), even in real-time . Additionally, some form of communication and processing is required, which in many cases can be feasible with minimal to no cost, via Smart Meters or Smart plugs and Apps [56, 57, 58]. As seen though in the previous Chapter, aggregated TCLs' demand change throughout the day, due various factors and includes some level of stochasticity according to heterogeneity and size. Thus, even 2 TCLs with identical parameters can have different duty cycles and thus consumption. For that reason, but also in order to

reduce monitoring (sensors) cost, one potential solution is State Estimation, such as the one suggested by Koch et al. [29] and investigated further in [40, 54, 55]. For this approach there are 2 main scenarios, no state information but knowledge of aggregated power and full state information with knowledge of aggregated power (in both scenarios TCL parameters are considered known as well). For the aggregated measurements of TCL demand, those are assumed on substation level, through disaggregation techniques [41, 42, 94], though technically possible, due to practical reasons (e.g. VPPs need approval System Operators to install equipment and perform such process) and the disaggregation accuracy are limiting factors.

Creating VPPs to provide DR services through the use of State Estimation and aggregated power measurements [54] could be technically possible. In these studies, central models have been adopted, where a lot of TCL info needs to be gathered centrally (raising privacy concerns), though as stated in [54], a decentralized approach can be more beneficial (such as for response times).

Kalman Filter (KF) has been used [29, 54] for State Estimation, which is known to be an instance of Sum-Product Algorithms [111]. One such Sum-Product Algorithm is the Belief Propagation (BP) algorithm, which is fundamentally a fully distributed algorithm, more generic and offers higher flexibility of system modelling due to Factor Graphs (FGs). This was the motivation to introduce BP in this framework as well as investigating more decentralised approaches. Essentially, FGs can be seen as generic Bayesian networks, utilizing factors nodes, which are functions. Apart from DR, BP can be applied for different purposes in power systems, such as in Cosovic's & Vukobratovic's work on distributed state estimation of power systems [112]. BP can be applied in either fully distributed, or centralised or mixed systems (i.e. hierarchical), and is easily integrated with KF or similar probabilistic frameworks.

This Chapter explores BP, for state estimation of thermal states in aggregated populations of TCLs, to provide balancing services in power systems. Thermal states can be described by hybrid-state Markov processes and hence aggregated populations behaviour by Markov Chains. If a method like the one suggested by Koch et al. [29] where to be used, one simple example of hierarchical architecture to visualize would be substations exchanging information (distributed), whilst each substation has its own aggregation and then partial measurements of loads (Figure 4.1). In which case BP can be used as a method to pass information between them.

The chapter is organized as follows: Section 4.2 explains how power state reading can be used via Smart Meters or Smart Plugs, Section 4.3 defines the transition probabilities of a TCL population and Section 4.4 the analytical calculation of the Probability Transition Matrix. In Section 4.5 the state space model is presented and KF basics. Section 4.6 details how factor graphs are formed and Section 4.7 the BP algorithm which is considered for this Chapter. Section 4.8 presents numerical results and Section 4.9 concludes.

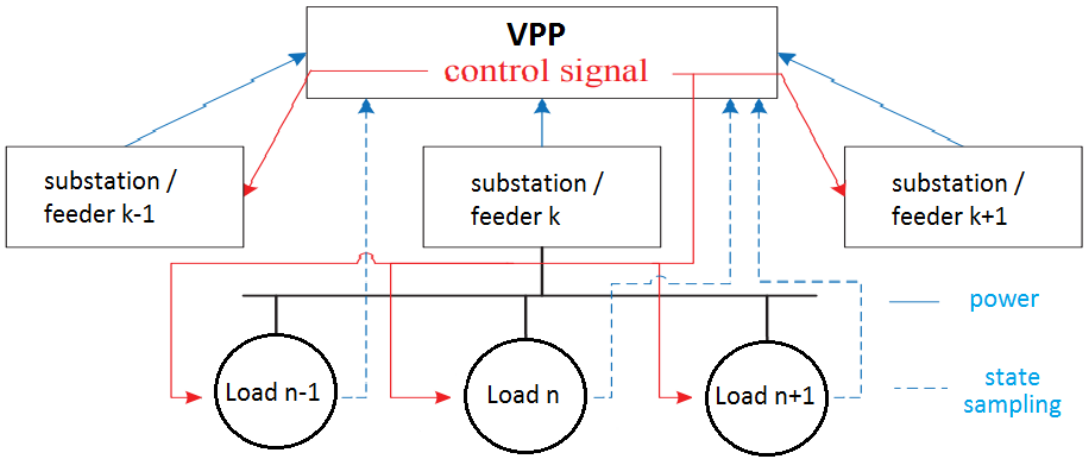


Figure 4.1: Distributed / Hierarchical architecture for VPP operation with load sampling and State Estimation

4.2 Thermal State Approximation via Power State Measurement

Literature models assume knowledge of thermal state from precise thermal sensors (i.e. [29, 54, 55]) and/or smart agents on appliance level which measures TCL parameters (e.g. λ, C, R, θ_g) in addition to accurate ambient temperature readings (i.e. [34, 39]). In some model there is the addition of devices within households which provide fast and accurate frequency readings (i.e. [36]), knowledge of the rest TCL population (in terms of available P) and knowledge of Power Systems characteristics to respond autonomously. In this thesis on of the main aims is to minimize such requirements and use primarily power readings in time (t_{on}, t_{off}, i.e. duty cycle), which here will be referred to as power state. The concept behind estimation of thermal state using duty cycle and power state (t_{on}, t_{off}) is shown in Figure 4.2. In Figure 4.2 the temperature dead-band is partitioned in n (5 for this example) equal intervals, state bins, for each state (on, off). If a very precise thermal sensor was installed in this TCL, the readings at different times (e.g. 45, 72, 119) would show the exact temperature (marked by X in Figure 4.2) and thus the bin for each of those times can be identified, assuming θ_{off}, θ_{on} are also known. If the time within each state is used instead and compared to t_{on}, t_{off} for this duty cycle, then the bin for each time can be approximated (marked by □ in Figure 4.2), with a small error.

Note that in this case relative position within the deadband is given, thus actual knowledge/readings of θ_{off}, θ_{on} are not required, e.g. a unit shows power on for 37minutes (out of 45minutes t_{on}), thus likely to be in bin 80%-100% on, i.e. bin 5 (on). On the other hand, compared to precise thermal sensors, error of actual state bin is higher (e.g. at time 72). Interestingly, in case of multi-compartment TCLs, thermal sensors will give different readings for each compartment; even

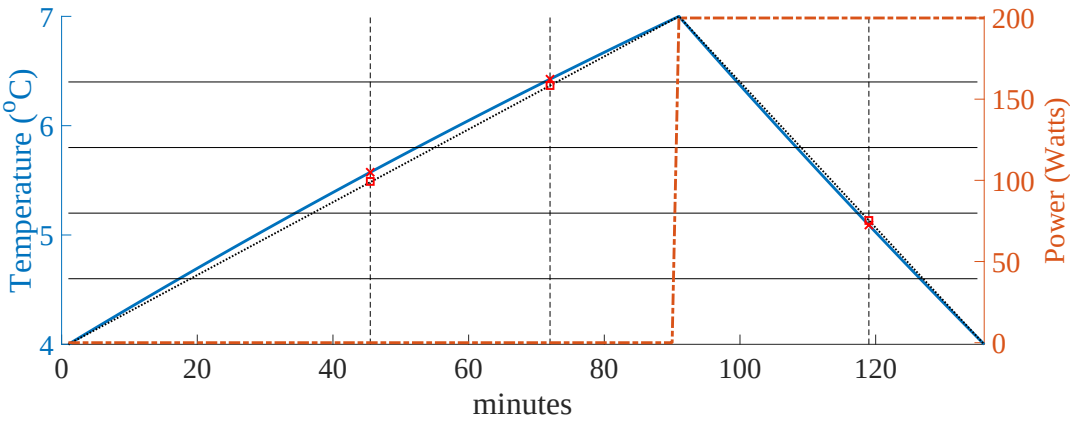


Figure 4.2: TCL cycle using real characteristics of cold load of Figure 3.10. The temperature deadband is partitioned in 10 state bins, 5 for on and for off state. Solid line shows the thermal state over time, whilst power demand is shown by bold dashed line.

though they tend to synchronize [109, 110], readings can give different state bins, one in off state, another in on, whilst TCL is either on or off, resulting in wrong input for the state space model. For which case, power readings at unit level are needed to avoid instances where commands to switch on is given to units whilst being already on or vice versa. Using power state instead, means directly looking at demand and "relative" thermal bin for the unit as a whole ("an equivalent"). Either way, using transition probability from previous states and state estimation is important in improving bin state estimations. The key advantage of using power states is how inexpensive this approach is compared to the cost of precise thermal sensors for **each** TCL. Also, with Smart Home Systems which are now supported by Smart Meters or Smart Plugs in combination with smart agents (also known as virtual assistants, well known ones have been developed by Amazon, Google, Apple), it is very simple to implement. Many households already use either, meaning no extra equipment costs for those.

Figure 4.3 shows 2 different TCLs (different parameters R, C, θ_g), but with the same duty cycle due to some external effect (similar to the real cold load of Figure

3.10). For both of these TCLs, approximating the thermal state using power state still has a small error. Another point seen in both Figures 4.2 and 4.3 is that for a small time step relative to a (e.g. 1 minute), the transition probabilities between bins is practically limited. For instance, a TCL with a normal cycle of 90minutes (30minutes on, 60 minutes off, see as example Figure 4.2) cannot realistically move from the beginning of one state to the middle of the other in 1 minute ($\tau = 1min$). Even when H is considered, for 1minute that TCL can practically move to the next bin, the previous (Figure 4.3) or most likely stay in its current one. Respectively, if a time step τ of 25minutes is assumed for the same TCL (Figure 4.2), a transition to the next bin or the one after are likely, but not staying in the same bin. When H is considered (Figure 4.3) though, there is one more possible bin to transition to. Small time step are preferable obviously and also required Balancing Services, thus the error minimal.

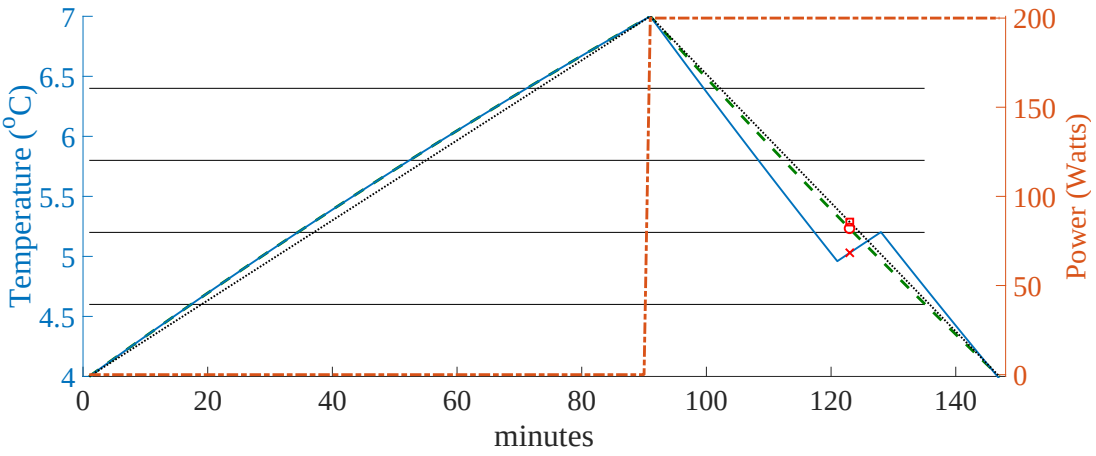


Figure 4.3: Two TCL cycles with different characteristics, yet same duty cycle. The temperature deadband is partitioned in 10 state bins, 5 for on and for off state. The first TCL (solid line) has the same parameters as in Figure 4.2 but with an external effect, whilst the dashed one without. The errors of estimating state bin through power state is higher than before but still small.

A few key points are:

1. As seen this method gives an approximation of the state bin, albeit very close, state estimation will be important in reducing errors. Given previous state and transition probability, state estimation is beneficial. Note that even when precise thermal sensors are considered, such as in [29, 40, 54, 55], state estimation is still deemed crucial.

2. The duty cycle (t_{on}, t_{off}) as stated in Chapter 3 is important since consumption directly depends on it. Not only that, but with this approach relative thermal state (indirectly) is tracked by power state and t_{on}, t_{off}. This is already done by devices and Apps which are becoming increasingly common [56, 57, 58]. Additionally, clustering based on t_{on}, t_{off} is more generic than based on TCL parameters (e.g. R, C, θ_g). In general clustering has shown improvement in accuracy of modelling aggregated populations of TCLs [113].

3. As someone can easily notice from Figure 4.3, there are times when previous state and a state space model will give an inaccurate estimation, be it either when measuring temperature or the proposed method. On an aggregated population though, this error will be lower, as TCLs are not synchronized and disturbances are randomly distributed. This means that in some of them the error will be towards one direction (in regards to state bins) and in some others towards the other one.

4. Units with different parameters, can have the same t_{on}, t_{off} (thus duty cycle). Clustering based on t_{on}, t_{off} (i.e. duty cycle) as proposed in Chapter 3 is beneficial. Some level heterogeneity is still sought after for control actions to dump partial-synchronization, as discussed in Chapter 3.

4.3 Trans. Probabilities for TCLs, Including External Factors

For the purposes of this chapter, BP is directly compared to KF for state estimation. As such the same state space model is used as in KF studies [29, 54]. Additionally, a similar method for calculating analytically the PTM is explored, though transition probabilities within each state are not equal and there is the addition of external factors. For comparison purposes, external factors are set to 0 during simulations, since they were not considered in the aforementioned studies. Three crucial points, that limit the potential of this particular aggregation model, as will be made evident later on, are:

1. Limited heterogeneity modelling (only parameter $e^{\tau/R\cdot C}$)
2. Impractical to calculate external factors in reality
3. Complex (if not infeasible) to perform analytical calculations for more parameters

Therefore, a new aggregation method for heterogeneous TCL populations is developed in the following chapter, as well as the corresponding state space model and FG. Yet for this chapter, a plain field between BP and KF was deemed necessary.

TCL models in this Chapter adhere to (but are not limited to) HVAC, cold loads, electric space and water heating, heat pumps, chiller etc. They are found in residential, commercial and industrial consumers. TCLs control is based on a hysteresis scheme within a deadband δ, where they change between *on* and *off* states as described in the previous chapter. During *on* state a TCL operates until it reaches a specific temperature θ_{off}, at which point it switches off (thermostat

command). After entering the *off* state, the temperature advances towards a specific θ_{on} temperature, where it switches on again and repeats the cycle.

This is the most common operation found in domestic, commercial and industrial thermal loads and gives them thermal "storage" capabilities; rendering them perfect candidates for providing short term balancing services as mentioned in previous chapters. Loads whose operation is characterized by a cycle and exhibit some form of inertia (or hysteresis) can be represented by a cycle of transition probabilities between *on* and off states (and sub-states of those). Aggregated TCLs, given the proper communication infrastructure, can operate as part of virtual power plants [114, 115]. The equations derived in Chapter 3 to represent the operation of a TCL will be used to calculate transition probabilities for a TCL and consequently for a TCL population.

Individual TCL Model:

The previous Chapter focused on accurate bottom-up MC models, with the aim to capture the dynamic changes in duty cycles. Yet, as mentioned, for coordinated control of large populations, aggregation models are required. In this and the following Chapter the focus will be on aggregation model, keeping in mind that they should not be static. The MC model developed in the previous Chapter can be used to validate aggregation models developed hereon, by simulating both and comparing results (aggregation models should ideally match the MC one with a small noise).

Equations (3.23) and (3.24) derived in Chapter 3 describe the (discrete-time) evolution of temperature of a TCL unit, where ZOH has been applied. When small time steps are assumed, such as those needed for Balancing Services, the term $\gamma\frac{a-1+\lambda\tau}{\lambda}$ of (3.23) tends to 0 and given that $\theta_a(n)$ is updated in each time step it can be excluded, thus (3.23) or (3.24) can be used indiscriminately.

TCL approximation for transition probabilities:

As observed by (3.24) and also Figures 3.10 and 4.3 (the reader is highly encouraged to see Figure 3.10 which is based on real world data), the effect of H can cause a TCL to move "backwards" in terms of thermal state when in *on* state or advance faster when in off state. Especially in the case of "on" state, this behaviour means that in reality, for a single TCL, there is the possibility of going to a previous state when in *on* state. Albeit this behaviour was captured in Chapter 3 and the Monte Carlo simulations of 10,000 units, in reality it is not practical to track and include in the same detail to calculate transition probabilities of large aggregated TCL populations; unless very precise thermal sensors are used to monitor the exact thermal state which is against the aims of this thesis.

Observing Figure 4.3, we can see that in such a case, the TCL can be approximated, with some error, with another TCL of equal t_{on}, t_{off} (thus Duty Cycle), resulting in the same consumption in time. For this particular example, if the TCL of equal t_{on}, t_{off} (dotted line) was used to approximate the thermal state, then before time $t = 126$ the real thermal state is lower (for *on* state), whilst afterwards the opposite. For most of the time, the real thermal state is lower in this case. If that same event had occurred around time $t = 100$ then for most part (of *on* state) the real thermal state would be higher. Given that within a TCL population the state of one unit is independent of another, then the distribution of these events can be considered random in terms of where they occur within each state (*on* or off). Based on that, we will use the TCL of equal t_{on}, t_{off} (see Figure 4.3 dotted) to calculate transition probabilities and that of course will introduce some error. Note that not all TCLs in a population will have such an event for every of *on* or off state.

The second TCL of Figure 4.3 is thus given by

$$\theta(n+1) = a_H \cdot \theta(n) + [1 - a_H][\mu(n) \cdot \theta_g + \theta_a(n)] \tag{4.1}$$

where $a_H = e^{-\lambda_H \tau}$ can change value in each state to account for H and is calculated using t'_{on} of (3.46) and t'_{off} of (3.48) as per below:

$$\lambda_H = -\frac{1}{t'_{off}} \cdot ln(\frac{\theta_{on} - \theta_a}{\theta_{off} - \theta_a}) \tag{4.2}$$

$$\lambda_H = -\frac{1}{t'_{on}} \cdot ln(\frac{\theta_{off} - (\theta_g + \theta_a)}{\theta_{on} - (\theta_g + \theta_a)}) \tag{4.3}$$

Note that when $\bar{H} = 0$ then $\lambda_H = \lambda$, for that state thus $a_H = e^{-\lambda\tau} = e^{-\frac{\tau}{R \cdot C}} = a$, otherwise a_H and a are different. Specifically, a_H now changes values in each state to account for the increased duty cycle and consumption, as per Figure 4.3 (dotted line).

Note that (4.1) is used **only for the transition probabilities and the resulting transition matrix** of a large TCL population as shown later in this section and not for the Monte Carlo simulations of this Chapter or others, which still use the model of Chapter 3. The aim is to create a transition matrix (via a way that does not require precise thermal sensors) that indirectly accounts for increased consumption due to H, to use together with State Estimation to estimate the state of the TCL population. That TCL population is modelled via mentioned Monte Carlo according to Chapter 3. In that sense, for a fraction of the TCL population (4.1) will overestimate the state and for another it will underestimate it (as shown in Figure 4.3), where random distribution of the events is assumed. It is hard to argue if the overall error will be biased towards either direction, unless extensive experimental data is used (but unfortunately was not available). Of course (4.1) is not equal to (3.24), and the above method introduces some error, albeit does not require precise thermal sensors.

TCL population: Observing and tracking a large collection of individual TCL models is not scalable. In [40], an aggregation method is proposed for a large collection of heterogeneous individual TCL models. The advantage of such a model is better tracking of the TCL population behaviour to external controls through system analysis and thermal state estimation, which proves to be more effective than model-free control algorithms [54, 55]. State estimation (on / off) is crucial for real time control as well as determining the ability of TCL populations to serve as VPPs and provide balancing services (in terms of both power and duration). In the absence of external signals, the dynamics of the discrete-valued state $\mu(t)$ are given by the standard thermostat operation (switch at θ_{on}, θ_{off}). The change from one state to another (or remaining within the same state) can be described by a Markov chain (Figure 4.4). This can be extended further by partitioning on, off states, such an example can be seen in Figure 4.5.

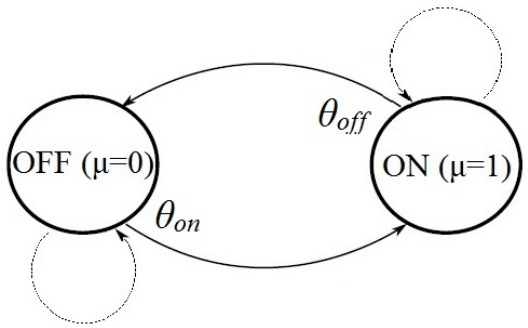


Figure 4.4: TCL dynamics described by a Markov chain.

Using that process, the operation of an aggregated TCL population can be described by the graph presented in Fig. 4.5, where each bin state has a fraction of the TCL population, randomly distributed.

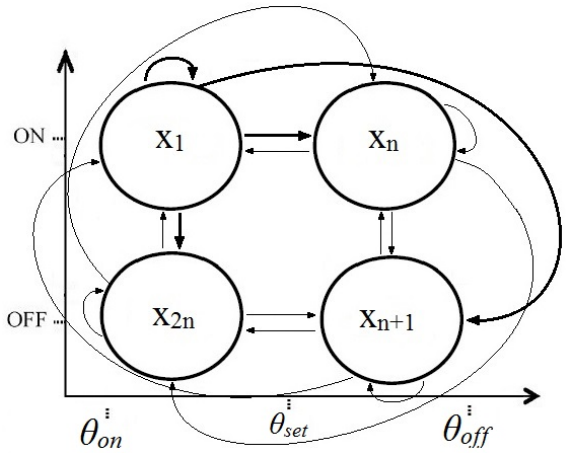


Figure 4.5: TCL aggregated population transition between 2n bin states, where n here equal to 2. Same process as above can be applied for n higher than 2. TCL state transition representation between $2n$ bin states (off and on)

The temperature deadband is divided in n equal temperature bins. It contains two sets of states: the set of *off* states $\mathcal{X}_{\text{off}} = \{x_1, x_2, \ldots, x_n\}$, and the set of *on* states $\mathcal{X}_{\text{on}} = \{x_{n+1}, x_{n+2}, \ldots, x_{2n}\}$. The states in both sets are defined as follows; for a given i-th bin, $1 \leq i \leq 2n$, the corresponding state x_i represents a random variable that describes the fraction of the TCL population that is currently in the i-th bin (temperature interval). The total number of states is $N = 2n$, as shown in Fig. 4.5. This way by separating each state in n bins (for each *on*/*off* state), the above can be mapped to probabilities and thus a TPM.

4.4 Transition Matrix Analytical Solution

Transition probabilities from one state bin to another can be calculated, albeit with limitations, based on a (consecutively a_H). A concept originally introduced by Koch et al. [29] and later used by others such as Mathieu et al. [40], Vrettos et al. [55]. Though the transition probabilities in [29] are not solved analytically but

numerically via simulations. The main basis to create probabilities will be used in Chapter 4 (but not in Chapter 5), but with a few distinct differences, such as that the transition probabilities within each state to **not be equal and constant** (for instance see [40]), with the addition of accounting indirectly for external factors as shown previously and the distribution of starting temperature (θ_{start}). Also, the probabilities here are analytically solved and not via simulations. This is highlighted as in the author's opinion the transition probabilities should **not be equal within each state** as the temperature evolution of TCLs is not linear (but exponential).

Assume that θ_g, θ_a are the same for all TCLs of the population and that a_H^i of (4.1) (i denoting the ith TCL within the population) is stochastic in a such a way that all a_H^i are uniformly distributed $a_H \sim U(a_{min}, a_{max})$ for given a_{min} and a_{max}. Note that a_H (as well as a) is a function of the time step τ and thus the factor $(1 - a_H)$ in front of the term $(\mu(n) \cdot \theta_g + \theta_a(n))$ defines the rate of temperature evolution towards θ_{off} or θ_{on} for that time step depending on state. As the time step τ increases, $(1-a_H)$ increases as well. The factor a_H in front of the term $\theta(n)$ has the opposite effect. These assumptions introduce two important limitations at this point; first the distribution of a_H and secondly the assumption of θ_g, θ_a being the same for all TCLs. Both introduce some error, maybe for θ_g, θ_a average values of the population could be used instead, which still would have some error.

By choosing a relatively small time step τ, the temperature evolution for each step is practically restricted within the starting bin or the next one (as shown earlier in Figure 4.2 and Figure 4.3 for the equivalent TCL), which is also affected by the size of the state bins.

A unit's advance from a temperature point to another one depends on a_H (4.1) as shown in Figure 4.6. Two units with $a_1 < a_2$, will go to θ_1, θ_2 accordingly.

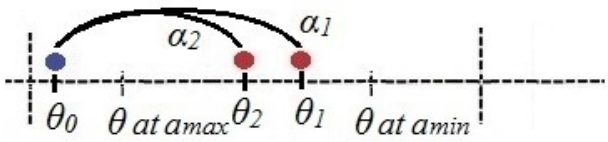


Figure 4.6: TCLs advancing to another temperature, based on a_H.

These a_1, a_2 are given by:

$$a_1 = \frac{\theta_1 - \theta_a(n) - \mu(n) \cdot \theta_g}{\theta_0 - \theta_a(n) - \mu(n) \cdot \theta_g} \tag{4.4a}$$

$$a_2 = \frac{\theta_2 - \theta_a(n) - \mu(n) \cdot \theta_g}{\theta_0 - \theta_a(n) - \mu(n) \cdot \theta_g} \tag{4.4b}$$

As seen by Figure 4.6, a_{min} defines the maximum temperature jump in one time step (which depends on τ) and the minimum bin size for which the assumption of restricting transition to current and next bin holds.

The probability of a TCL going from θ_0 to a temperature between θ_2 and θ_1 can be given by:

$$Pr(a_1 < a_H < a_2) = \int_{a_1}^{a_2} p(a_H)\, \mathrm{d}a_H \tag{4.5a}$$

$$p(a_H) = \begin{cases} \frac{1}{a_{max} - a_{min}} & for \quad a_H \in (a_{min}, a_{max}) \\ 0 & otherwise \end{cases} \tag{4.5b}$$

Now consider state bins a group of TCLs, that are either all *on* or off ($\mu^i(n) = \mu(n)$). The probability (Pr) of going from a $\theta_s \in (\theta_{sj}, \theta_{ej})$ to a $\theta_{end} \in (\theta_{si}, \theta_{ei})$ in one time step can be given by:

$$Pr = \int_{\theta_{sj}}^{\theta_{ej}} \frac{1}{\theta_{ej} - \theta_{sj}} \int_{a_1}^{a_2} p(a_H) da_H d\theta_s \tag{4.6a}$$

$$p(a_H) = \begin{cases} \frac{1}{a_{max} - a_{min}} & for \qquad a_H \in (a_{min}, a_{max}) \\ 0 & otherwise \end{cases} \tag{4.6b}$$

where a_1, a_2 depend on θ_s as per (4.4a), (4.4b). Also note that in (4.6a) there is the assumption that θ_s^i are uniformly distributed within the state bin, which looking at Figures 4.2 and 4.3 (dotted) it is not a big error, but again here extensive experimental data should be used and the appropriate distribution should be used in real life applications with (4.6a) being adjusted accordingly.

Now, using the above the aim is to derive the TPM-A, thus find the solution of 4.6a. Solving for the *off* state, there are 5 cases regarding the given starting state bin and possible new state bin in the next time step, where $\theta_1 = \theta_{si}$, $\theta_2 = \theta_{ei}$ and $\mu = 1$ if the TCL is in *on* state, $\theta_1 = \theta_{ei}$, $\theta_2 = \theta_{si}$ and $\mu = 0$ if the TCL is in off state. Since the state bin size is larger than the maximum temperate jump (the one for a_{min}) due to small simulation step, which means $a_{max} - a_{min} < a_2 - a_1$. This can be visualized in both Figures 4.6 and 4.7, but also in Figures 4.2, 4.3. If a larger step is chosen, then there will be transition probabilities in more bins. For models designed for Balancing Services, small steps are expected to be used, even smaller than the one used here.

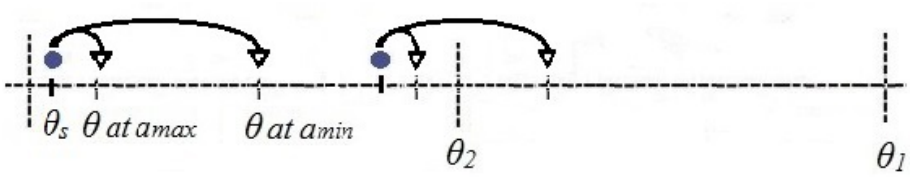


Figure 4.7: TCLs advancing within state bins, based on a. Note that the more the TCLs advance in each state the slower the rate they do so. Cases 1 (leftmost) and 2 (rightmost) are shown

$$Pr(a_1 < a_H < a_2) = \begin{cases} 0, & \text{if} \quad a_2 < a_{min} \\ \frac{a_2 - a_{min}}{a_{max} - a_{min}}, & \text{if} \quad a_{min} < a_2 < a_{max} \ (a_1 < a_{min} \text{ implied}), \\ 1, & \text{if} \quad a_{max} < a_2 \qquad (a_1 < a_{min} \text{ implied}), \\ \frac{a_{max} - a_1}{a_{max} - a_{min}}, & \text{if} \quad a_{min} < a_1 < a_{max} \ (a_{max} < a_2 \text{ implied}), \\ 0, & \text{if} \quad a_{max} < a_1 \end{cases} \tag{4.7}$$

Note that α_1, α_2 **are functions of** θ_s (4.4a, 4.4b), as such, we can calculate the above inequalities for θ_s. In that way, we can see the possible positions of θ_s for the above cases.

Case 1, *off* state, $a_2 < a_{min}$

$$\begin{gathered} a_2 < a_{min} \Rightarrow \frac{\theta_2 - \theta_a - \mu \cdot \theta_g}{\theta_s - \theta_a - \mu \cdot \theta_g} < a_{min} \Leftrightarrow \\ \frac{\theta_2 - \theta_a - \mu \cdot \theta_g - a_{min} \cdot (\theta_s - \theta_a - \mu \cdot \theta_g)}{\theta_s - \theta_a - \mu \cdot \theta_g} < 0 \end{gathered} \tag{4.8}$$

For the OFF state $\theta_s - \theta_a - \mu \cdot \theta_g < 0$ (because $\mu = 0$ and $\theta_s < \theta_a$ always for

cooling/cold loads). Equation 4.8 during the *off* state becomes:

$$\theta_2 - \theta_a > a_{min} \cdot (\theta_s - \theta_a) \Leftrightarrow \theta_s < \frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}} \tag{4.9}$$

The rest cases for the *on* state are calculated similarly to 4.8 and 4.9.

Off state θ_s limits for each case:

$$\theta_s < \frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}} \tag{4.10a}$$

$$\frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}} < \theta_s < \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}} \tag{4.10b}$$

$$\frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}} < \theta_s < \frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}} \tag{4.10c}$$

$$\frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}} < \theta_s < \frac{\theta_1 + (a_{max} - 1) \cdot \theta_a}{a_{max}} \tag{4.10d}$$

$$\frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}} < \theta_s \tag{4.10e}$$

For the *on* state, $\theta_s - \theta_a - \mu \cdot \theta_g > 0$ (because $\mu = 1$ and $\theta_s - \theta_g > \theta_a$ always). Thus 4.8 during the *on* state becomes:

$$\theta_s > \frac{\theta_2 + (a_{min} - 1) \cdot (\theta_a + \mu \cdot \theta_g)}{a_{min}} \tag{4.11}$$

The rest cases for the *on* state are calculated similarly to 4.8 and 4.11. *On* state:

$$\theta_s > \frac{\theta_2 + (a_{min} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{min}} \tag{4.12a}$$

$$\frac{\theta_2 + (a_{min} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{min}} > \theta_s > \frac{\theta_2 + (a_{max} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{max}} \tag{4.12b}$$

$$\frac{\theta_2 + (a_{max} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{max}} > \theta_s > \frac{\theta_1 + (a_{min} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{min}} \tag{4.12c}$$

$$\frac{\theta_1 + (a_{min} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{min}} > \theta_s > \frac{\theta_1 + (a_{max} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{max}} \tag{4.12d}$$

$$\frac{\theta_1 + (a_{min} - 1)(\theta_a - \mu \cdot \theta_g)}{a_{min}} > \theta_s \tag{4.12e}$$

The solution of (4.6a) from state bin I $(\theta_{sj}, \theta_{ej})$ to state bin Ξ $(\theta_{si}, \theta_{ei})$, in *off* state can now be calculated. Note, that the above (4.10) and (4.12) show the values θ_s can take for each case.

Case 1, *off* state, $a_2 < a_{min} \Rightarrow \theta_s < \frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}}$

The relation $a_2 < a_{min}$ means obviously that the probability is zero, as a cannot take a value out of limits. This in practical terms means that for the given time step (τ), a cannot take such a small value and transition to a bin so far from the starting bin, as shown in Figures 4.7 and 4.2. A small a_H results to large steps, but that requires a relatively large time step as well.

$Pr(a_1 < a_H < a_2) = 0$, thus

$$Pr^1_{I,\Xi} = 0 \tag{4.13}$$

Case 2, *off* state, $a_{min} < a_2 < a_{max} \Rightarrow \frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}} < \theta_s < \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}}$

The solution of the first integral is:

$$\frac{a_2}{a_{max} - a_{min}} - \frac{a_{min}}{a_{max} - a_{min}}$$

Supplementing a_2 from 4.4b (θ_s the variable of the second integral):

$$\frac{1}{\theta_{ej} - \theta_{sj}} \int_{\theta_{\mathrm{sj}}}^{\theta_{\mathrm{ej}}} \frac{\theta_2 - \theta_a}{(\theta_s - \theta_a)(a_{max} - a_{min})} - \frac{a_{min}}{a_{max} - a_{min}} \mathrm{d}\theta_s$$

Remember that $\frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}} < \theta_s < \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}}$, hence if $\theta_{sj} < \frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}}$, the above integral from θ_{sj} to $\frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}}$ is zero. Similarly if $\theta_{ej} > \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}}$.

Denoting $\theta_{s1} = max(\theta_{sj}, \frac{\theta_2 + (a_{min} - 1) \cdot \theta_a}{a_{min}})$ and $\theta_{s2} = min(\theta_{ej}, \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}})$, the

solution of the above is:

$$Pr^2_{I,\Xi} = \left[\frac{(\theta_2 - \theta_a) \cdot \ln(\theta_s - \theta_a) - \theta_s \cdot a_{min}}{(\theta_{ej} - \theta_{sj})(a_{max} - a_{min})} \right]_{\theta_{s1}}^{\theta_{s2}} \tag{4.14}$$

An illustration of the probability calculated in (4.14) (moving to next state bin between θ_2 and θ_1) can be seen in Figure 4.7 (rightmost example). The probability of staying in the same state bin for the same example is calculated in case 4 below.

Case 3, *off* state, $a_{max} < a_2, a_1 < a_{min} \Rightarrow \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}} < \theta_s < \frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}}$

Denoting $\theta_{s1} = max(\theta_{sj}, \frac{\theta_2 + (a_{max} - 1) \cdot \theta_a}{a_{max}})$ and $\theta_{s2} = min(\theta_{ej}, \frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}})$, the probability is:

$$Pr^3_{I,\Xi} = \left[\frac{\theta_s}{\theta_{ej} - \theta_{sj}} \right]_{\theta_{s1}}^{\theta_{s2}} \tag{4.15}$$

An illustration of the probability calculated in (4.15) (staying in the same bin, before the illustrated θ_2) can be seen in Figure 4.7 (leftmost example) when θ_2, θ_1 are chosen for the starting bin instead (e.g. the illustrated θ_2 is now the θ_1).

Case 4, *off* state, $a_{min} < a_1 < a_{max} \Rightarrow \frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}} < \theta_s < \frac{\theta_1 + (a_{max} - 1) \cdot \theta_a}{a_{max}}$

Similarly to case 2, denoting $\theta_{s1} = max(\theta_{sj}, \frac{\theta_1 + (a_{min} - 1) \cdot \theta_a}{a_{min}})$ and $\theta_{s2} = min(\theta_{ej}, \frac{\theta_1 + (a_{max} - 1) \cdot \theta_a}{a_{max}})$

$$Pr^4_{I,\Xi} = \left[\frac{\theta_s \cdot a_{max} - (\theta_1 - \theta_a) \cdot \ln(\theta_s - \theta_a)}{(\theta_{ej} - \theta_{sj})(a_{max} - a_{min})} \right]_{\theta_{s1}}^{\theta_{s2}} \tag{4.16}$$

An illustration of the probability calculated in (4.16) (staying in the same bin, before the illustrated θ_2) can be seen in Figure 4.7 (rightmost example) when θ_2, θ_1 are chosen for the starting bin instead (e.g. the illustrated θ_2 is now the θ_1).

Case 5, *off* state, $a_{max} < a_1 \Rightarrow \frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}} < \theta_s$

$$Pr^5_{I,\Xi} = 0 \tag{4.17}$$

The above practically means that a_H cannot take a value large enough (out of bounds) to stay in the same state bin.
A simple example is given below for the readers:

Consider a TCL (air conditioner) with with $C_{min} = 1.5$, $R_{min} = 1.5$, $C_{max} = 1.5$, $R_{max} = 1.5$, $\theta_a = 32$, $\theta_g = -2$, temperature dead-band $\delta = 1$, with 4 bins and their respective temperature points [19.50, 19.75, 20, 20.25, 20.5]. For simulation step $\tau = 1minute$, $a_{max} = 0.997336885730501$, $a_{min} = 0.992619959819751$.

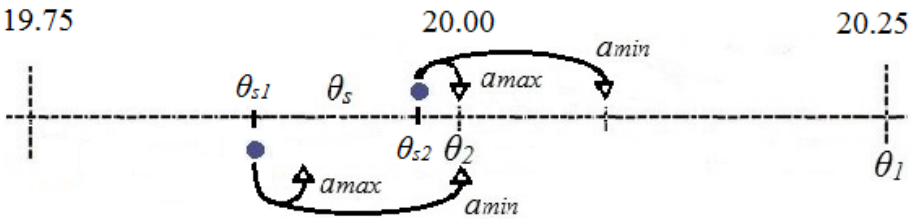


Figure 4.8: θ_s for case 2

The probability of going from bin [19.75, 20] $(\theta_{sj}, \theta_{ej})$ to bin [20, 20.25] $(\theta_{si}, \theta_{ei})$ is given by (since $Pr^4_{I,\Xi}$ gives the probability to stay in the same bin, e.g. θ_s out of bounds, and $Pr^1_{I,\Xi}, Pr^5_{I,\Xi} = 0$):

$$Pr_{I,\Xi} = Pr^2_{I,\Xi} + Pr^3_{I,\Xi}$$

First calculate limits for case 2 (these can be seen in Figure 4.8):
$\theta_{s1} = max(19.75, \frac{\theta_2+(a_{min}-1)\cdot\theta_a}{a_{min}})$ and $\theta_{s2} = min(20.00, \frac{\theta_2+(a_{max}-1)\cdot\theta_a}{a_{max}})$.

$$\frac{\theta_2+(a_{min}-1)\cdot\theta_a}{a_{min}} = \frac{20+(0.992619959819751-1)\cdot 32}{0.992619959819751} = 19.910781078612327,$$

thus $\theta_{s1} = 19.910781078612327$

$$\frac{\theta_2+(a_{max}-1)\cdot\theta_a}{a_{max}} = \frac{20+(0.997336885730501-1)\cdot 32}{0.997336885730501} = 19.967957295382110,$$

thus $\theta_{s2} = 19.967957295382110$

Then calculate $Pr^2_{I,\Xi}$, using (4.14):

$$Pr^2_{I,\Xi} = \left[\frac{(\theta_2-\theta_a)\cdot\ln(\theta_s-\theta_a)-\theta_s\cdot a_{min}}{(\theta_{ej}-\theta_{sj})(a_{max}-a_{min})}\right]_{\theta_{s1}}^{\theta_{s2}} \Rightarrow \frac{(\theta_2-\theta_a)\cdot\ln(\frac{\theta_{s2}-\theta_a}{\theta_{s1}-\theta_a})-(\theta_{s2}-\theta_{s1})\cdot a_{min}}{(\theta_{ej}-\theta_{sj})(a_{max}-a_{min})} =$$

$$\frac{(20-32)\cdot\ln(\frac{\theta_{s2}-32}{\theta_{s1}-32})-(\theta_{s2}-\theta_{s1})\cdot a_{min}}{(20.25-20)(a_{max}-a_{min})} = \frac{(-12)\cdot(-0.004740740740741)-(0.057176216769783)\cdot a_{min}}{(20.25-20)(a_{max}-a_{min})} =$$

$$\frac{0.056888888888889-0.056754253992667}{(0.25)(0.004716925910750)} = \frac{0.028542932148801}{0.25}$$

Similarly for case 3 limits:
$\theta_{s1} = max(19.75, \frac{\theta_2+(a_{max}-1)\cdot\theta_a}{a_{max}})$ and $\theta_{s2} = min(20.00, \frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}})$.

$$\frac{\theta_2+(a_{max}-1)\cdot\theta_a}{a_{max}} = \frac{20+(0.997336885730501-1)\cdot 32}{0.997336885730501} = 19.967957295382110,$$

thus $\theta_{s1} = 19.967957295382110$

$$\frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}} = \frac{20.25+(0.992619959819751-1)\cdot 32}{0.992619959819751} = 21.400556929841688,$$

thus $\theta_{s2} = 20$

Then calculate $Pr^3_{I,\Xi}$, using (4.15):

$$Pr^3_{I,\Xi} = \left[\frac{\theta_s}{\theta_{ej}-\theta_{sj}}\right]_{\theta_{s1}}^{\theta_{s2}} \Rightarrow \frac{\theta_{s2}-\theta_{s1}}{\theta_{ej}-\theta_{sj}} = \frac{0.032042704617890}{0.25}$$

The probability of moving from bin [19.75, 20.00] to bin [20.00, 20.25] is thus $\frac{0.028542932148801}{0.25} + \frac{0.032042704617890}{0.25} = \frac{0.060585636766691}{0.25} = 0.242342547066765$

The probability of staying in bin [19.75, 20] ($\theta_{sj} = \theta_{si}, \theta_{ej} = \theta_{ei}$) is given by:

$$Pr_{I,I} = Pr^3_{I,I} + Pr^4_{I,I}$$

Alternatively, it could be given as $1-(Pr^2_{I,\Xi}+Pr^3_{I,\Xi})$, which was calculated before. Yet, we will calculate it as $Pr^3_{I,I} + Pr^4_{I,I}$ and check that the sum is equal to 1.

For case 3 limits:
$\theta_{s1} = max(19.75, \frac{\theta_2+(a_{max}-1)\cdot\theta_a}{a_{max}})$ and $\theta_{s2} = min(20.00, \frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}})$.

$$\frac{\theta_2+(a_{max}-1)\cdot\theta_a}{a_{max}} = \frac{19.75+(0.997336885730501-1)\cdot 32}{0.997336885730501} = 19.717289739035905,$$

thus $\theta_{s1} = 19.75$

$$\frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}} = \frac{20.00+(0.992619959819751-1)\cdot 32}{0.992619959819751} = 19.910781078612327,$$

thus $\theta_{s2} = 19.910781078612327$

Then calculate $Pr^3_{I,I}$, using (4.15):

$$Pr^3_{I,I} = \left[\frac{\theta_s}{\theta_{ej}-\theta_{sj}}\right]_{\theta_{s1}}^{\theta_{s2}} \Rightarrow \frac{\theta_{s2}-\theta_{s1}}{\theta_{ej}-\theta_{sj}} = \frac{0.160781078612327}{0.25}$$

Similarly for case 4 limits:
$\theta_{s1} = max(19.75, \frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}})$ and $\theta_{s2} = min(20.00, \frac{\theta_1+(a_{max}-1)\cdot\theta_a}{a_{max}})$.

$$\frac{\theta_1+(a_{min}-1)\cdot\theta_a}{a_{min}} = \frac{20.00+(0.992619959819751-1)\cdot 32}{0.992619959819751} = 19.910781078612327,$$

thus $\theta_{s1} = 19.910781078612327$

$$\frac{\theta_1+(a_{max}-1)\cdot\theta_a}{a_{max}} = \frac{20.00+(0.997336885730501-1)\cdot 32}{0.997336885730501} = 19.967957295382110,$$

thus $\theta_{s2} = 19.967957295382110$

Then calculate $Pr^4_{I,I}$, using (4.16):

$$Pr^4_{I,I} = \left[\frac{\theta_s\cdot a_{max}-(\theta_1-\theta_a)\cdot\ln(\theta_s-\theta_a)}{(\theta_{ej}-\theta_{sj})(a_{max}-a_{min})}\right]_{\theta_{s1}}^{\theta_{s2}} \Rightarrow \frac{(\theta_{s2}-\theta_{s1})\cdot a_{max}-(\theta_1-\theta_a)\cdot\ln(\frac{\theta_{s2}-\theta_a}{\theta_{s1}-\theta_a})}{(\theta_{ej}-\theta_{sj})(a_{max}-a_{min})}$$

$$\frac{0.057176216769783\cdot 0.997336885730501-(-12)\cdot(-0.004740740740741)}{(0.25)(0.004716925910750)} =$$

$$\frac{0.028633284620309}{0.25}$$

The probability of staying in bin [19.75, 20.00] is thus
$\frac{0.160781078612327}{0.25} + \frac{0.028633284620309}{0.25} = \frac{0.189414363232636}{0.25} = 0.757657452930544$

Summing the above probabilities $Pr_{I,I} + Pr_{I,\Xi}$ gives $\frac{0.060585636766691}{0.25} + \frac{0.189414363232636}{0.25} = \frac{0.249999999999327}{0.25} = 1$

Now the solution of 4.6a from bin $(\theta_{sj}, \theta_{ej})$ to bin $(\theta_{si}, \theta_{ei})$ in *on* state for each case accordingly:

Case 1, *on* state, $a_2 < a_{min}$

$$Pr^1_{I,\Xi} = 0 \tag{4.18}$$

Case 2, *on* state, $a_{min} < a_2 < a_{max}$

$$Pr^2_{I,\Xi} = \left[\frac{(\theta_2-\theta_a-\theta_g)\cdot\ln(\theta_s-\theta_a-\theta_g)-\theta_s\cdot a_{min}}{(\theta_{ej}-\theta_{sj})(a_{max}-a_{min})}\right]_{\theta_{s1}}^{\theta_{s2}} \tag{4.19}$$

Case 3, *on* state, $a_{max} < a_2$

$$Pr^3_{I,\Xi} = \left[\frac{\theta_s}{\theta_{ej}-\theta_{sj}}\right]_{\theta_{s1}}^{\theta_{s2}} \tag{4.20}$$

Case 4, *on* state, $a_{min} < a_1 < a_{max}$

$$Pr^4_{I,\Xi} = \left[\frac{\theta_s \cdot a_{max} - (\theta_1 - \theta_a - \theta_g) \cdot \ln(\theta_s - \theta_a - \theta_g)}{(\theta_{ej} - \theta_{sj})(a_{max} - a_{min})}\right]^{\theta_{s2}}_{\theta_{s1}} \tag{4.21}$$

Case 5, *on* state, $a_{max} < a_1$

$$Pr^5_{I,\Xi} = 0 \tag{4.22}$$

The full TPM $\mathbf{A}$ of the example is given below (rounded to $4th$ digit):

$$\begin{bmatrix}
0.7526 & 0 & 0 & 0 & 0 & 0 & 0 & 0.3130 \\
0.2474 & 0.7577 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0.2423 & 0.7627 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0.2373 & 0.7678 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0.2322 & 0.6718 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.3282 & 0.6769 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.3231 & 0.6819 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.3181 & 0.6870
\end{bmatrix}$$

As TCLs advance in each state, the rate of temperature change reduces, thus transition probabilities to next bins reduce, which is reflected in the example above. The general form of matrix $\mathbf{A}$ is (for 4 state bins):

$$\boldsymbol{A} = \begin{bmatrix}
Pr_{1,1} & 0 & 0 & 0 & 0 & 0 & 0 & Pr_{1,8} \\
Pr_{2,1} & Pr_{2,2} & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & Pr_{3,2} & Pr_{3,3} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & Pr_{4,3} & Pr_{4,4} & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & Pr_{5,4} & Pr_{5,5} & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & Pr_{6,5} & Pr_{6,6} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & Pr_{7,6} & Pr_{7,7} & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & Pr_{8,7} & Pr_{8,8}
\end{bmatrix} \tag{4.23}$$

Looking at (4.13) to (4.22) we can see that transition probabilities from one state bin to another can be calculated if θ_{ie}, θ_{is}, θ_a, θ_g, a_H are known. The temperature set-point/dead-band and number of bins gives θ_{ie}, θ_{is} without the need of extra precise thermal sensors (e.g. cold loads with internal sensor of accuracy < 0.1 ^oC), whilst θ_g can be calculated from (3.40) and (3.41) if θ_a is approximately known. In houses the approximate ambient room temperature is known, simply by the heating settings, though if Smart Home Systems are available it would be easier and straightforward to communicate such information. In any case, some sort of simple "intelligent" infrastructure (even only a Smart Meter) is required. Obviously, H is not known and cannot be measured directly, though it affects t_{on}, t_{off} (as per (3.46), (3.48)) and thus is captured indirectly to some extent indirectly via a_H as described in previous section. This shows that transition probabilities should be updated in regular (sub hourly) intervals. This topic is further investigated in Chapter 5, Section 5.5.

At this point is also deemed important to note differences with [29, 40] in the last part of this section:

1. The effect of external factors is assumed via a noise process (defined as "$\omega_{i,t}$") with mean value 0, thus not considered in transition probabilities. This means that two TCL with same physical characteristics (R, C) will have the same probabilities, whilst here that is not the case and a_H changes values to account for increased duty cycle and consumption. It results in increased t_{on}, thus lower transition probabilities in *on* state compared to the actual a (if physical characteristics of TCLs are only considered); the opposite for off state.

2. Also, (4.6a) and (13) of [29] (and similar papers by the same authors, e.g. [40]) are different. Moreover, instead of solving the integrals (here solutions are given in (4.13) to (4.22)), as state in [29] "For each combination of starting and ending

bins, we evaluate (13) numerically to generate the analytically-derived **A**-matrix". The exact numerical process is given in more detail in [116] with steps:

- discretize the starting temperature bin into 1,000 'starting temperatures'
- compute for each the range of reachable temperatures
- discretize the total range of reachable temperatures into 1,000 'ending temperatures
- sum the number of starting temperatures that could end up in each ending temperature
- normalize the sums so that the total probability of going from all starting temperatures to all ending temperatures is 1
- map the ending temperatures to the ending temperature bin

3. The form of **A** matrix as noted in [40] is the following:

$$\boldsymbol{A} = \begin{bmatrix} \alpha_1 & 0 & 0 & 1-\alpha_2 \\ 1-\alpha_1 & \alpha_1 & 0 & 0 \\ 0 & 1-\alpha_1 & \alpha_2 & 0 \\ 0 & 0 & 1-\alpha_2 & \alpha_2 \end{bmatrix} \tag{4.24}$$

where α_1, α_2 in this matrix are probabilities (not be confused with a). Similar matrices with equal (static) transition probabilities in each state are used in [117, 118]. It is an important difference with the one given here, since the further a TCL population advances in each state the slower the rate of temperature change (as easily seen by thermal equations), meaning lower transition probabilities, which is clearly shown in the earlier example of this section. The smaller the number of state bins the smaller the error of assuming a form like (4.24); as shown in [40], increasing the number of bins from $N_{bin} = 2$ to $N_{bin} = 80$ reduced the accuracy of KF estimates, which highlights the issue. The opposite can be argued for the form proposed in this section (based on (4.13) to (4.22)).

4.5 State Space Model

The dynamics of the TCL population can be modelled using a state space model, illustrating system evolution between two consecutive time instances k and $k+1$:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{v}(k) \tag{4.25a}$$

$$\mathbf{y}(k) = \mathbf{C}\mathbf{x}(k) + \mathbf{w}(k), \tag{4.25b}$$

where the vector $\mathbf{x} = (x_1, \ldots, x_{2n}) \in \mathbb{R}^N$ represents the state vector, $\mathbf{u} \in \mathbb{Z}^n$ is the control input (zero-vector in normal operation), $\mathbf{y} \in \mathbb{R}^N$ is the output vector (power demand), vectors $\mathbf{v} \in \mathbb{R}^N$ and $\mathbf{w} \in \mathbb{R}^N$ are defined as random noise vectors containing i.i.d. (Independent and identically distributed) zero-mean Gaussian random variables. The control vector $\mathbf{u}$ influences a TCL's *on*/*off* state as an external command. By broadcasting a signal to switch TCLs between *on* and off states, they move from their current state bin to its mirror bin which has the opposite state ($on \leftrightarrow off$). The upper part of the matrix $\mathbf{B} \in \mathbb{Z}^{N \times n}$ is defined as a negative identity matrix, while the lower part is defined as an exchange matrix:

$$\boldsymbol{B} = \begin{bmatrix} -1 & & 0 \\ & \ddots & \\ 0 & & -1 \\ 0 & & 1 \\ & \iddots & \\ 1 & & 0 \end{bmatrix} \tag{4.26}$$

It is assumed that the only available measurement is aggregated power, thus $\mathbf{y}$ is a scalar and $\mathbf{C} \in \mathbb{R}^N$ is a vector:

$$\mathbf{C} = c_p \cdot \Big[0 \ldots 0, 1, \ldots 1\Big], \tag{4.27}$$

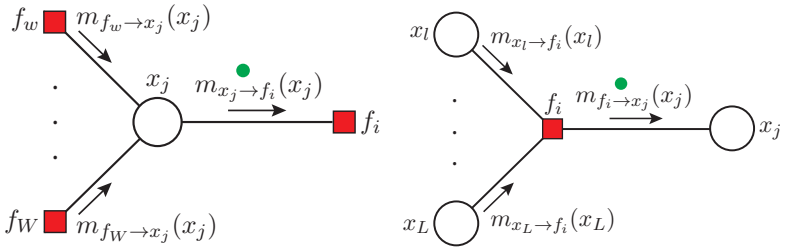
where $c_p = N_{\text{TCL}} \cdot \bar{P}_{\text{ON}}$ is a constant, N_{TCL} is the number of TCLs in the population and $\bar{P}_{\text{ON}}$ is the mean power consumption of TCLs in the *on* state. For units in the *off* state the corresponding element of the matrix $\mathbf{C}$ is equal to zero.

To estimate the state of the above state space model from the observed measurements, standard KF can be applied. Modelling and control of virtual power plants either relies on centralised or distributed approaches, where for the former, KF has been traditionally used [54, 55, 119]. In this work, the aim is to introduce BP for state estimation, a distributed approach, which can also be used centrally. Then, the aim is to demonstrate that BP can be used instead of KF with the same efficiency, while providing higher flexibility.

4.6 Factor Graphs and Belief Propagation

A factor graph is a bipartite graph that describes factorization of a global function of many variables and contains a set of variable nodes $\mathfrak{X}$ and a set of factor nodes $\mathcal{F}$. Using the BP algorithm in factor graphs, it is possible to determine exactly or approximately marginal functions derived from the global function. The BP algorithm over factor graphs is based on exchanging two types of messages along the edges of the factor graph: i) variable node to a factor node, and ii) factor node to a variable node messages. Both variable and factor nodes process the incoming messages and calculate outgoing messages [111]. If all factors of the global function and all the inputs to the BP algorithm represent the Gaussian distributions, then the corresponding BP algorithm is known as the Gaussian BP algorithm. Each message exchanged in Gaussian BP is represented using only two values: the mean and the variance [120].

The message $m_{x_j \to f_i}(x_j)$ from the variable node x_j to the factor node f_i, shown in Figure 4.9a, is equal to the product of all incoming messages $m_{f_w \to x_j}(x_j)$, ...,



(a) **(b)**

Figure 4.9: The message $m_{x_j \to f_i}(x_j)$ from the variable node x_j to the factor node f_i (subfigure a) and the message $m_{f_i \to x_j}(x_j)$ from the factor node f_i to the variable node, x_j(subfigure b)

$m_{f_W \to x_j}(x_j)$ from factor nodes to the variable node arriving from all the other adjacent edges.

If all incoming messages are Gaussian represented by their mean-variance pairs $(z_{f_w \to x_j}, \sigma^2_{f_w \to x_j})$, $\ldots$, $(z_{f_W \to x_j}, \sigma^2_{f_W \to x_j})$, it can be shown that the message $m_{x_j \to f_i}(x_j)$ is proportional ($\propto$) to the Gaussian function:

$$m_{x_j \to f_i}(x_j) \propto \mathcal{N}(z_{x_j \to f_i} | x_j, \sigma^2_{x_j \to f_i}), \tag{4.28}$$

with mean $z_{x_j \to f_i}$ and variance $\sigma^2_{x_j \to f_i}$:

$$z_{x_j \to f_i} = \left(\sum_{f_a \in \mathcal{F}_j \setminus f_i} \frac{z_{f_a \to x_j}}{\sigma^2_{f_a \to x_j}} \right) \sigma^2_{x_j \to f_i} \tag{4.29a}$$

$$\frac{1}{\sigma^2_{x_j \to f_i}} = \sum_{f_a \in \mathcal{F}_j \setminus f_i} \frac{1}{\sigma^2_{f_a \to x_j}}, \tag{4.29b}$$

where $\mathcal{F}_j \setminus f_i \subset \mathcal{F}$ defines the set of factor nodes adjacent to the variable node x_j, excluding the factor node f_i.

The message $m_{f_i \to x_j}(x_j)$ from the factor node f_i to the variable node x_j, shown

in Figure 4.9b, is defined as a product of all incoming variable node to factor node messages arriving from all the other adjacent edges, multiplied by the function associated to the factor node f_i, and marginalized over all of the variables associated with the incoming messages. The message $m_{f_i \to x_j}(x_j)$ can be computed only when all other incoming messages are known, denoted by:

$$\begin{aligned} m_{x_l \to f_i}(x_l) &\propto \mathcal{N}(z_{x_l \to f_i} | x_l, \sigma^2_{x_l \to f_i}) \\ &\vdots \\ m_{x_L \to f_i}(x_L) &\propto \mathcal{N}(z_{x_L \to f_i} | x_L, \sigma^2_{x_L \to f_i}). \end{aligned} \tag{4.30}$$

Furthermore, the Gaussian function associated with the factor node f_i is given by:

$$\mathcal{N}(z_i | x_j, x_l, \ldots, x_L, \sigma_i^2) \propto \exp\left\{ \frac{[z_i - \eta_i(x_j, x_l, \ldots, x_L)]^2}{2\sigma_i^2} \right\} \tag{4.31}$$

where $\eta_i(x_j, x_l, \ldots, x_L)$ is a linear function:

$$\eta_i(x_j, x_l, \ldots, x_L) = g_{x_j} x_j + g_{x_l} x_l + \cdots + g_{x_L} x_L. \tag{4.32}$$

Due to the linearity of function $h_i(\cdot)$, it can be shown that the message $m_{f_i \to x_j}(x_j)$ is proportional to the Gaussian function:

$$m_{f_i \to x_j}(x_j) \propto \mathcal{N}(z_{f_i \to x_j} | x_j, \sigma^2_{f_i \to x_j}), \tag{4.33}$$

with mean $z_{f_i \to x_j}$ and variance $\sigma^2_{f_i \to x_j}$:

$$z_{f_i \to x_j} = \frac{1}{g_{x_j}} \left(z_i - \sum_{x_b \in \mathcal{X}_i \setminus x_j} g_{x_b} z_{x_b \to f_i} \right) \tag{4.34a}$$

$$\sigma^2_{f_i \to x_j} = \frac{1}{g^2_{x_i}} \left(\sigma_i^2 + \sum_{x_b \in \mathcal{X}_i \setminus x_j} g^2_{x_b} \sigma^2_{x_b \to f_i} \right). \tag{4.34b}$$

where $\mathcal{X}_i \setminus x_j \subset \mathcal{X}$ is the set of variable nodes adjacent to the factor node f_i, excluding the variable node x_j.

4.7 Belief Propagation for Decentralized State Estimation

Factor graphs can be used to depict a large variety of algorithms, such as the forward-backward algorithm for hidden Markov models, probability propagation in Bayesian networks, the Viterbi algorithm and KF, to name a few. The later, KF, can be represented by a factor graph as described in [111], which connects any time instance k with the next one $(k+1)$. KF can be seen as the forward sum-product recursion through the factor graph (i.e. Figure 4.10) and yields the posterior probability distribution of the state given observations up to time k.

4.7.1 The Choice of Belief Propagation

Various aggregation and control methods have been suggested in literature, mainly fully central or decentralised ones, but also some mixed ones. In a central approach, the aggregation and the control are determined in a server, which gathers information on units and computes available Power for DR. It usually assumes high system observability (precise sensors) [28, 29, 55]. Even so, State Estimation, via KF (Kalman Filtering) has been suggested in central models to improve accuracy [29, 54, 55]. Note that in these models ([29, 54, 55]) there is the assumption of knowledge of the total aggregated TCL power (P_{tot}) per substation, which assumes that all TCLs in one substation belong to the same aggregator/VPP (meaning also same utility or same third party contracts). This is unrealistic since not all consumers will opt-in in DR, neither be contracted

to the same companies. Moreover, it would require each Distributed System Operator to install measuring devices with high resolution in each substation and then use disaggregation techniques, before finally sending such information to the aggregator/VPP. Besides the previous issues, this also raises privacy concerns, unless all consumers have agreed to it and opted-in.

On the other end of the spectrum, decentralised methods are based on smart agents and to the best of the author's knowledge no other decentralised State Estimation has been suggested besides the author's [121]. Instead of State Estimation, they rely on assumptions of near perfect knowledge of the state of TCLs, based on very precise measurements (i.e. thermal sensors), as well as knowledge of the grid's state in order to respond. For Balancing Services these models (e.g. [36, 37, 38]) require accurate frequency readings, in high resolution and processing of rate of change within 200ms, the power system's inertia or RoCoF in case of unbalance events, power system's damping and the system's demand or infeed loss, in order to calculate the required aggregated DR response. Also, the response of other TCLs (agents) must be predicted/known, which means total number, average power and duration. Therefore, the authors usually assume that such information is been updated from some central point on some sub-hourly intervals and that units, via their local processors (agents), take decisions accordingly to frequency changes [36, 37, 47]. Note that in reality, the response of other Balancing Services providers must be known as well, which is why such Balancing Services are contracted, coordinated and controlled by the grid operators. Besides that crucial issue, there is also the cost, when not only accurate thermal sensors but also high resolution frequency readers are required per agent. On the flip-side, agents (virtual assistants) or applications that monitor the demand, report it and control it are becoming very common. In addition, privacy concerns are less of an issue since data per consumer is not

sent to some central server but only receive power system information (consumers decide themselves to use such applications and opt-in).

Last, but not least, somewhere in the middle ground mixed models exist. There are various types, some of the most common described by Totu et al. [35] and Vinther et al. [39] assume an aggregator collecting consumer (TCL) data, receiving information from the grid operators as commercial aggregator operate currently (business as usual), computing available Power for DR and sending a reference signal during a demand response event. TCLs in this case take a local decision of how to respond. This approach is closer to a central model. Some aggregators already include DR in their portfolio in a similar set-up, albeit industrial loads are used mainly, in combination with other back-up sources, ensuring that contractual obligations, response level in MW and duration will be met.

Looking at the rapid adoption of Smart Home Systems which operate with devices that already provide readings of power states and have basic processing capabilities, the proposed approach in this thesis fits seamlessly. A few other approaches can be feasible, with some common points. Those points are that an aggregator must exist, which has updated information on available total Power for DR (and a state space model), which operates in a similar manner as current aggregators (preferably including other sources as well) and can be part of the market. As such, the signal will be central but decision local as described by Totu et al [69]. For instance a central signal -0.3 would mean that 30% of units in *on* state between [0.7, 1.0] of their deadband will switch off (similarly with a positive signal for the opposite). According to the example signal, if 10 bins for each state are assumed (20 in total), it would mean that units in the last 3 *on* bins will switch off (though not all from the bin third from the end, as that would be a bit more than 30%). Architecture examples are:

- If consumers agree to opt-in, an aggregator would collect all relevant data (power states) and TCL parameters as per Section 4.3. State Estimation is performed centrally to compute the available Power for DR. The VPP responds to System Operator's requests according to contracts (business as usual), the signal is computed centrally for DR, but the decision is local. In case of such a model, which is central, anonymous data from TCLs might be possible in some way (maybe randomized TCL ID's which update daily or so). This could work with KF as well.

- In a more decentralised approach, the agents would do the decentralised State Estimation as follows: Initial state for all TCLs is according to power state (time in given state preferably or simply on/off). In an N_{TCL} population with a given state space model (A TPM) and number of units in state *on*, the fraction of TCLs X in each state bin can be estimated. This is done via a Sum-Product Algorithm, Gaussian BP in this thesis (the readers are encouraged to read [111]). The BP message passes on aggregated information (fraction of units in each state bin) and not individual information (minimizing privacy concerns), given transition probability, previous state, current estimate and total number of units *on*. In practice, together with the BP message, transition probability (calculated as per Section 4.4), total number of units *on* (N_{on}) within N_{TCL} and P_{aver} should be send as well in a fully decentralised State Estimation, where aggregated data (state X, N_{on} within N_{TCL}, P_{aver}) should be sent to the aggregator as well.

- Given that only aggregated information needs to be send to the aggregator, another possible approach is an (Hierarchical) aggregation on some local level, potentially substation level and then on aggregator level. For example, a few agents can act as an aggregation medium and then aggregated data can be send via a Sum-Product Algorithm. Aggregated data per substation would also be beneficial if known, which can be given to System Operators

to estimate the change in Power (resulting in changes in Voltage in the network) in case of a DR action, especially if multiple VPPs exist in the Distribution Network (even when not in the same substation). Note that for small time steps (as per Section 4.4), we can make the assumption (with small error), that the fraction of units ($x_i(n)$) in one bin depends on the previous time step's fractions of itself ($x_i(n-1)$) and its preceding bin ($x_{i-1}(n-1)$) given transition probabilities accordingly ($a_{i,i}$, $ai, i-1$). As such, an agent per bin can act as the local aggregation point, where TCLs with a power state belonging to that bin will send data to the according medium and then BP messages are exchanged between those.

The first and third options can be seen as similar, where in the third option there is an intermediate step to gather information for on, off states and transition probabilities. Then only aggregated information is sent to the VPP (fraction of units in each bin, P_{ave} and TPM probabilities), meaning no states on/off or TCL parameters or even ambient temperature settings for households are sent to the VPP.

4.7.2 The Factor Graph Construction

The set of variable nodes $\mathcal{X}$ consists of two sets of bins $\mathcal{X}_{\text{off}}$ and $\mathcal{X}_{\text{on}}$. In general, each factor node from the set $\mathcal{F}$ is defined with (noisy) linear equation as defined in (4.31), and the set contains two sets of factor nodes. The set of factor nodes $\mathcal{F}_{\text{t}} \subset \mathcal{F}$ which is defined by transition equations is given in (4.25a) and the set of factor nodes $\mathcal{F}_{\text{m}} \subset \mathcal{F}$ defined by measurement equations provided by (4.25b).

To explain the concept of the proposed algorithm, a simple example with $N = 4$ bins is used. More specifically, the TCL population is divided in two temperature ranges where $\mathcal{X}_{\text{off}} = \{x_1, x_2\}$ and $\mathcal{X}_{\text{on}} = \{x_3, x_4\}$. The factor nodes

$\mathcal{F}_{\mathrm{t}} = \{f_1, f_2, f_3, f_4\}$ are defined according to (4.25a) and they describe transition from the time step k to the time step $k+1$. The factor node $\mathcal{F}_{\mathrm{m}} = \{f_y\}$ is defined by (4.25b) and represents the real-time measurement (e.g. aggregated active power) in the time step k. Hence, the factor graph in Figure 4.10 represents the discrete time state space model (4.25).

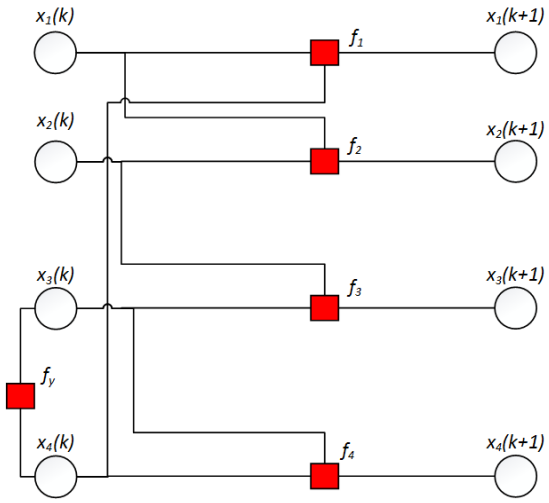


Figure 4.10: Factor graph of the state-space model for $N_{bin} = 4$

4.7.3 Algorithm

The form of the algorithm is shown in Figure 4.11. The initialization of $x_i \in \mathcal{X}$ at the time instance $k = 0$ can be done in different ways. The simplest one is to give each state bin an equal fraction of the population, thus $x_i = 1/N$. Otherwise, by having an initial assumption of the duty cycle, D (fraction of *on* state duration during one cycle), each $x_i \in \mathcal{X}_{\mathrm{on}}$ can be given an equal fraction of D and each $x_i \in \mathcal{X}_{\mathrm{off}}$ an equal fraction of $1 - D$.

Subsequently, the algorithm propagates messages through a sequence of segments of the factor graph. In the following, the segment defined by the sets of state variables in two consecutive time instances, k and $k+1$, is considered (example in

Figure 4.10). For simplicity of notation, the time indices are not used in further description of the algorithm, as they are made evident from the context. Each variable node $x_i \in \mathcal{X}$ is initialized at time instance k from messages arriving from the factor graph segment processing in time instance $k-1$.

The first step is processing messages from the factor node f_y to the variable node $x_j \in \mathcal{X}_{\text{on}}$ which can be computed using (4.34a) and (4.34b):

$$z_{f_y \to x_j} = \frac{1}{c_p}\left(z_{f_y} - c_p \sum_{x_b \in \mathcal{X}_i \setminus x_j} z_{x_b \to f_y} \right) \tag{4.35a}$$

$$\sigma^2_{f_y \to x_j} = \frac{1}{c_p^2}\left(\sigma^2_{f_y} - c_p^2 \sum_{x_b \in \mathcal{X}_i \setminus x_j} \sigma^2_{x_b \to f_y} \right). \tag{4.35b}$$

where $\mathcal{X}_i \setminus x_j \subset \mathcal{X}_{\text{on}}$ is the set of variable nodes adjacent to the factor node f_y, excluding the variable node x_j. Equations (4.35a) and (4.35b) can be simplified if z_{f_y} is expressed as a function of c_p, $z_{f_y} = z'_{f_y} \cdot c_p$. This is achieved by normalizing the measurement first. By normalizing the demand it is also easier to check the aggregated TCLs' behaviour independently of population size and demand (c_p), which are but scaling constants practically.

The second step is processing messages from variables nodes $x_j \in \mathcal{X}_{\text{on}}$ to factor nodes $f_i \in \mathcal{F}_{\text{t}}$ according to (4.29a) and (4.29b):

$$z_{x_j \to f_i} = \left(\frac{z_{x_j}}{\sigma^2_{x_j}} + \frac{z_{f_y \to x_j}}{\sigma^2_{f_y \to x_j}} \right) \sigma^2_{x_j \to f_i} \tag{4.36a}$$

$$\sigma^2_{x_j \to f_i} = \frac{\sigma^2_{f_y \to x_j} \cdot \sigma^2_{x_j}}{\sigma^2_{f_y \to x_j} + \sigma^2_{x_j}} \tag{4.36b}$$

Lastly, messages are propagated from all factor nodes $f_i \in \mathcal{F}_{\text{t}}$ to the next time

step $k+1$ using

$$z_{f_i \to x_i} = \alpha_{i,i-1} \cdot z_{x_{i-1} \to f_i} + \alpha_{i,i} \cdot z_{x_i \to f_i} + b_i \cdot u_i^k \tag{4.37a}$$

$$\sigma^2_{f_i \to x_i^{k+1}} = \sigma^2_{x_i} + \alpha^2_{i,i-1} \cdot \sigma^2_{x_{i-1} \to f_i} + \alpha^2_{i,i} \cdot \sigma^2_{x_i \to f_i} + b_i^2 \cdot \sigma^2_{u_i} \tag{4.37b}$$

where $\alpha_{i,i}$, $\alpha_{i,i-1}$ represent the elements of the Markov transition matrix $\boldsymbol{A}$.

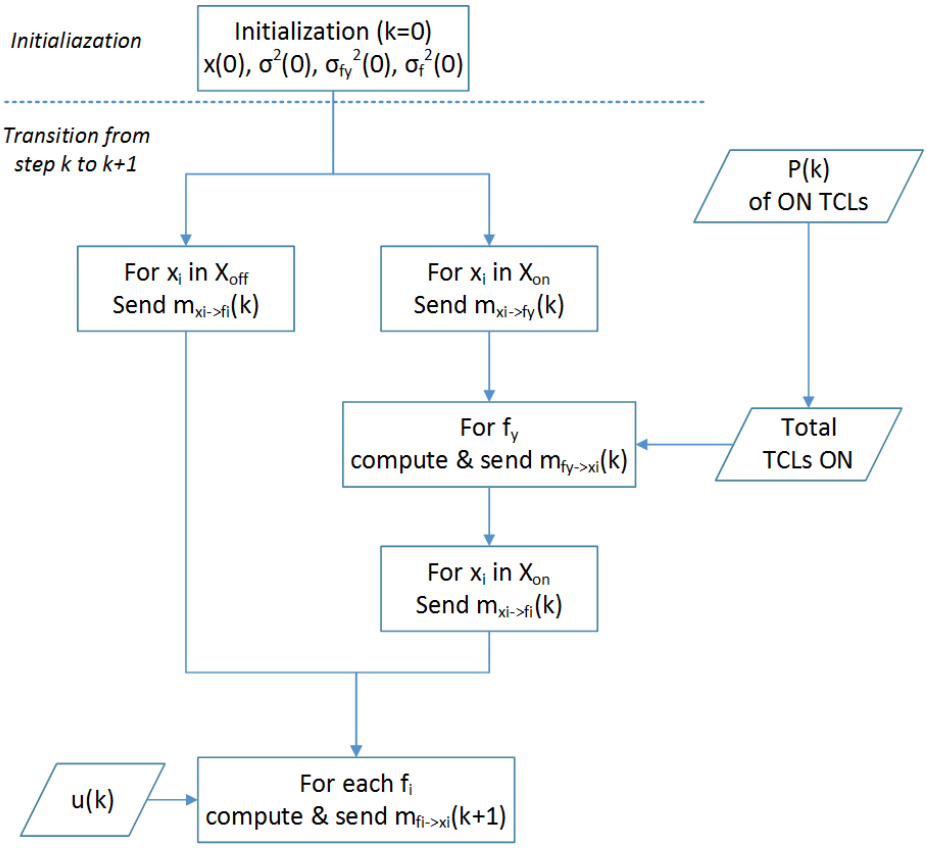


Figure 4.11: BP algorithm for State Estimation of Factor Graph 4.10

From the algorithm in Figure 4.11, the information that is given is the number of *on* units and Power, as in the base case in [29]. But here this is not via disaggregation but from Smart Meters or Smart Plugs. In some sub-hourly interval, aij should also be updated (or the info required to compute it according to Section 4.4). If time in power *on* state where to be used (as per example in Figures 4.2 & 4.3) then this could be used for initialization as well (not done in this

Chapter). Note that when control signals are sent, for TCLs that have switched state, this information will no longer be as accurate unless the expected state during the switch is considered. For example, a TCL which is halfway through the off state and is switched on, is already starting at 50% of the relative dead-band (for example see Figure 2.21) and the appropriate state bin. It would thus take approximately half the time to reach θ_{off} again. Considering the expected state at the time of switch and the duty cycle, it could be calculated in a simple manner. In general, if not done appropriately it could introduce errors during control actions.

4.8 Simulation Results

The Monte Carlo model of Chapter 3 is simulated for 10,000 TCLs using the data listed in Table 4.1 [54], as well as for the cold loads used in Chapter 3. Then, the state space model is derived for each case and simulated, using State Estimation, where the total number of units on with some noise is given by the Monte Carlo model accordingly.

Table 4.1: TCL (Air Conditioner) Parameters

Parameter	Value$^{(1)}$	Value$^{(2)}$
θ_{set}	20 ^oC	$\mathcal{U}$(15 ^oC,25 ^oC)
δ	1 ^oC	$\mathcal{U}$(0.25 ^oC,1 ^oC)
θ_a	32 ^oC	32 ^oC
R	$\mathcal{U}$(1.5 $^oC/kW$,2.5 $^oC/kW$)	$\mathcal{U}$(1.5 $^oC/kW$,2.5 $^oC/kW$)
C	$\mathcal{U}$(1.5 $kWh/^oC$,2.5 $kWh/^oC$)	$\mathcal{U}$(1.5 $kWh/^oC$,2.5 $kWh/^oC$)
θ_g	28 ^oC	$\mathcal{U}$(15 ^oC,41 ^oC)

(1) lower heterogeneity, (2) higher heterogeneity (as [54])

Air condition load data from [54], both cases have the same state space model

Monte Carlo is used to randomly draw parameters from the Uniform distributions according to Table 4.1, initial temperature θ is taken randomly from a Uniform distribution within the dead-band. Initial state m is randomly taken with a Uniform distribution between [0,1], where a successful trial (trial value > mean duty cycle) returns value $m_i = 1$; otherwise $m_i = 0$. The small heterogeneity and high heterogeneity difference is taken by using values from the according columns of Table 4.1, where in higher distribution 5 parameters taken random values within Uniform distributions instead of only 2.

For comparison purposes, $H = 0$ for the data listed in Table 4.1 [54]. The Monte Carlo on cold loads of Chapter 3 are modelled normally, as in Chapter 3. The output is normalized (with respect to c_p) in order to track only the state of the population (number of units *on*, off). Essentially, for large enough populations ($N_{TCL} >> 1,000$), behaviour of TCLs is independent of the constant c_p. Comparing KF and BP, it is evident that both algorithms yield practically the same state estimation (as expected).

4.8.1 KF and GBP Results

A comparison between GBP and KF on steady state measurement and number of iteration/time steps is performed. The initial sates are given as $x_i = 1/N_{bin}$, the population size is assumed very large ($N_{TCL} \to \infty$), thus noise due to population size tends to 0, and for a small period no changes in population are assumed.

Results are shown in Figures 4.12, 4.13 and Table 4.2. KF reaches the steady state of each bin after 15 steps, whilst BP does so after 44 steps, both give practically the same results. Number of steps between BP and KF are hard to compare since one iteration of BP is not equal to one iteration of KF, still BP should improve, something investigated in the next Chapter. Looking at these

figures also highlights the importance of a good initial state estimation and more importantly of the preceding time step (k) in order to estimate the following $(k+1)$ accurately.

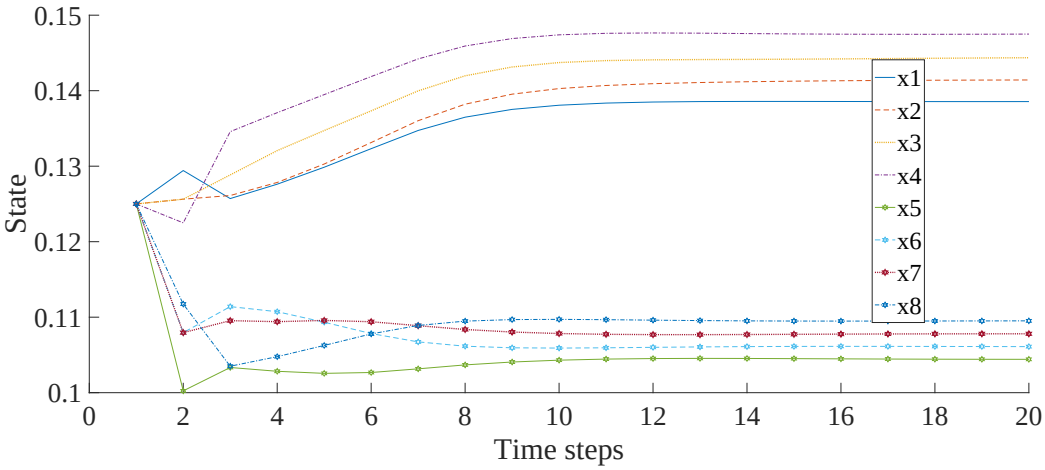


Figure 4.12: KF State Estimation. Assume $N_{TCL} \to \infty$, thus no noise due to population size, also assume no external interactions for a short period, thus the population in steady state. From initialization $x_i = 1/N_{bin}$ to X_{ss}, number of time steps are observed.

In reality, the initial state estimation of course will not be $x_i = 1/N_{bin}$, but more accurate (and obviously between time steps information of the previous step is used). Under normal operation the state $(\mathbf{x}(k))$ and aggregated demand between consecutive time steps will not vary greatly, but when control actions are used there will be some "drops/jumps" (e.g. Figures 4.14 - 4.19). Good estimation of state $(\mathbf{x}(k))$ and thus accurate use of control $(\mathbf{u}(k))$ are important in estimating that change of state $(\mathbf{x}(k+1))$ and therefore subsequent control actions and demand output.

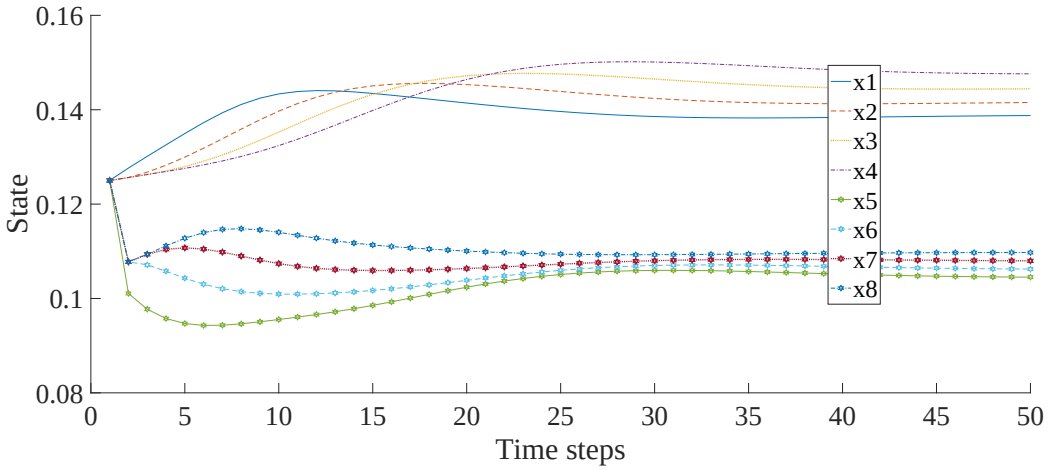


Figure 4.13: BP State Estimation. Assume $N_{TCL} \to \infty$, thus no noise due to population size, also assume no external interactions for a short period, thus the population in steady state. From initialization $x_i = 1/N_{bin}$ to X_{ss}, number of time steps are observed.

Table 4.2: KF & BP Steady State

	KF	BP
x_1	0.13856	0.13857
x_2	0.14142	0.14144
x_3	0.14443	0.14444
x_4	0.14760	0.14757
x_5	0.10448	0.10447
x_6	0.10610	0.10608
x_7	0.10780	0.10778
x_8	0.10967	0.10965

KF at 15 iterations, BP at 44, rounded at the 5 decimals

4.8.2 Simulation Results of Control Commands

Starting with the loads of Table 4.1, for a populations with lower and higher heterogeneity, at time $t = 120min$ an external signal is given to the whole TCL population to switch off. The state space model follows the MC model, with a small lag and initial error which increases at first and then reduces. The delay to "catch up" between actions (commands) and output is inherent due to the state space model itself, something explored in the next Chapter. Finally, it is obvious when comparing Figures 4.14 and 4.15 that the rebound effect depends on heterogeneity level, where higher heterogeneity results in faster damping of oscillations.

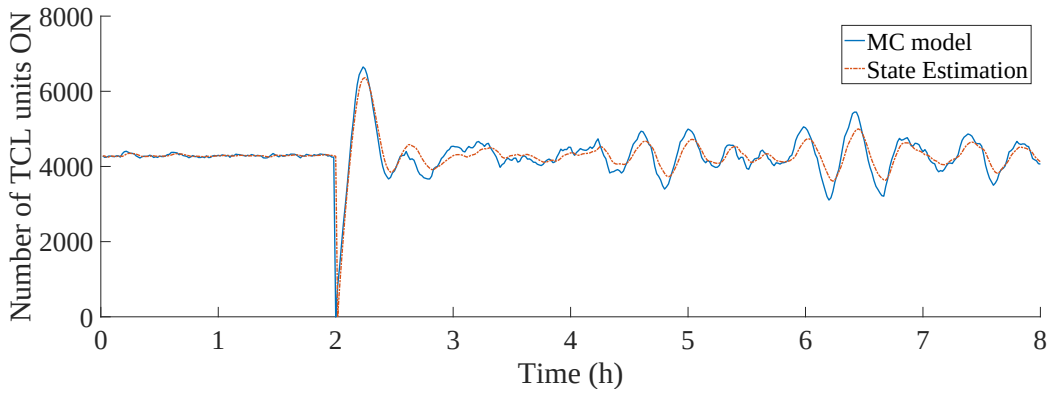


Figure 4.14: 10,000 TCLs, according to Table 4.1, lower heterogeneity, switch off signal at time 120' for whole population, short duration

In a 100% homogeneous population this oscillation would remain indefinitely and in the case of Figure 4.14 where a population with low heterogeneity (high homogeneity) is modelled, the partial synchronization remains for hours afterwards. The same conclusion comes from aggregated models based on CFPE, as shown in [28].

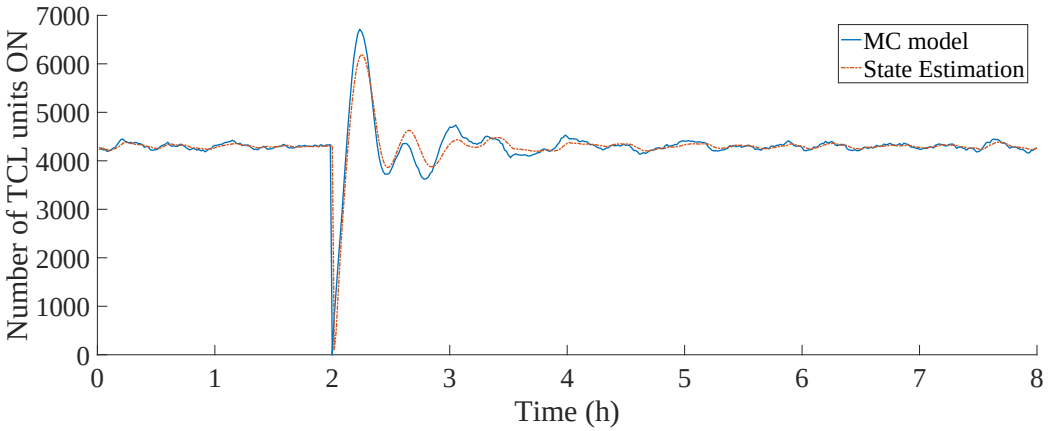


Figure 4.15: 10,000 TCLs, according to Table 4.1, higher heterogeneity, switch off signal at time 120' for whole population, short duration

A few more control commands are explored in Figures 4.16 - 4.19, with commands given to populations of TCLs of Table 4.1 as well as TCLs of Chapter 3. These commands vary in duration and percentage of units *on*. The higher the duration and number of units used, the higher the rebound afterwards.

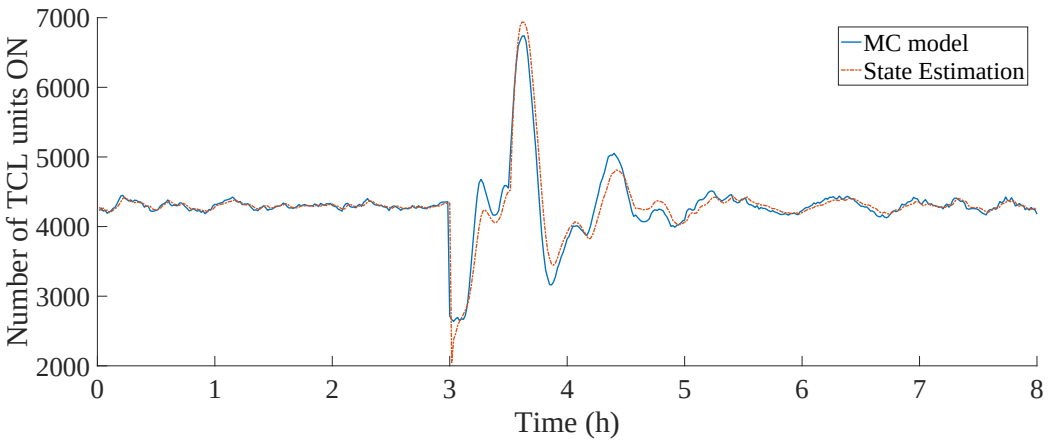


Figure 4.16: 10,000 TCLs, higher heterogeneity (according to Table 4.1), switch off signal at time 180' for 40% of units on, 10' duration.

Also, higher heterogeneity results in faster damping of oscillations. More

sophisticated control algorithms can improve this behaviour and many such have been developed in literature as discussed, thus not developed here.

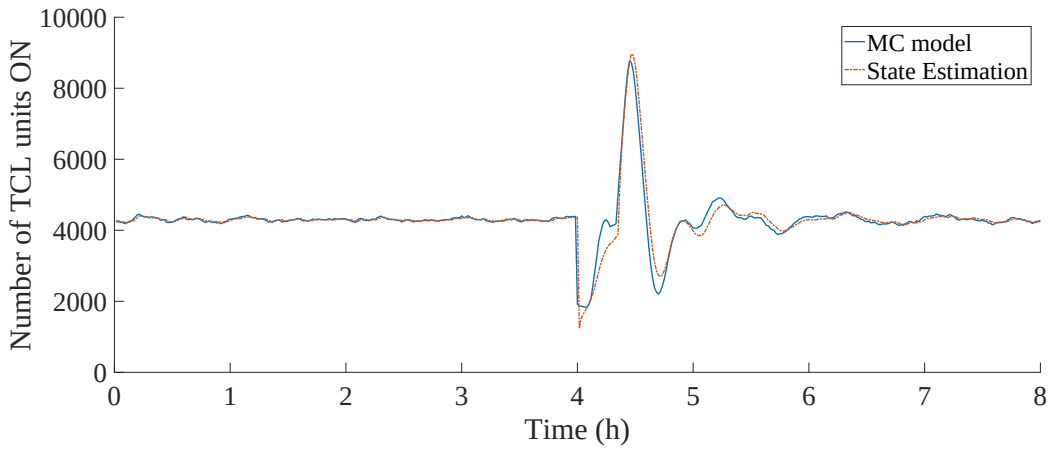


Figure 4.17: 10,000 TCLs, higher heterogeneity (according to Table 4.1), switch off signal at time 240' for 60% of units on, 6' duration.

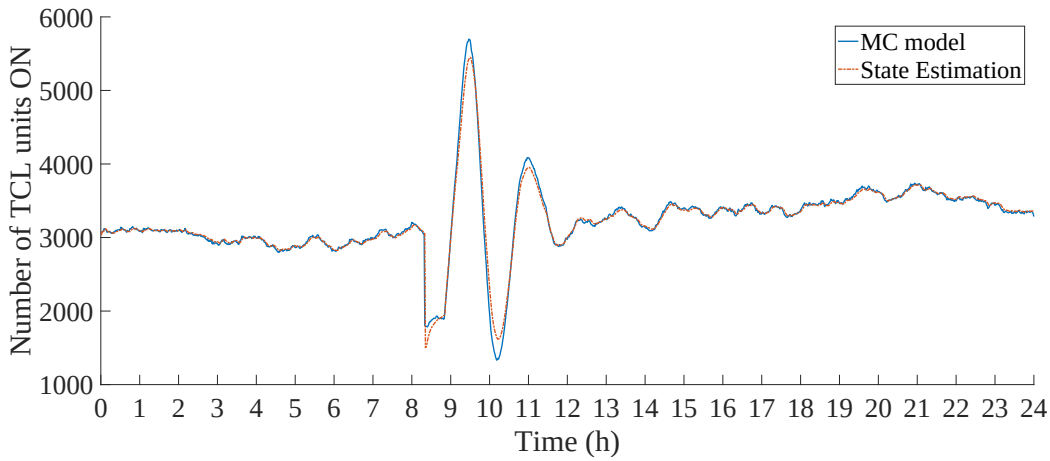


Figure 4.18: 10,000 TCLs of Chapter 3, switch off signal at time 500' for 40% of units on, 30' duration.

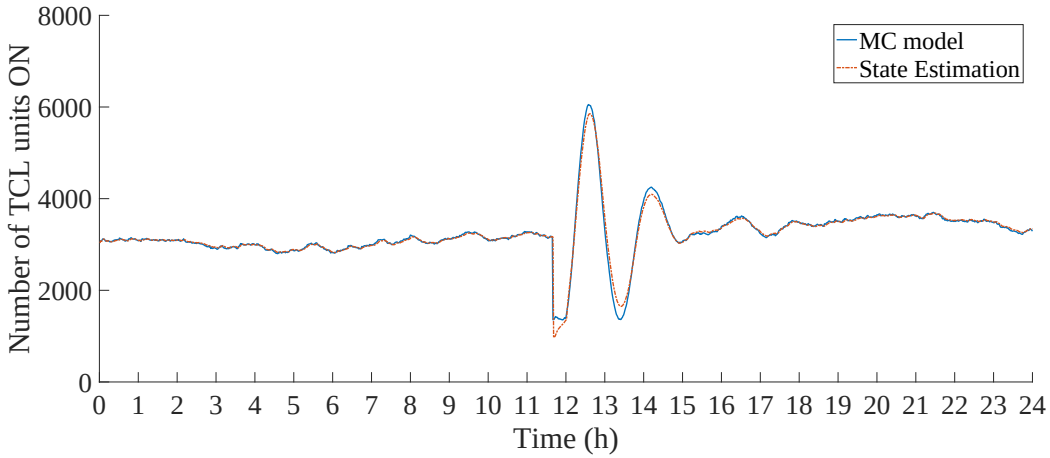


Figure 4.19: 10,000 TCLs of Chapter 3, switch off signal at time 700' for 60% of units on, 20' duration.

4.9 Conclusions

This Chapter introduces BP for state estimation in large scale populations of TCLs, as well a state space model to represent aggregated heterogeneous TCL populations, based on transition probabilities. The transition probabilities according to equations of Section 4.4 include external factors, albeit in reality, it is not possible to track them, directly that is at least.

The effect of the parameter H on duty cycle (i.e. consumption) is extending t_{on} or shortening t_{off}, as explained in Chapter 3, (3.46), (3.48), and modelled here via a_H. As observed in Figure 4.3, due to this two TCLs with different parameters can have the same duty cycle and t_{on}, t_{off}. As such, the TCL can be virtually seen as operating with different parameters (λ', θ'_g) for that period. What matters at the end of the day is tracking those t_{on}, t_{off} changes, which practically means tracking a_H and ambient changes indirectly. This can be done via tracking P_{on} duration, through a Smart Plug or Smart Meter. These could also used to approximate

thermal state (Figures 4.2, 4.3) and can be used as initial estimation, which will be explored more in the next chapter.

Another important point is that the analytical derivation method of Section 4.4 has limitations. Heterogeneity is modelled on α only and assumptions of uniform distribution of TCL parameters are used. Despite these, the results shown in simulations show that State Estimation has a good level of accuracy.

In the following chapter, a more generic aggregation method is developed, to fit a highly heterogeneous population, with heterogeneity in more than one parameters, whilst being simple and easy to calculate. Additionally, an updated FG and BP algorithm are presented, where not only information of number of *on* units is used but also of number of off units. Last, but not least, methodologies for on-line updating of the aggregation model are developed which do not rely on precise thermal sensors but only power readings.

Chapter 5

Aggregation of Heterogeneous TCLs Using Power Rates

5.1 Introduction

In state of the art models where aggregation methods are explored, heterogeneity is regarded as partial heterogeneity for modelling purposes, such as in [40] where it was only considered for a (excluding parameters θ_g, θ_a and limiting C, R) or in [28] where only small and static heterogeneity is considered. Afterwards, simulations and case studies are performed on data with heterogeneity in more parameters though, to test models derived with partial heterogeneity. As noted by the authors [40]; "If we allow R and P_{trans} to vary across TCLs, the **A**-matrix becomes harder to derive." Looking at the method examined in the previous chapter, for each extra parameter the calculation of an additional integral in (4.6a) is required, which results to an extra 3 integrals, as well as requiring the knowledge of each parameter's distribution. It is hard to add external factors or reflect their dynamic behaviour (non-static parameters and demand in time,

Chapter 3). In addition, in order to track aggregated populations (such as for State Estimation [29]) the information of units *on* and aggregated Power are measured, whilst information of units off is not considered.

In this chapter, a new aggregation approach is developed, which allows for a simple calculation of the TPM, with heterogeneity added in multiple parameters, to better represent real world scenarios. Additionally, improvements to the state space model and state estimation are introduced alongside the new aggregation method. As shown, seemingly different types of TCLs can be aggregated, given similar behaviour to control actions, something indicated in Chapter 3. Hence, a universal, flexible aggregation model is advantageous for aggregators. Stochastic errors diminish as population size increases, leading to higher accuracy. In the case of balancing services, such as FFR, high accuracy is important. TPM and heating/cooling rates are the base of the aggregated model and BP for state estimation. The result is a more accurate, flexible, and universal model.

The developed aggregation model was created with the intention to be easily updated in time in order to match the dynamic behaviour of TCLs. Lastly, any type of loads which follow a similar behaviour or cycles, such as batteries, can be modelled in similar frameworks [7, 23], using a TPM.

The chapter is organized as follows. Section 5.2 details how the aggregation model is formed. Section 5.3 describes factor graphs and the BP algorithm for this model. Markov aggregation model is formed. Section 5.4 presents simulation results and a comparison with the previous model. Section 5.5 describes the on-line updating methodologies and Section 5.6 concludes.

5.2 *On/off* Duration to Transition Probability

The method used in Chapter 4 has the following limitations:

- Partial heterogeneity (assumed on one parameter only, a)
- Requirement of uniform distribution of that parameter (4.6b)
- Complex to model and calculate heterogeneity in more parameters as shown in [40]
- Unpractical to measure/consider H (external influence) directly, as discussed in Chapter 4, where a_H is used instead

The new method is based on deriving transition probabilities based on t_{on}, t_{off} and the relative temperature advance in time, as shown in Figures 4.2, 4.3. If a TCL was in *on* state which had a duration t_{on} equal to $30'$, then we could say that for that state, every minute, the TCL advanced on average 1/30 of the dead-band (between θ_{on} and θ_{off}). That means on average 3.33% relative temperature advance every minute. If we separated the dead-band in 10 equal bins, on average, every minute the TCL would have advanced 1/3 of each bin. We could say, that given a random point within a bin, on average, the probability of advancing to the next bin is 1/3 (in a similar manner to Chapter 4).

Of course, the rate ($\frac{d\theta(t)}{dt}$) reduces as units advance towards the switching points, θ_t being an exponential function, thus the transition probability reduces as TCLs near the end of each state. This shows that units in the first state bins (following a switch) have a higher probability of moving to the next state bins. This has being captured in the TPM of Chapter 4, one such TPM can be seen in Table 5.1. The deviation from that average practically depends on parameters θ_a , θ_g, λ. Yet TPMs used in state of the art use a static (same) probability in the TPM

for each state, practically an approximate average value [29, 40], (this would be correct if $\theta(t)$ was a linear function and not an exponential).

In reality during either the heating or cooling part of the cycle, an individual TCL may experience a change in temperature opposing its state or hastening it, due to external factors (see (3.46), (3.48)). Overall though, each part of the cycle will be completed. Also, when large aggregated populations are considered and their relative dead-band (0% at $\theta(on)$, 100% at θ_{off}), which are not synchronized, then these interactions or changes in ambient temperature can be considered randomly distributed within the relative dead-band in any point in time.

Using (3.18) the *on* and *off* rates at different temperatures are calculated:

$$\varrho_{on} = -\lambda \cdot (\theta(t) - \theta_a(t) - \theta_g) + H(t) \tag{5.1a}$$

$$\varrho_{off} = -\lambda \cdot (\theta(t) - \theta_a(t)) + H(t) \tag{5.1b}$$

where $H(t) \equiv \lambda \cdot \upsilon(t) \cdot \theta_e(t)$.

Parameters ϱ_{on} and ϱ_{off} represent the rate at which a TCL' temperature "advances", consecutively, so does a population of TCLs and these equations can be assumed for each individual TCL.

The starting (switch on), ending (switch off) and mean rates are:

$$\varrho_{on,start} = \frac{d\theta(t_0)}{dt} = -\lambda \cdot (\theta_{on} - \theta_a(t_0) - \theta_g) + H(t_0) \tag{5.2a}$$

$$\varrho_{on,end} = \frac{d\theta(t_{on})}{dt} = -\lambda \cdot (\theta_{off} - \theta_a(t_{on}) - \theta_g) + H(t_{off}) \tag{5.2b}$$

$$\varrho_{on,mean} = \frac{\varrho_{on,start} + \varrho_{on,end}}{2} = -\lambda \cdot (\frac{\theta_{off} + \theta_{on}}{2} - \bar{\theta_a} - \theta_g) + \bar{H} \tag{5.2c}$$

$$\varrho_{off,start} = \frac{d\theta(t_0)}{dt} = -\lambda \cdot (\theta_{off} - \theta_a(t_0)) + H(t_0) \tag{5.3a}$$

$$\varrho_{off,end} = \frac{d\theta(t_{off})}{dt} = -\lambda \cdot (\theta_{on} - \theta_a(t_{off})) + H(t_{off}) \tag{5.3b}$$

$$\varrho_{off,mean} = \frac{\varrho_{off,start} + \varrho_{off,end}}{2} = -\lambda \cdot (\frac{\theta_{off} + \theta_{on}}{2} - \bar{\theta_a}) + \bar{H} \tag{5.3c}$$

This approximation will have a small error, as seen in Figure 4.3.

At this point, by partitioning the dead-band in N_{bin} number of bins ($N_{bin}/2$ for each state) we may calculate the mean rate of the nth bin (θ_n, θ_{n+1}) as

$$\varrho_{on,n} = -\lambda \cdot (\frac{\theta_n + \theta_{n+1}}{2} - \bar{\theta_a} - \theta_g) + \bar{H} \tag{5.4}$$

$$\varrho_{off,n} = -\lambda \cdot (\frac{\theta_n + \theta_{n+1}}{2} - \bar{\theta_a}) + \bar{H} \tag{5.5}$$

The transition probability from one state bin (θ_n, θ_{n+1}) can be calculated as the probability of a random TCL, within the state bin, moving out of it in one time step. Temperature change in one time step τ is:

$$\Delta\theta_{on,n} = \varrho_{on,n} \cdot \tau \tag{5.6a}$$

$$\Delta\theta_{off,n} = \varrho_{off,n} \cdot \tau \tag{5.6b}$$

The above practically gives the expected temperature change of TCLs, within the nth bin, in one time step τ. Thus the probability of moving out of the n-th state bin in one time step is the fraction of temperature change with respect to the size of the bin.

$$Pr_{on,n} = \frac{\Delta\theta_{on,n}}{\theta_{n+1} - \theta_n} \tag{5.7a}$$

$$Pr_{off,n} = \frac{\Delta\theta_{off,n}}{\theta_{n+1} - \theta_n} \tag{5.7b}$$

The probability of staying in the same state bin is obviously $1 - Pr_n$, given the same assumption as in Chapter 4, that for a small time step the potential move in bin is limited to the next or the existing bin.

TCLs with similar probability transitions, independently of actual TCL parameters, can be aggregated together if they have similar TPMs. As such, heterogeneity in multiple TCL parameters is not an issue. Moreover it is a simpler method where when seemingly **different** TCLs (λ, θ_a, θ_g, θ_{max}, θ_{min}), have similar TPMs, their response to control actions is similar. This allows for aggregation of larger TCL populations (even mixed heating and cooling loads), therefore coordinated control actions should be more accurate due to lower stochasticity related to population size (Section 3.4).

5.3 State Space Model and Factor Graph

The state space model used in Chapter 4 and literature ([29, 40]), as shown in Figures 4.14 - 4.19, assumes that the effects of actions ($\mathbf{u}(k)$) in time step "k" take place and are seen in next time step "k+1". Therefore, in a simulation with a time step of 1 minute, when a control action ($u(k)$) is given within that minute, it will be accounted in the next minute ($x(k+1)$), even though a change in the Monte Carlo has already occurred. The State Estimation will try to respond to that according to measurement $y(k)$ and a small error is observed for time step $k+1$. The expected practical delay (between command and action) in most VPPs' will be below $200ms$ since they use internet connection (internet connection is usually around $50ms$). If the state space model runs with a control step and measurement update step <1second (around $200ms$) then (4.25) should be used. Whether in reality measurements and data on TCLs would be available in such time resolution is unknown. For this thesis 1 minute resolution is assumed. As such, using (4.25) will have that 1 time step lag and initial error.

Note that for Balancing Services the response time should be fast and in the case of Frequency Response Services the requirements can be response <1second (EFR, FFR etc.). Given these specification, a VPP would probably need to have state space models in the order of a few seconds, if not faster, with control actions capture at the same time step.

Otherwise, there will be a lag equal to one time step and some related error. In which case, it is more accurate to use the following, where control actions are seen in the same time step [122]:

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k+1) + \mathbf{v}(k+1) \tag{5.8a}$$

$$\mathbf{y_1}(k+1) = \mathbf{C_1}\mathbf{x}(k+1) + \mathbf{w}(k+1) \tag{5.8b}$$

$$\mathbf{y_2}(k+1) = \mathbf{C_2}\mathbf{x}(k+1) + \mathbf{w}(k+1), \tag{5.8c}$$

where $\mathbf{C_1} = N_{TCL} \cdot \Big[0 \dots 0, 1, \dots 1\Big]$ and $\mathbf{C_2} = N_{TCL} \cdot \Big[1, \dots 1, 0 \dots 0\Big]$. Here we can use $\mathbf{y} = [\mathbf{y_1}; \mathbf{y_2}]$ and $\mathbf{C} = [\mathbf{C_1}; \mathbf{C_2}]$ to get the same form as used in Chapter 4. Estimation of the next time step (k+1), based on the current measurement (k), is important for control too. A good estimation for the next time steps $(k+1)$ allows for control choices with lower error. Therefore, less corrective actions are required, a key point for open-loop control or control where measurements cannot be updated fast enough. The actual difference between (5.8) and (4.25) is that the former accounts for control actions in the same time step and not the next one, which is preferred here for the given simulation time step as explained above.

5.3.1 Factor Graph and Algorithm

The factor graph in this case uses the same concept as described in the Chapter 4, with the addition of an extra factor node. An example with $N = 4$ bins is

shown in Figure 5.1. The new factor node $\mathcal{F}_{\mathrm{m}} = \{f_{y_2}\}$ is defined as $c_p - y(k)$. The factor graph in Fig. 5.1 represents the discrete time state space model (5.8).

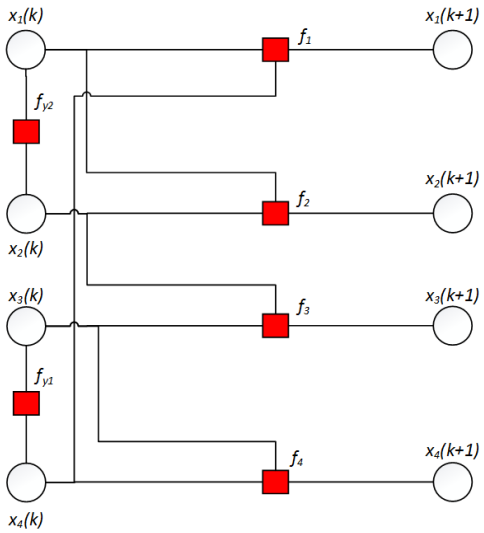


Figure 5.1: Factor graph of the state-space model for $N_{bin} = 4$

The form of the algorithm of the above FG is shown in Algorithm 5.2, where besides the extra factor node, a new initialization is used. The point of the extra factor is increased information from the TCL population, where now not only number of units *on* is tracked for State Estimation, but also number of units off. The extra information of number of units in off state can be used in a similar manner to number of units in *on* state for state estimation, thus adding information to the system. Assuming a population of N units, by tracking the number of units in *on* state, N_{on}, someone can easily calculate N_{off}. In reality though, there will be units which are in the off state of the duty cycle (meaning idle) and units not in operation, thus there will be an error. As such it is best to track units in off state. The initialization of $x_i \in \mathcal{X}$ at the time instance $k = 0$ can be done as in Chapter 4. The most accurate method though is via the Steady-State vector or better via t_{on}, t_{off} if power state duration is considered, as shown and discussed in Figures 4.2, 4.3. The former would be used if only power states *on*, off are available and not current t_{on}, t_{off}; though it would be easy to

calculate t_{on}, t_{off} by having power states updated every minute or so, it might not be done for privacy issues. In which case, the available info is only state $(on\ /\ off)$, potentially send with randomized ID. The extra factor makes the BP algorithm more accurate, as $x_i \in X_{off}$ are more accurate, meaning also the transition from the last bin off to the first bin on is more accurate. In addition this has value for TCLs that are not operating 24/7 (like cold loads do) and the actual number of TCLs N_{TCL} in operation changes dynamically.

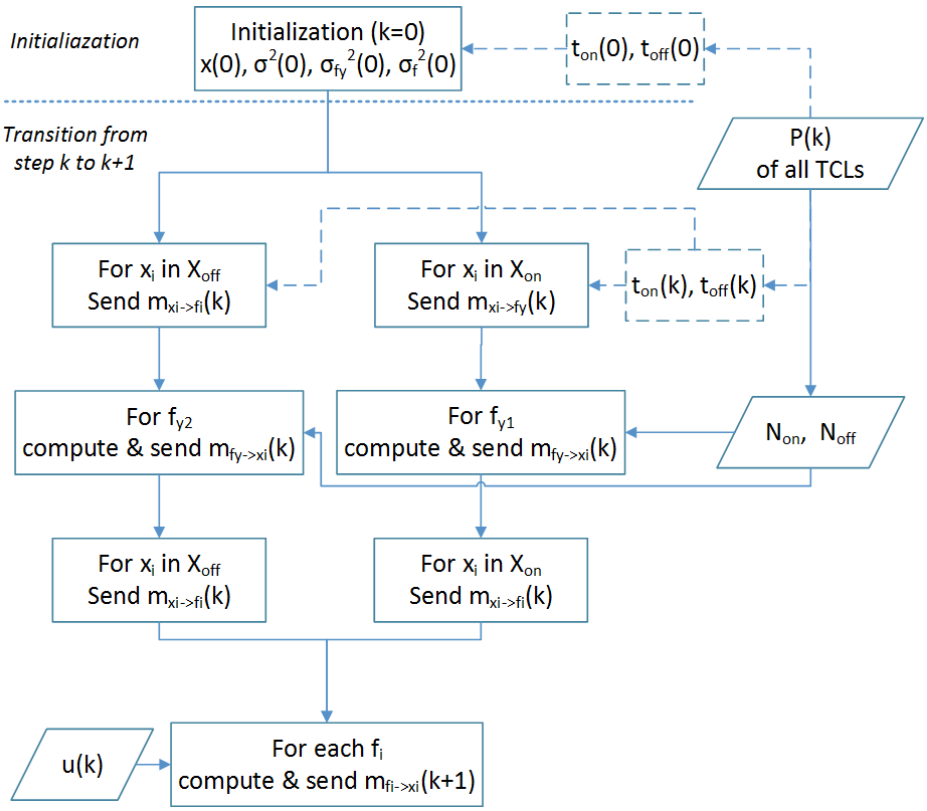


Figure 5.2: BP algorithm for State Estimation of Factor Graph 5.1. Dotted lines denote extra info which may be used but not a necessity.

Regular Transition Matrix: A transition matrix is regular when there is a power of that matrix that contains all positive non-zero entries. The 3 main rules of a Regular Transition Matrixes are:

1. If the transition matrix is not irreducible (not connectable), then it is not regular
2. If the transition matrix is irreducible (connectable) and at least one entry of the main diagonal is nonzero, then it is regular
3. If all entries on the main diagonal are zero, but T^n (after multiplying by itself n times) contains all positive entries, then it is regular.

In this case, the TPM (matrix $\mathbf{A}$ is regular because of condition 2. Alternatively, after multiplying with itself $r-2$ times (where r the size of rows), $\mathbf{A}^{(r-2)}$ has no 0 entries.

5.3.2 Initialization

The Steady-State Vector: A regular Transition Matrix is one that always achieves a steady state. The probability vector which remains unchanged when it is multiplied by the transition matrix is the Steady-State vector. The Steady-State vector, $\mathbf{x}_{ss}$ satisfies the equation:

$$\mathbf{A}\mathbf{x}_{ss} = \mathbf{x}_{ss} \leftrightarrow (\mathbf{A} - \mathbf{I})\mathbf{x}_{ss} = 0 \tag{5.9}$$

where $\mathbf{A}$ here is used in a generic way. We will be solving for the $\mathbf{A}$ defined in (5.8), to get its steady state vector. As long as the steady-state vector of a transition matrix is found, it can be considered regular, regardless of the entries in the transition matrix (excluding matrices with negative integers, since probabilities are always positive).

Solving the above for $\mathbf{A} \in \mathbb{Z}^{2n \times 2n}$ gives (where α_{ij} elements of $\mathbf{A}$):
$(1 - \alpha_{11})x_{s1} = \alpha_{1n}x_{s2n}$
$(1 - \alpha_{22})x_{s2} = \alpha_{21}x_{s1}$
.

.

$(1-\alpha_{2n2n})x_{s2n} = \alpha_{2n2n-1}x_{s2n-1}$

It should be noted here that α_{ii} is the probability of not moving within the next time step, while $\alpha_{i+1,i}$ is the probability of moving. Thus $\alpha_{i+1,i} = 1-\alpha_{ii}$. Substituting to the above:

$$(1-\alpha_{11})x_{s1} = (1-\alpha_{22})x_{s2} = ... = (1-\alpha_{2n2n})x_{s2n} \tag{5.10}$$

This equation shows that in a "steady state", the flow of units (from one bin to the next one) remains steady. In an infinite number of units the above should hold true for any given moment, assuming no external influence.

This proves the quasi-equilibrium condition of the natural diversity of a cycling load, in steady state (3.10). In steady state, the flow of units, switching from *on* state to *off* state has to match those switching to the opposite directions, in steady state, otherwise the aggregated consumption would change. Duty cycle is the sum of all fractions of units in *on* state (these are given by the first n bins), therefore:

$$D = \sum x_i \in X_{on} = \sum_{n+1}^{2n} x_n \tag{5.11}$$

$$1-D = \sum x_i \in X_{off} = \sum_{1}^{n} x_n \tag{5.12}$$

Equations (5.11), (5.12) and (5.10) can be used to express $\mathbf{x}_{ss}$ (x_{si}) as a function of A (α_{ij}). Solution:

$1-D = x_{s1} + x_{s2} + x_{s3} + ... + x_{sn} \Rightarrow$

$1-D = x_{s1} + \frac{1-\alpha_{11}}{1-\alpha_{22}}x_{s1} + \frac{1-\alpha_{11}}{1-\alpha_{33}}x_{s1} + ... + \frac{1-\alpha_{11}}{1-\alpha_{nn}}x_{s1} \Rightarrow$

$1-D = x_{s1}(1-\alpha_{11})\sum_1^n \frac{1}{1-\alpha_{NN}}$

$x_{s1} = \frac{1-D}{1-\alpha_{11}} \frac{1}{\sum_1^n \frac{1}{1-\alpha_{NN}}}$

For $i \in [1, n]$:

$$x_{si} = \frac{1 - D}{1 - \alpha_{ii}} \frac{1}{\sum_1^n \frac{1}{1-\alpha_{NN}}} \tag{5.13}$$

For $i \in [n+1, 2n]$, similarly:

$$x_{si} = \frac{D}{1 - \alpha_{ii}} \frac{1}{\sum_{n+1}^{2n} \frac{1}{1-\alpha_{NN}}} \tag{5.14}$$

5.3.3 BP Message Propagation

The algorithm propagates messages through a sequence of segments of the factor graph, as described in Chapter 4. Each variable node $x_i \in \mathcal{X}$ is initialized at time instance k from messages arriving from the factor graph segment processing in time instance $k-1$ apart from the first initialization, using the Steady-State vector, $\mathbf{x}_{ss}$.

The first step is processing messages from the factor nodes f_{y_1} and f_{y_2} to the variable nodes $x_j \in \mathcal{X}_{\text{on}}$ and $x_j \in \mathcal{X}_{\text{off}}$ accordingly:

$$z_{f_{y_1} \to x_j} = \frac{1}{c_p} \left(z_{f_{y_1}} - c_p \sum_{x_b \in \mathcal{X}_{on,i} \setminus x_j} z_{x_b \to f_{y_1}} \right) \tag{5.15a}$$

$$\sigma^2_{f_{y_1} \to x_j} = \frac{1}{c_p^2} \left(\sigma^2_{f_{y_1}} - c_p^2 \sum_{x_b \in \mathcal{X}_{on,i} \setminus x_j} \sigma^2_{x_b \to f_{y_1}} \right) \tag{5.15b}$$

$$z_{f_{y_2}\to x_j} = \frac{1}{c_p}\left((c_p - z_{f_{y_1}}) - c_p \sum_{x_b \in \mathcal{X}_{off,i}\setminus x_j} z_{x_b \to f_{y_2}}\right) \tag{5.16a}$$

$$\sigma^2_{f_{y_2}\to x_j} = \frac{1}{c_p^2}\left(\sigma^2_{f_{y_2}} - c_p^2 \sum_{x_b \in \mathcal{X}_{off,i}\setminus x_j} \sigma^2_{x_b \to f_{y_2}}\right) \tag{5.16b}$$

where $\mathcal{X}_{on,i} \setminus x_j \subset \mathcal{X}_{\text{on}}$ is the set of variable nodes adjacent to the factor node f_{y_1}, excluding the variable node x_j (similarly for $\mathcal{X}_{off,i} \setminus x_j \subset \mathcal{X}_{\text{off}}$).

The second step is processing messages from variables nodes $x_j \in \mathcal{X}$ to factor nodes $f_i \in \mathcal{F}_{\text{t}}$:

$$z_{x_j\to f_i} = \left(\frac{z_{x_j}}{\sigma^2_{x_j}} + \frac{z_{f_y\to x_j}}{\sigma^2_{f_y\to x_j}}\right)\sigma^2_{x_j\to f_i} \tag{5.17a}$$

$$\sigma^2_{x_j\to f_i} = \frac{\sigma^2_{f_y\to x_j}\cdot\sigma^2_{x_j}}{\sigma^2_{f_y\to x_j} + \sigma^2_{x_j}} \tag{5.17b}$$

Lastly, messages are propagated from all factor nodes $f_i \in \mathcal{F}_{\text{t}}$ to the next time step $k+1$ using

$$z_{f_i\to x_i} = \alpha_{i,i-1}\cdot z_{x_{i-1}\to f_i} + \alpha_{i,i}\cdot z_{x_i\to f_i} + b_i\cdot u_i^k \tag{5.18a}$$

$$\sigma^2_{f_i\to x_i^{k+1}} = \sigma^2_{x_i} + \alpha^2_{i,i-1}\cdot\sigma^2_{x_{i-1}\to f_i} + \alpha^2_{i,i}\cdot\sigma^2_{x_i\to f_i} + b_i^2\cdot\sigma^2_{u_i} \tag{5.18b}$$

where $\alpha_{i,i}$, $\alpha_{i,i-1}$ the elements of the TPM $\boldsymbol{A}$.

The output, x_i^{k+1} states, is corrected through normalization over the sum $\Sigma_{i=1}^{N} x_i^{k+1}$, where N here is the summation variable.

5.4 Simulation Results

TPM comparison

The model of Chapter 4 (parameters listed in Table 4.1 [54]) is used to compare TPMs between Chapter 4 and Chapter 5. Tables 5.1, 5.2 show these TPMs according to Chapter 4 (Section 4.4) and Chapter 5 (Section 5.2) methods. Analytical calculation in the first case, as mentioned before, is limited to heterogeneity in one parameter(α). The difference between the 2 TPM is minimal, this is expected due to the choice of Gaussian and Uniform distributions for TCL parameters (Table 4.1). In case of different parameter distribution (e.g. Log-Normal distributions where the mode value is a better choice than the average value), the gap is wider. Note, that as stated in [54], the analytical calculation of the TPM requires assumption of Uniform distribution of parameter α, for the rest TCL parameters within the population average values are assumed.

Convergence comparison

Another change in this chapter is the FG (Figure 5.1) and its algorithm (algorithm 5.2). Factor node f_{y_2} adds extra information, which causes $x_{j \in X_{off}}$ to be updated faster and normalisation converges all states ($x_{j \in X}$) to fractions whose sum equals 1. A comparison between the previous model (Chapter 4) [121] and this one can be seen between Figures 4.13 and 5.3. Both are given a constant measurement (equal to the steady state in this case) and have the same initialization (for comparison purposes) as per Chapter 4. Note, the initialization here is the same only for comparison purposes and for the remaining simulations of this chapter they will be as described in the previous section.

Table 5.1: TPM according to Section 4.4

	X_{off}				X_{on}			
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	0.7526	0	0	0	0	0	0	0.3130
x_2	0.2474	0.7577	0	0	0	0	0	0
x_3	0	0.2423	0.7627	0	0	0	0	0
x_4	0	0	0.2373	0.7678	0	0	0	0
x_5	0	0	0	0.2322	0.6718	0	0	0
x_6	0	0	0	0	0.3282	0.6769	0	0
x_7	0	0	0	0	0	0.3231	0.6819	0
x_8	0	0	0	0	0	0	0.3181	0.6870

Heterogeneity can be assumed on $a = e^{-\tau/RC}$ only, fixed uniform distribution

Rounding at 4 decimals(per MatLab). Columns may sum to 0.9999 or 1.0001

From the very first iteration the new model gives an output close to the steady state and converges about 3 times faster for $x_i \in \mathcal{X}_{\text{off}}$ and about twice as fast for $x_i \in \mathcal{X}_{\text{on}}$. This shows the importance of tracking not only *on* units, as commonly suggested in literature ([29, 40, 54, 55]) but also off units. The steady states are identical, as seen in Table 5.3.

NB.: If (5.13), (5.14) are applied, according to transition probabilities of Tables 5.1 and 5.2, the resulting steady states are practically the same as in Table 5.3.

Table 5.2: TPM according to Section 5.2

	X_{off}				X_{on}			
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	0.7507	0	0	0	0	0	0	0.3148
x_2	0.2493	0.7557	0	0	0	0	0	0
x_3	0	0.2443	0.7607	0	0	0	0	0
x_4	0	0	0.2393	0.7658	0	0	0	0
x_5	0	0	0	0.2342	0.6701	0	0	0
x_6	0	0	0	0	0.3299	0.6751	0	0
x_7	0	0	0	0	0	0.3249	0.6801	0
x_8	0	0	0	0	0	0	0.3199	0.6852

Heterogeneity TCL parameters of table 4.1, Gaussian & Uniform distributions Rounding at 4 decimals(per MatLab). Columns may sum to 0.9999 or 1.0001

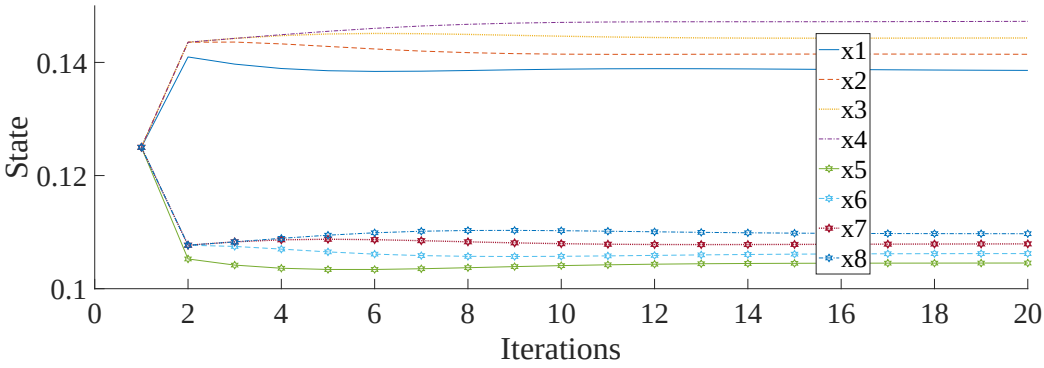


Figure 5.3: New factor graph and algorithm, state estimation for "steady" state

State space models performance comparison

Using the TCL parameters listed in Table 4.1 (high heterogeneity), Mathieu et al. [54] simulated a population of 10,000 TCLs and randomly drew each

Table 5.3: BP Steady State, Chapter 4 and 5 comparison

	BP (Chapter 5)	BP (Chapter 4)
x_1	0.13880	0.13857
x_2	0.14147	0.14144
x_3	0.14428	0.14444
x_4	0.14719	0.14757
x_5	0.10449	0.10447
x_6	0.10610	0.10608
x_7	0.10783	0.10778
x_8	0.10982	0.10965

Chapter 5 BP at 14 iterations (Chapter 4 BP at 44, Table 4.2), rounded at the 5 decimals

parameter from the uncorrelated uniform distributions between the minimum and maximum values. Information of aggregate power demand was used for state estimation. The performance of the state space model to predict the aggregate response to control signals was evaluated using the RMS (Root Mean Square) error for the first 5 minutes of the simulation following a control action. In more detail (quoting Mathieu et al. [54]): "Two randomly generated open-loop control sequences (drawn from uniform distributions) were used to force the population: high forcing, in which up to 12.5% of the TCLs were switched in one time step, and low forcing, in which up to 2.5% of the TCLs were switched in one time step. The control was applied such that TCLs in bins nearer to the dead-band were switched preferentially. We then evaluated the ability of the model to predict the aggregate power consumption of the plant. We assume that the model knows the state perfectly when the prediction horizon is zero and gains no additional state information over time." The RMS prediction error was normalized by the steady state power consumption of the TCL population, for different prediction horizons and plotted. The reported RMSE for $N_{bin} = 10$

for high forcing (12.5% of the TCLs) was approximately up to 80% and for low forcing (2.5% of the TCLs) was approximately up to 18%. Unfortunately no further information or exact numbers were given. For more details the readers are prompted to read [54].

For comparison purposes, the same procedure to calculate the RMS error will be used. The RMS error is calculated using the standard formula:

$$RMSE = \sqrt{\frac{\Sigma_{i=1}^{n}(\mathbf{y_{1,i}} - N_{on,i})^2}{n}} \tag{5.19}$$

where n the prediction horizon (in minutes) starting at the minute the control signal is sent ($i = 0$), as done in [54]. N_{on} is given by the MC simulation with actual number of units *on*. Then the RMS error is normalized by the steady state number of TCL units *on* ($D \cdot N_{TCL}$). Note that simulations of Chapter 4 the RMS error is calculated using $\mathbf{y_i}$ instead of $\mathbf{y_{1,i}}$.

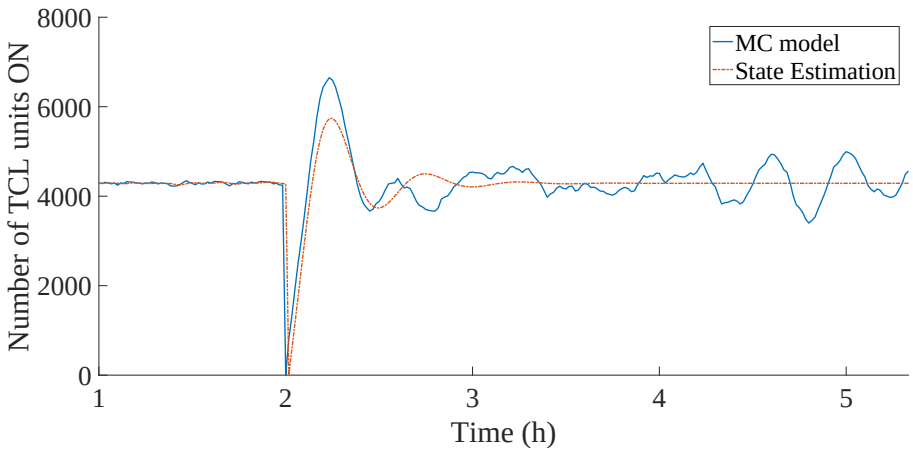


Figure 5.4: Chapter 4 model performance, without state estimation after signal, lower heterogeneity

At time $t = 120min$ a short switch off command is given to all TCL units. At that point there is **no further state information over time** (as in [54]).

Results can be seen in Figures 5.4 - 5.9, where the RMS error of Figures 5.4, 5.5 are shown in Figures 5.6, 5.7.

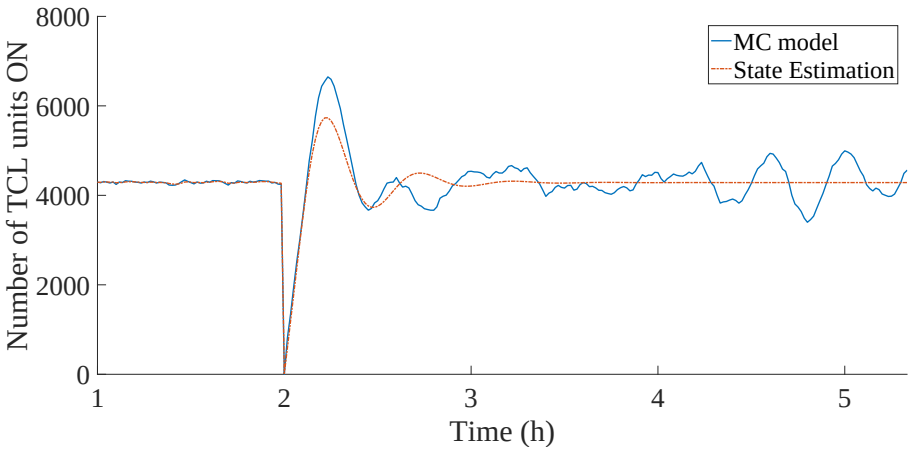


Figure 5.5: Chapter 5 model performance, without state estimation after signal, lower heterogeneity

The RMS error of Figures 5.4, 5.5 are very similar in shape and numbers, thus excluded here but included in Table 5.4. The prediction horizon of up to 120 minutes post signal was chosen since there is little change in the RMS error afterwards, which is expected when observing the State Estimation in Figures 5.4, 5.5 (or Figures 5.8, 5.9).

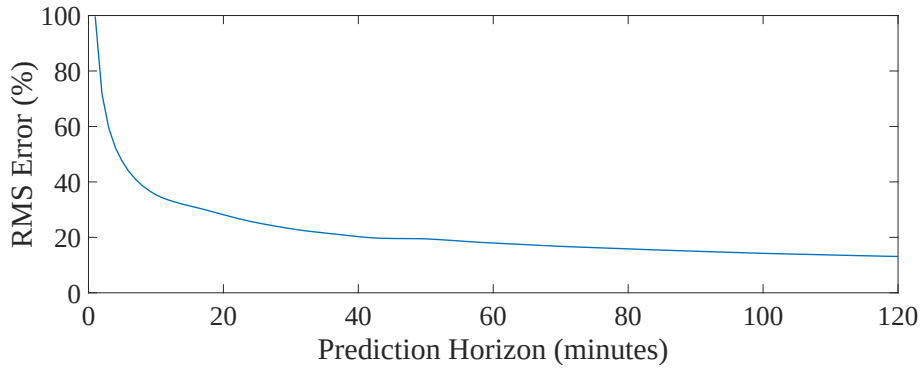


Figure 5.6: RMS error of Figure 5.4, for prediction horizons up to 120 minutes.

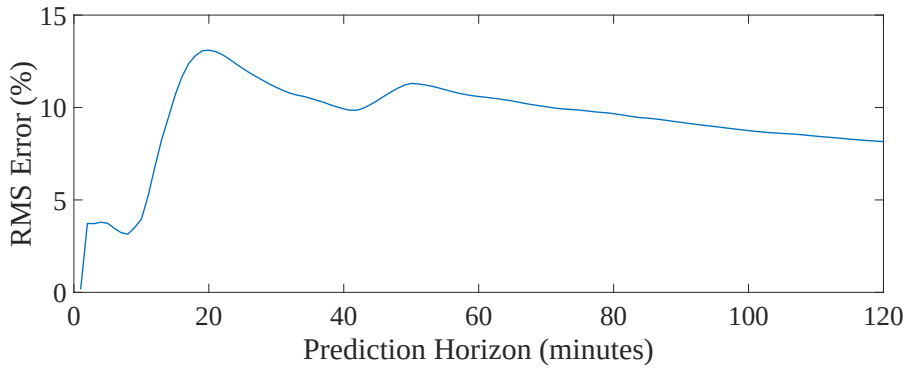


Figure 5.7: RMS error of Figure 5.5,for prediction horizons up to 120 minutes.

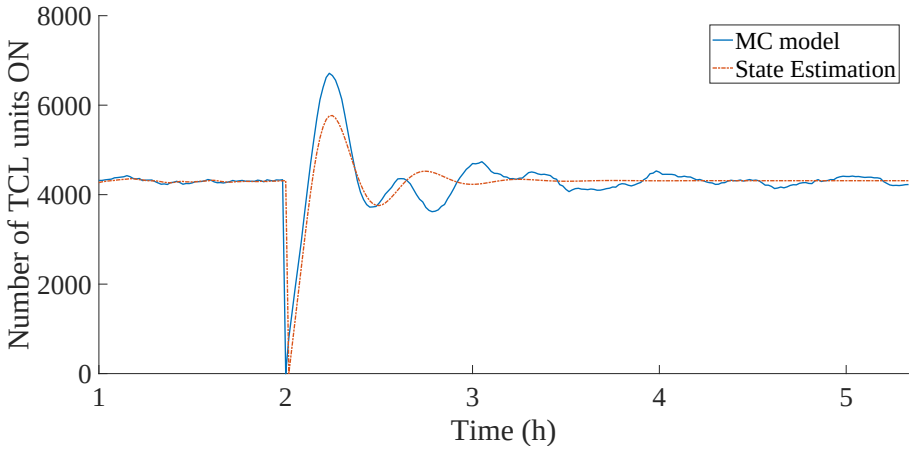


Figure 5.8: Chapter 4 model performance, without state estimation after signal, higher heterogeneity

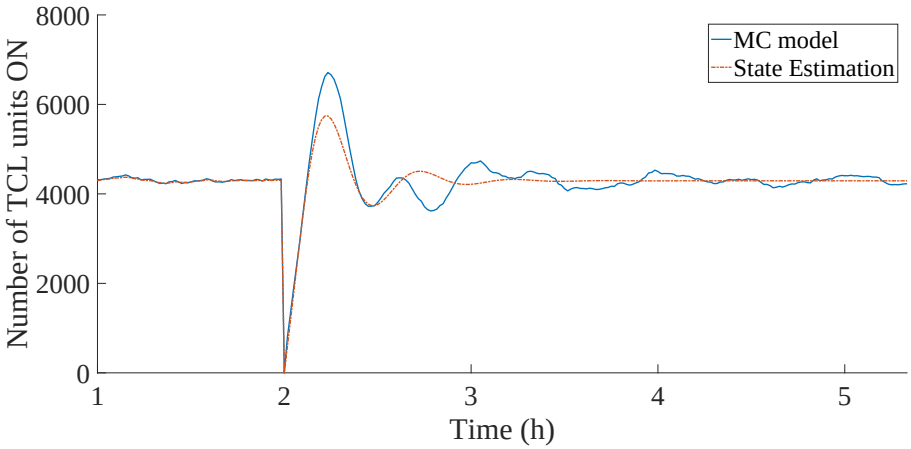


Figure 5.9: Chapter 5 model performance, without state estimation after signal, higher heterogeneity

The results show an improvement compared to the method used in Chapter 4 and in [54]. Also, it is apparent that the state space model itself is not enough and State Estimation is crucial for accuracy.

Simulation results with same control commands as in Chapter 4

The same control commands and same TCL populations as in Section 4.8.2 are used for comparison purposes.

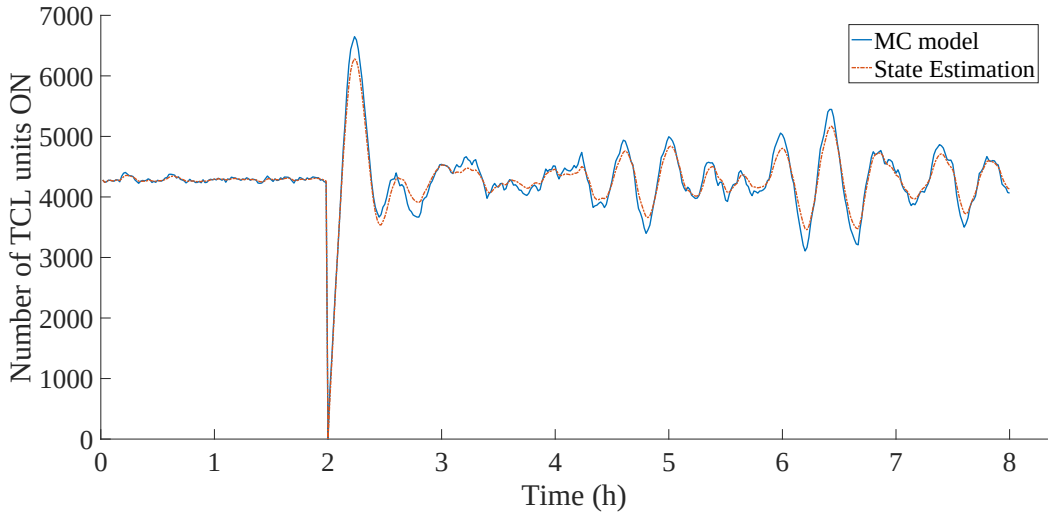


Figure 5.10: 10,000 TCLs, according to Table 4.1, lower heterogeneity, switch off signal at time 120' for whole population, short duration

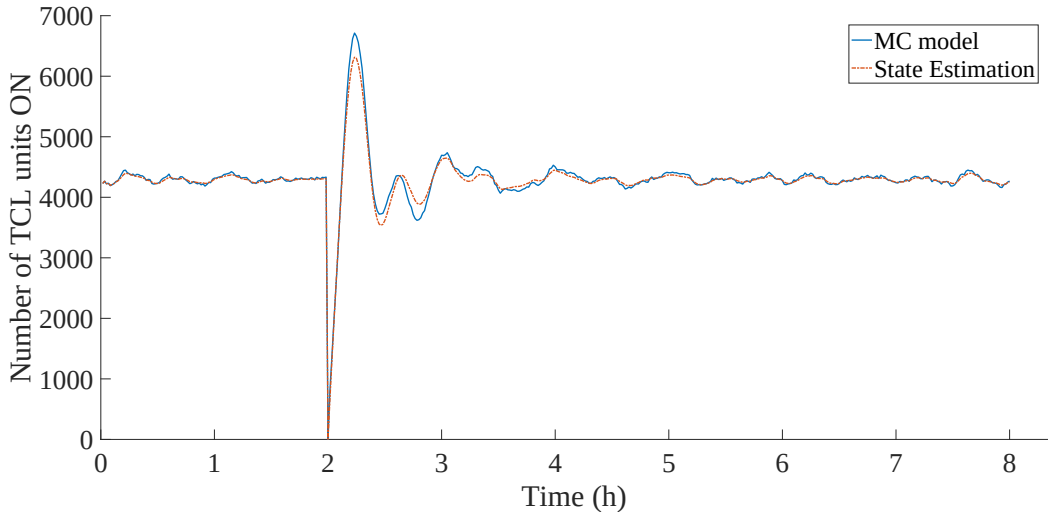


Figure 5.11: 10,000 TCLs, according to Table 4.1, higher heterogeneity, switch off signal at time 120' for whole population, short duration

Compared to the method of Chapter 4 there is an improvement in the first few steps after the control command is send, which results in more accuracy thereafter as well.

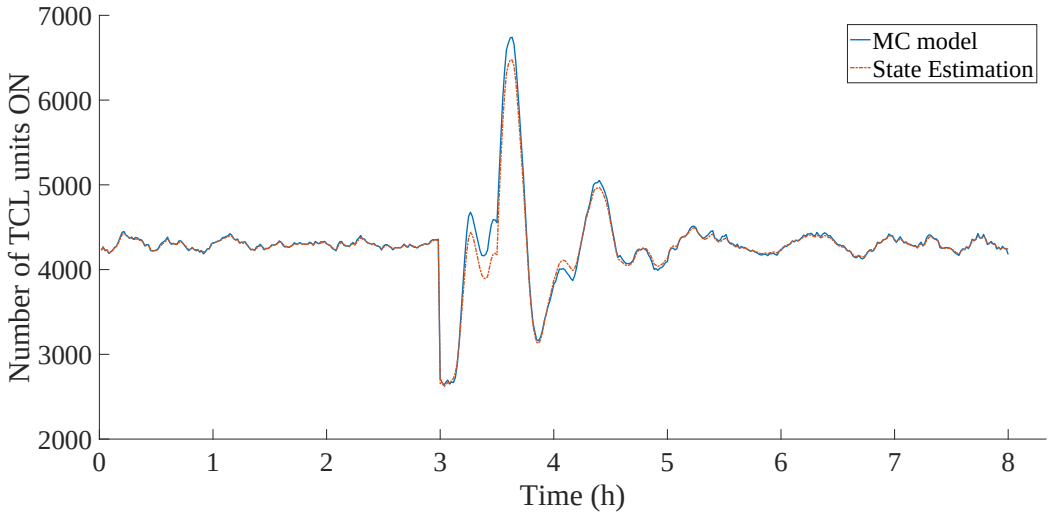


Figure 5.12: 10,000 TCLs, higher heterogeneity (according to Table 4.1), switch off signal at time 180' for 40% of units on, 10' duration.

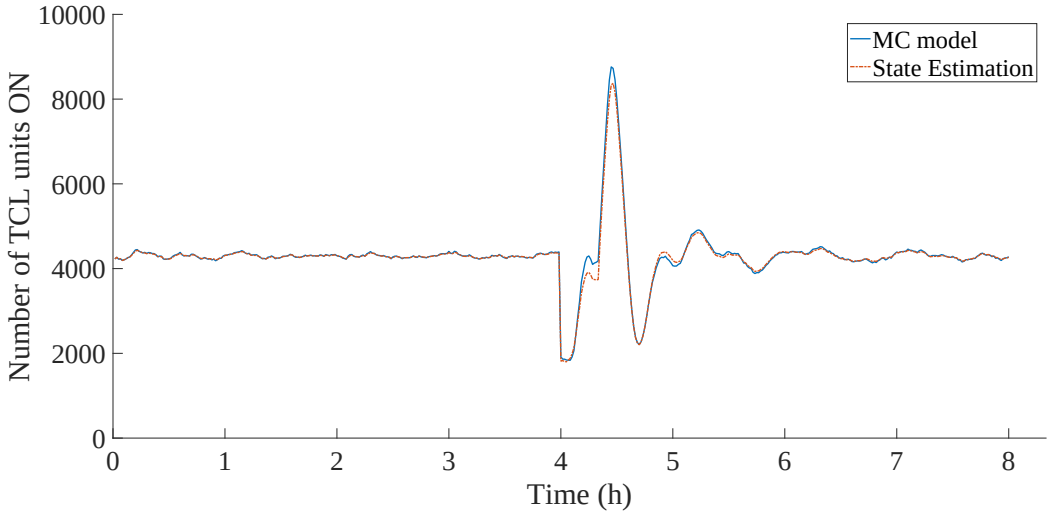


Figure 5.13: 10,000 TCLs, higher heterogeneity (according to Table 4.1), switch off signal at time 240' for 60% of units on, 6' duration.

In case of steps around 1 second or less (based on expected time for the command signal to be send and the time for TCLs to respond), the state space model should change where control actions in step k ($\mathbf{u}(k)$) are observed in step $k+1$ ($\mathbf{x}(k+1)$).

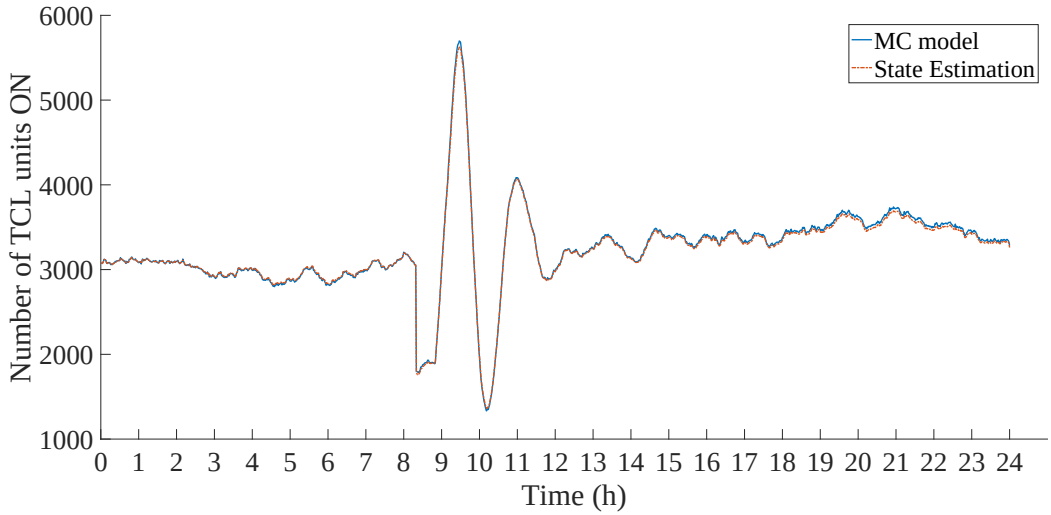


Figure 5.14: 10,000 TCLs of Chapter 3, switch off signal at time 500' for 40% of units on, 30' duration.

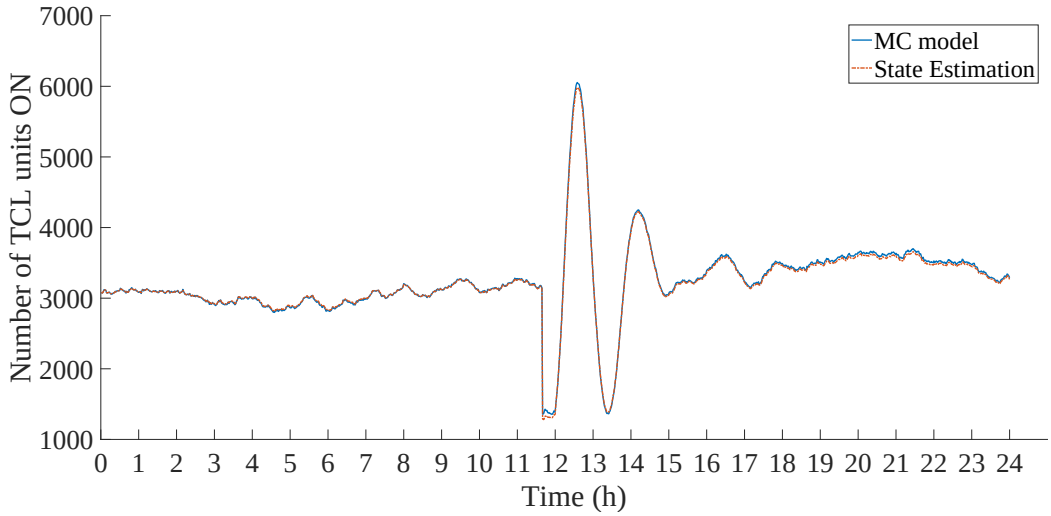


Figure 5.15: 10,000 TCLs of Chapter 3, switch off signal at time 700' for 60% of units on, 20' duration.

Finally, Table 5.4 shows the RMS error of simulations in Chapter 4 and 5. It is hard to have a direct comparison with [54] due to lack of data, yet the closest comparison would be the high forcing (%12.5 of TCL population) with the

simulation of 40% with a duration of 10'. In which case both methods described in Chapter 4 and 5 have better results. Overall the results of the method described in Chapter 5 has the lowest RMS error between all. The larger the percentage of the TCL population used to switch state the greater the RMS error as expected. It is hard to comment how much would be used in real world applications without the appropriate data.

Table 5.4: RMS errors of simulations in Chapters 4 and 5

Figures	Chapter 4(max)	Chapter 4(120')	Chapter 5(max)	Chapter 5(120')
5.4, 5.5	99.7	13.1	13.1	8.2
5.8, 5.9	100.6	13.4	14.1	8.5
4.14, 5.10	99.7	10.7	5.3	3.4
4.15, 5.11	100.6	10.7	5.7	3.4
4.16, 5.12	37.7	8	5.5	3.4
4.17, 5.13	57.3	9.9	6.8	3.7
4.18, 5.14	40.9	7.8	1.3	1.1
4.19, 5.15	58.7	9	2.9	1.5

5.5 Online Aggregated Parameter Identification

As investigated in Chapter 3, loads vary during the day due to various external factors. Therefore, it is crucial for aggregation models to be able to keep up with those changes and update in real-time if possible. El-Férik & Malhamé [98] proposed an identification algorithm to calculate and update the parameters (C, R, P, θ etc.) of their proposed CFPE model, so that it could cope with real world changes and concluded that sampling intervals of 15.198 minutes or less is required. Trovato et al. [22], Vinther et al. [39] and Totu et al. [34] in their respective works have also reported similar conclusions, where precise sensors for tracking of current temperature and system parameters are required for all loads.

On the other hand, Koch et al.'s work [29] suggests a model with reduced sensoring requirements, where sampling of the state of a percentage of the population is only needed. Still for state (temperature for the case of TCLs) thermal sensors are needed even if not as precise. Mathieu et al. [54] build on that work and investigate a joint parameter/state estimation through EKF to further reduce sensor requirements but without success.

A method to derive the TPM is introduced which requires no thermal sensors, but only using knowledge of power on state (duration), as it is provided by smart meters, smart appliances or even from disaggregation techniques as described earlier (variation of this method). The assumptions are the following:

- Power *on* is known in time (i.e. from smart meter). Thus t_{on} and t_{off} are known.
- Set-point is known ($\theta_{set}, \theta_{on}, \theta_{off}$). This does not require thermal sensors.
- Either θ_g is known or external temperature θ_a is estimated (can be known for indoors, from the set-point of other devices)

A more generic calculation of average rates over the whole period of each state is (using the notations $\rho_{on,aver}$, $\rho_{off,aver}$ to avoid confusion):

$$\rho_{off,aver} = -\frac{\theta_{off} - \theta_{on}}{t_{off}} \tag{5.20a}$$

$$\rho_{on,aver} = -\frac{\theta_{on} - \theta_{off}}{t_{on}} \tag{5.20b}$$

NB: An important aspect is that they require less information and that external factors are actually included. Since the average rates depend on t_{on}, t_{off}, any effect of external factors and/or variations is factored in.

Consider for example Figure 4.3. Both TCLs have the exact same $\rho_{on,aver}$, $\rho_{off,aver}$ (and $\rho_{on,mean}$, $\rho_{off,mean}$ obviously), even though they clearly have different TCL

parameters. Using equations (5.2c) and (5.3c) the values λ and θ_g (or θ_a) can be calculated. The calculated values **do not** reflect the actual parameters of the device, but rather the ones that the system should have to account for noise, temperature changes, external factors etc. This holds true especially for $\lambda(t)$ which adjusts as cycles change, but that actually makes the system robust (to avoid confusion notations λ' and θ'_g will be used, as in the parameters of the TCL without noise of Figure 4.3). Solving (5.2c) and (5.3c) for λ' and θ'_g gives:

$$\lambda' = -\frac{\rho_{off,mean}}{\frac{\theta_{on}+\theta_{off}}{2} - \bar{\theta_a}} \tag{5.21}$$

$$\theta'_g = \frac{\theta_{on}+\theta_{off}}{2} - \bar{\theta_a} + \frac{\rho_{on,mean}}{\lambda'} \tag{5.22}$$

Alternatively, (4.2) can be used to calculate λ' (λ' replacing λ_H in the equation) and (3.46) to calculate θ'_g (solving for θ_g and using λ' value in place of λ_H). For a large population of TCLs, the error of this assumption will be minimal. This is easily observed in Figure 4.3 where such disturbances are distributed randomly since the TCLs are not synchronized. Thus, in some TCLs the bin estimation will have an error towards one direction and in some other TCLs towards the other.

Given a heterogeneous population, various readings of P_{on} and subsequently t_{on} and t_{off} measurements are expected. The first step is to determine the distribution of those and process them in order to create the TPM. It might also be favourable to cluster them when deemed appropriate, independently of the load type. The readings reflect recent data but not exactly real-time data, as the t_{on}, t_{off} are the ones from states that have just ended, yet the aggregated state should not change significantly.

5.5.1 Case Studies

The proposed method's accuracy was tested using the simulated MC model employed before with a population of 10,000 TCLs (Table 4.1), whose TPM is already known as a basis for comparison (Tables 5.1, 5.2). A sampling every 15 minutes is taken, where duration of power *on* (t_{on}) and power off states (t_{off}) is measured, as it could be done in real life (note that Smart Plugs track power state every second and these updates can be done in smaller intervals than 15 minutes). Some of those can be seen in Figures 5.16, 5.18, 5.17 and 5.19, where for blocks of 15 minutes the duration of t_{on} and t_{off} of TCL units is tracked.

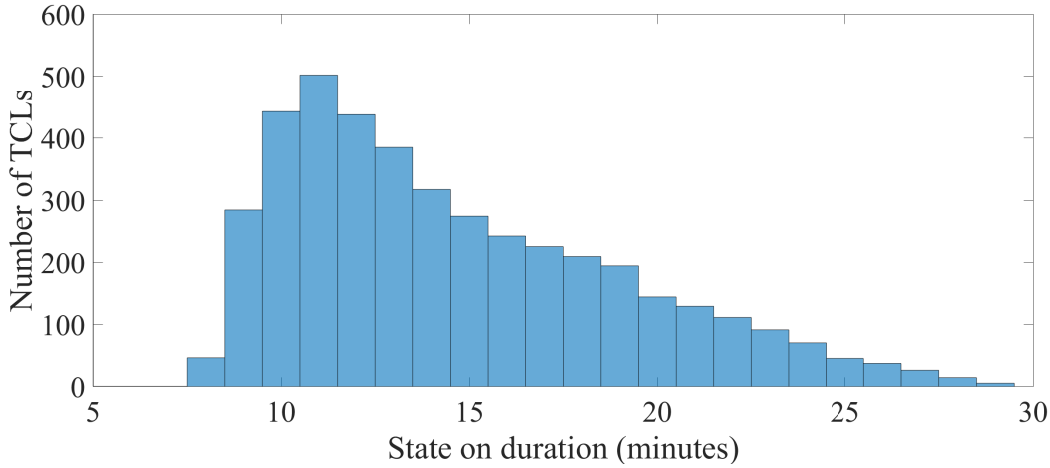


Figure 5.16: Distribution of power *on* state duration t_{on}. Sampling between simulation time 30' and 45'.

In these figures the distribution (histogram) can be seen, where for instance in Figure 5.16, between simulation time 30' and 45', an approximate number of 300 TCLs had t_{on} equal to 15 minutes and a total number of units *on* is above 4000, as expected when looking at Figure 5.12 in simulation time 45'. Additionally, knowledge of set-point (θ_{off}, θ_{on}) and estimation of ambient temperature (θ_a with noise, e.g. from the settings of heating or an air-conditioner, which requires

no thermostat) are considered. Examining the distribution of Figures 5.16, 5.18, 5.17 and 5.19 it is apparent that the Log-Normal distribution is a good fit. Now using the t_{on} and t_{off} information we will calculate the TPM as described in the previous section.

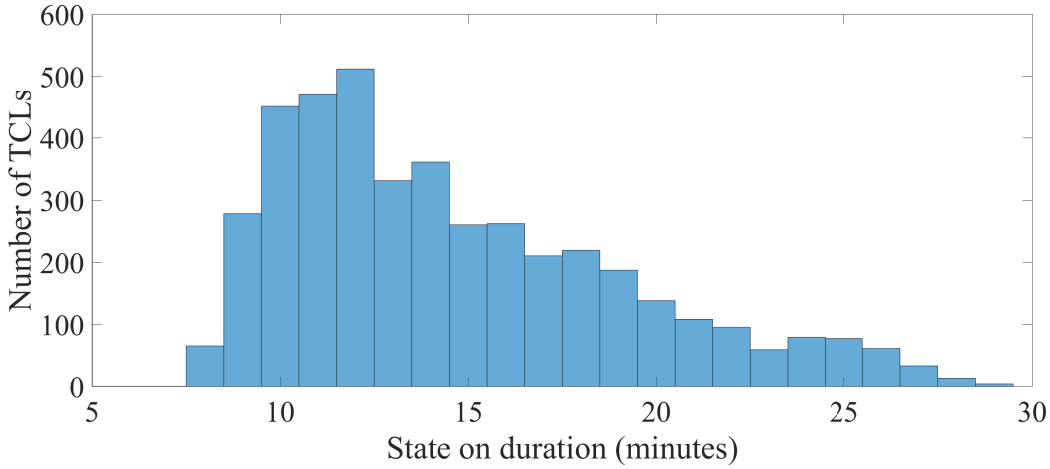


Figure 5.17: Distribution of power on state duration t_{on}. Sampling between simulation time 255' and 270'.

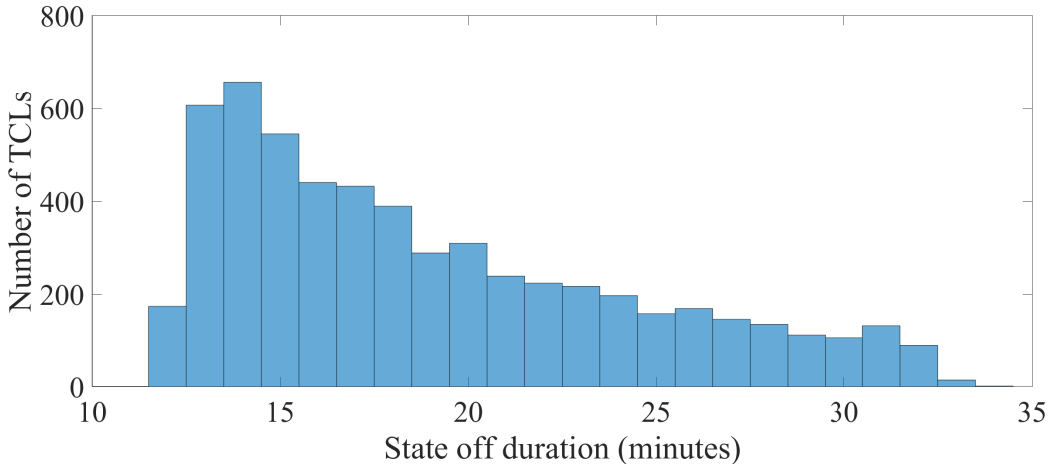


Figure 5.18: Distribution of power off state duration t_{on}. Sampling between simulation time 30' and 45'.

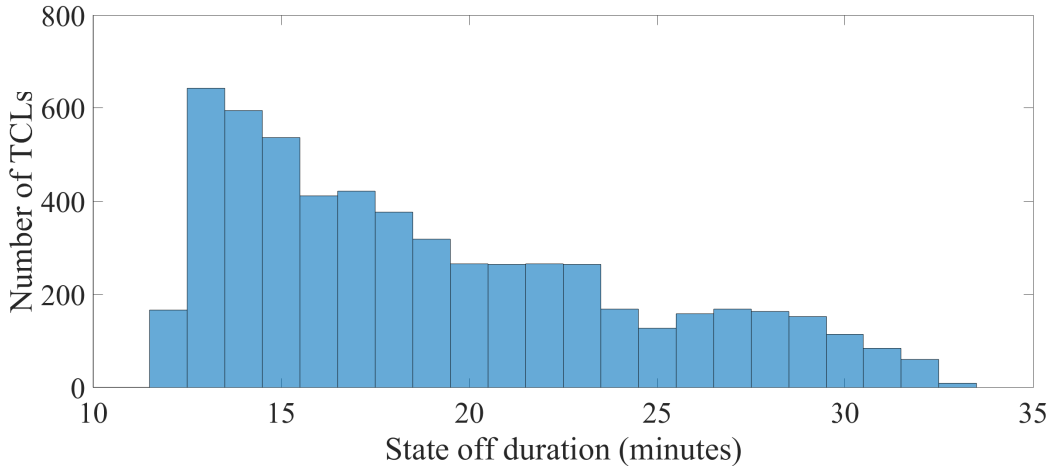


Figure 5.19: Distribution of power off state duration t_{on}. Sampling between simulation time 255' and 270'.

A few cases will be examined on representative t_{on} , t_{off} values to be used in order to derive the TPM closest to the ones of Table 5.1 and Table 5.2. According to the data observed (part of it seen in Figures 5.16, 5.18, 5.17 and 5.19), these cases are:

- Case 1: First and simplest case, use the mean values of t_{on}, t_{off} to calculate the TPM (would fit well Gaussian distributions).
- Case 2: Secondly, use the median values of t_{on}, t_{off} to calculate the TPM.
- Case 3: Lastly, given the observation that the distribution of t_{on}, t_{off} values follows a Log-normal distribution trend, examine between the mean $(\exp\left(\mu + \frac{\sigma^2}{2}\right))$, median $(\exp(\mu))$ and mode $(\exp(\mu - \sigma^2))$. A note here is that even though the input parameters (Table 4.1) do not follow a Log-normal distribution, the resulting t_{on}, t_{off} do. This is easily explained by looking at (3.46), (3.48) of Chapter 3.

The resulting TPMs for each of these cases are shown below. For the case of

Log-Normal distribution the best result was given using the mode and the others are not shown.

Table 5.5: TPM, transition from column to row - case 1

	X_{off}				X_{on}			
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	0.7831	0	0	0	0	0	0	0.2666
x_2	0.2169	0.7875	0	0	0	0	0	0
x_3	0	0.2125	0.7918	0	0	0	0	0
x_4	0	0	0.2082	0.7962	0	0	0	0
x_5	0	0	0	0.2038	0.7203	0	0	0
x_6	0	0	0	0	0.2797	0.7247	0	0
x_7	0	0	0	0	0	0.2753	0.7290	0
x_8	0	0	0	0	0	0	0.2710	0.7334

Rounding at 4 decimals(per MatLab). Columns may sum to 0.9999 or 1.0001

Table 5.6: TPM, transition from column to row - case 2

	X_{off}				X_{on}			
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	0.7708	0	0	0	0	0	0	0.2788
x_2	0.2292	0.7755	0	0	0	0	0	0
x_3	0	0.2245	0.7801	0	0	0	0	0
x_4	0	0	0.2199	0.7847	0	0	0	0
x_5	0	0	0	0.2153	0.7073	0	0	0
x_6	0	0	0	0	0.2927	0.7120	0	0
x_7	0	0	0	0	0	0.2880	0.7166	0
x_8	0	0	0	0	0	0	0.2834	0.7212

Rounding at 4 decimals(per MatLab). Columns may sum to 0.9999 or 1.0001

Table 5.7: TPM, transition from column to row - case 3

	X_{off}				X_{on}			
	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8
x_1	0.7576	0	0	0	0	0	0	0.3035
x_2	0.2424	0.7625	0	0	0	0	0	0
x_3	0	0.2375	0.7674	0	0	0	0	0
x_4	0	0	0.2326	0.7723	0	0	0	0
x_5	0	0	0	0.2277	0.6818	0	0	0
x_6	0	0	0	0	0.3182	0.6867	0	0
x_7	0	0	0	0	0	0.3133	0.6916	0
x_8	0	0	0	0	0	0	0.3084	0.6965

Rounding at 4 decimals(per MatLab). Columns may sum to 0.9999 or 1.0001

As it is easily observed the Log-Normal distribution with mode gives the closest result to the base TPM, which is expected. The ones corresponding to *on* states have an error of approximately 1-2% compared to Tables 5.1 and 5.2. This showcases that accurate enough TPM can be derived via t_{on} and t_{off} and limited information, without the need of high precision thermal sensors or other expensive equipment, in contrast to other methods (e.g. [54, 22, 34, 39]). The duty cycle in steady state of these TPMs can be obtained as described in Section 5.4.

5.5.2 Switches Tracking

The method described above can be used when the aggregated population is in a steady state; steady state here is defined as having no considerable changes on aggregated level in a short period of time (i.e. permitting small noise). Of course that is not the case all the time, as analysed in detail in Chapter 3. As such it is important to include some method of tracking it.

Equation (5.10) describes the condition of a steady state, since the number (or fraction of the population) of units that switch from *on* state to off state have to be equal to the ones that switch the other way around, otherwise there would be no steady state. In reality, this will be the case, but with a small level of noise. The parameter z can be defined as the fraction of switches between states (and state bins).

$$(1-\alpha_{11})x_{s1} = (1-\alpha_{22})x_{s2} = ... = (1-\alpha_{2n2n})x_{s2n} \tag{5.23}$$

The number of *on* switches and off switches can be tracked via Smart Meters and Smart Plugs (or even disaggregation on feeder level [41, 42, 94] though not probable as discussed earlier). In one time step h, the average (expected) number of *on* and off switches are equal to the fraction of h to t_{on}, t_{off} of the population within each corresponding state:

$$sw_{off}(h) = \frac{h}{t_{off}} N_{off} = \frac{h}{t_{off}}(1-D)\cdot N \tag{5.24a}$$

$$sw_{on}(h) = \frac{h}{t_{on}} N_{on} = \frac{h}{t_{on}}(D)\cdot N \tag{5.24b}$$

Using $D = \frac{t_{on}}{t_{on}+t_{off}}$, $1-D = \frac{t_{off}}{t_{on}+t_{off}}$, it easy to show that $w_{off}(h) = sw_{on}(h)$ during steady state, as expected. Expressing (5.24a) and (5.24b) as fractions (divided by N), they are equal to $(1-\alpha_{nn})x_{s,n}$ (off switches) and $(1-\alpha_{2n2n})x_{s,2n}$ (*on* switches) respectively. When (5.24a) and (5.24b) are (almost) equal, steady state can be assumed.

Another interesting application of tracking switches is identifying the TPM which describes the aggregated behaviour. One simple way of doing so, is by using the equations (5.24a) and (5.24b) and having knowledge of the most recent duty cycle $D(t)$ (i.e. 15minutes intervals as described in the method above) to calculate t_{on} and t_{off} in almost real-time. Another way of doing so is by estimating the duty cycle in real-time. Given the average demand per unit P_{aver} and the current

aggregated demand P_{tot}:

$$P_{tot} \approx D \cdot N \cdot P_{aver} \leftrightarrow D \cdot N \approx \frac{P_{tot}}{P_{aver}} \tag{5.25}$$

Replacing (5.25) into (5.24a) and (5.24b):

$$t_{off} \approx \frac{P_{tot}}{P_{aver}} \cdot \frac{h}{sw_{off}} \tag{5.26a}$$

$$t_{on} \approx \frac{P_{tot}}{P_{aver}} \cdot \frac{h}{sw_{on}} \tag{5.26b}$$

5.6 Conclusions

This chapter introduces a new method to model aggregated heterogeneous populations of TCLs and loads with similar operational characteristics. TPM based models are used in conjunction with *on/off* rates, referring to the operational behaviour and a methodology to analytically derive the TPM is presented. Afterwards, the equivalent FG is described as well as the BP algorithm which runs on top. The results are compared to existing models and the one used in the previous chapter. Finally, simulation results and comparison to the existing models and previous work are illustrated, which highlight the advantages of the proposed framework. These can be summarised as:

1. A new methodology, to derive the TPM, modelling heterogeneity in multiple parameters in a simple manner, as well being external factors.
2. Can be used universally for various TCLs types which have similar TPMs (or even other loads with similar TPMs).
3. A more appropriate state space model is suggested and its equivalent factor graph, resulting in higher accuracy on predicting new time steps and control

actions. The BP state estimation algorithm is also updated facilitating faster convergence and higher accuracy.

4. Methodologies for online robust updating have been proposed, without requirements of (precise) thermal sensor, but only power measurements from smart meters or disaggregation on feeder / substation level.
5. A simple metric to identify steady state was also created.

Chapter 6

Thesis Conclusions

6.1 Thesis Summary

This thesis made a contribution in the four key-aspects of the DR via TCLs (and similar loads with storage capabilities): dynamic behaviour modelling, heterogeneous aggregation, state estimation of large populations and dynamic tracking. The second Chapter examines the DR potential of loads, mainly residential but can be extended to commercial ones as well, under different types of DR services. DR balancing services are crucial for the grid's stability, especially under high intermittent RES penetration; later Chapters of this thesis focus on TCLs in particular as loads with the highest potential for such DR given that they are available 24/7 and have high demand throughout the day, thus high reliability as there is always a significant amount that be used for DR. The third Chapter deals with the derivation of dynamic thermal models to describe aggregated behaviour of TCLs. Different types of cold loads are modelled via Monte Carlo with high level of accuracy with regard to the scope of this work. The fourth Chapter introduces a distributed state estimation algorithm, Belief Propagation,

over Factor Graphs and in the fifth Chapter a novel aggregation model is combined with BP and online tracking and updating algorithms. It is concluded that a large population of heterogeneous TCLs can be accurately modelled over different time scales and without the use of expensive and sophisticated sensors or monitoring equipment; This is based on approximate (with noise) knowledge of power states which can infer relative thermal states, meaning reduced CAPEX and OPEX for DR services.

Reducing infrastructure and communication cost helps to minimize total system cost, but also when distributed method are considered privacy concerns are minimized [123, 124]. Instead of transmitting usage data – whether real-time or not – to a central controller, distributed models can be used which exchange between them aggregated data and afterwards it is send to the VPP/aggregator. Alternatively, anonymous or randomised data might be possible to be utilized, such via randomized IDs assigned every few hours or so, though such an case has not been studied here for its validity and this only a suggestion. Disaggregation on substation level, as suggested in literature, would have to be done by System Operators and then data would have to be sent to VPPs/aggregators, though that would have various implications, besides privacy concerns (in such case **all** connected end users would have to give consent). For instance, when in the same substation, there are clients who have contracted to more than one VPP/aggregator, differentiating the information to send to each VPP/aggregator is impossible.

State estimation is imperative in minimizing the errors and state space models can be used for stochastic control. Probabilistic signals can be sent to (part of) the cluster and then decision is taken locally based on the power state of the TCLs.

6.2 Potential Implications

6.2.1 Dynamic Load Modelling

Data acquisition of (aggregated) demand profiles on a daily basis, when linked with affecting factors (such as weather conditions, time, day of the week, season, holidays, etc.), can improve load forecasting. This, combined with almost real-time aggregated demand readings can greatly increase intra-day forecasting. The outcome is minimization of errors in day-ahead and intra-day markets, thus reduced dispatch and re-dispatch costs.

6.2.2 Proposed DR Framework

There are a few key points, which can be seen as requirements for successful deployment of residential DR, as the one investigated in this thesis:

- Viability is probably the foremost requirement, which means minimizing cost where possible; Precise thermal sensors, high sampling frequency measuring devices, sophisticated sensoring and monitor equipment etc., can be a feasible but not viable solution.
- It is not only about minimizing costs, but also about privacy and security concerns associated with how advanced metering and consumers' data are used. Transmitting consumers' data centrally could cause lower levels of participation in DR. Realistic solutions must aim on anonymous or limited data or intervals of readings, which means utilizing only a fraction of the communication and computation capabilities available in a Smart Grid [125, 126].

- Yet, control strategies must be robust, with tracking and real-time evaluation DR availability. Aggregators/VPPs require feedback or closed loop control for accuracy.

The statistical modelling and on-line aggregated parameter identification allow for minimal metering infrastructure, relying on power measurements and not precise thermal sensors which are costly. The framework proposed allows for flexible design from central to distributed designs, whilst hierarchical ones seems to be a good niche. As such, the aggregated data reaching the central level (VPP/aggregator) is practically anonymous, whilst being sufficient. High-tech measurement equipment (i.e. frequency readings) on local level are not required, since the signal can be broadcast centrally, yet decision is taken locally on unit level.

6.2.3 Large Scale DR Implications

The analysis of the residential demand in flexible, deferrable and base demand gives an insight on the potential for both balancing services and dynamic pricing. DR is but the utilization of already existing resources within the system, which are ubiquitous and large in numbers, available 24/7, thus highly reliable and do not require fuel.

As such, DR can benefit Power Systems in various ways, either by sculpting the daily demand profile in favourable system conditions or providing the ever more needed balancing services. The DR in such cases can be seen as "borrowed" energy. Loads have to provide some specific services which require a set amount of energy within some time and constraints. Power ratings can vary within those, as long as the energy and service is delivered; i.e. a kettle's power rating might be held a) constant or b) vary 5% higher and 5% lower on equal time scales, the

thermostat will activate when the same total thermal energy is delivered in both cases. Varying though the total demand by 5% when a generation in-feed loss occurs could determine whether the system maintains stability or not. To that extend DR's benefits are:

- Reduced cost since there is no fuel consumption, compared to spinning reserves which tend to be expensive (running below rated value and fast ramping engines).
- Reduced emissions accordingly, by avoiding use of conventional spinning reserves, running in inefficient states.
- RES absorption through negative reserves (e.g. Demand Turn Up in UK), further assisting in cheap and clean energy.
- Dynamic pricing to regulate demand and supply, such as for peak demand reduction, reducing generation dispatch costs.
- Enhanced power system stability via frequency services (e.g. Firm Frequency Response, Enhanced Frequency Response in UK).
- Extra system flexibility for system operators, such as in case of congestion management.

It is important to note that the above are already provided in various countries and adding residential DR will further enhance these capabilities. Studies have shown that significant cost savings can be achieved by adding an extra control parameter. For the UK the potential was estimated up to £813 million annually in under a Gone Green scenario (2030) from the use of TCLs for DR whilst also greatly reducing greenhouse gases emissions [127].

6.2.4 Aggregators

There are a few extra benefits of aggregators utilizing DR, arising from demand monitoring, which applies mostly to industrial and commercial units and not to residential ones:

- Power charges management (i.e. Triad management in UK). Triad periods are defined as half hourly periods of peak consumption, occurring between November and February each year, with at least 10 days between each triad period. Businesses are charged according to their consumption during those periods. VPPs monitor and reduce consumption during those periods.
- Energy savings (energy efficiency), by monitoring usage and demand of loads, as well as optimising consumption where possible.
- Abnormal operation / fault detection, using historic load data.

6.3 Limitations

The proposed models in this thesis are designed for DR for large scale populations, using statistical models, state estimation and with specific data requirements. They were based on specific assumptions and as such, have the following limitations:

- First and foremost, as in every statistical model, stochasticity produces some error. In the case studies investigated in this thesis, it was shown that for populations below a few thousands (2,000 approximately), the error can be prohibiting. As such, one limitation is the need for a large number of loads (more than 20,000 preferably).

- Another related limitation is the examined distributions of TCLs parameters, given the experimental data available. For other distributions the accuracy is unknown, yet same principles can be used, for instance clustering according to t_{on}, t_{off} parameters and use of representative population values to calculate the TPMs.

- For the state space model and state estimation, Gaussian (white) noise was assumed. In reality, coloured noise is more likely to be observed, in which case the GBP algorithm has to be adjusted accordingly.

- Power measurements are a requirement, they can be provided from Smart Meters and/or Smart Plugs and their Apps, or even from Smart Appliances directly. Alternatively, (thermal) state readings can be used (though not advised due to the associated cost of extra sensors) or disaggregation.

- The intervals of readings (due to privacy concerns) were assumed to be at 15 minutes, for significantly longer intervals the accuracy might not be satisfactory.

Although the proposed methods can achieve relatively highly accurate results, there are always errors when approximate and stochastic models are used. Even with the most accurate dynamic models, exact knowledge of DR availability in advance (as in day ahead) is not possible, but only some forecast of minimum/maximum DR availability with level of confidence. This is precisely why on-line tracking and updating is crucial. Still, DR cannot be a "stand alone" solution and some form of backup will be required to ensure the minimum balancing requirements are always met.

Another important limiting factor during the course of this thesis was the scarce availability of DR and residential load real world data. Additionally, no field trials took place to verify the results in real world.

6.4 Further Work

Further improvement in the accuracy of aggregated behaviour of loads, as those considered in this work and not only, is possible. First and foremost, pilot studies are essential for commercial acceptance; main aspects to focus on would be studying the distribution of real-world loads, preferably in the order of thousands and state estimation noise, which is expected to be non-Gaussian.

Dynamic updating in real-time can also be studied more in depth and improved, especially during periods when aggregated behaviour is shifting, as well as the effect on control actions. For frequency DR services, these should be combined with analysis of the effect of communication constraints, delays and packet loss.

Moreover, similar models should be tested and validated on other loads with storage capabilities and/or similar characteristics, notably EVs and potentially wet loads. Then the impact on national level should be examined. Lastly, the effect of both active and reactive power changes associated to the control of loads should be investigated in realistic networks; on voltage profiles (distribution networks - active power, transmission networks - reactive power).

List of References

[1] F. Teng, M. Aunedi, G. Strbac, V. Trovato, and A. Dallagi, "Provision of ancillary services in future low-carbon uk electricity system," in *Innovative Smart Grid Technologies Conference Europe (ISGT-Europe), 2017 IEEE PES*, pp. 1–6, IEEE, 2017.

[2] V. Trovato, S. H. Tindemans, and G. Strbac, "Demand response contribution to effective inertia for system security in the gb 2020 gone green scenario," in *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES*, pp. 1–5, IEEE, 2013.

[3] A. Ulbig, T. S. Borsche, and G. Andersson, "Impact of low rotational inertia on power system stability and operation," *arXiv preprint arXiv:1312.6435*, 2013.

[4] National Grid, "Frequency response services." (Online): https://www.nationalgrid.com/uk/electricity/balancing-services/frequency-response-services. Accessed March 2018.

[5] National Grid ESO, "List of balancing services." (Online): https://www.nationalgrideso.com/balancing-services/list-all-balancing-services. Accessed August 2020.

[6] L. Wu and D. Infield, "Modelling the provision of inertial response from

variable speed wind turbines," *IET Conference on Renewable Power Generation (RPG 2011)*, pp. 223–229, 2011.

[7] P. Fortenbacher, G. Andersson, and J. L. Mathieu, "Optimal real-time control of multiple battery sets for power system applications," in *PowerTech, 2015 IEEE Eindhoven*, pp. 1–6, IEEE, 2015.

[8] S. Ihara and F. C. Schweppe, "Physically based modeling of cold load pickup," *IEEE Transactions on Power Apparatus and Systems*, no. 9, pp. 4142–4150, 1981.

[9] R. Malhame and C.-Y. Chong, "Electric load model synthesis by diffusion approximation of a high-order hybrid-state stochastic system," *IEEE Transactions on Automatic Control*, vol. 30, no. 9, pp. 854–860, 1985.

[10] Global-Roam, "Explaining demand response." (Online): http://www.demandresponse.com.au/explaining-dr/. Accessed March 2018.

[11] "ADVANTAGE FP7 project." (Online): http://fp7-advantage.eu/. Accessed March 2018.

[12] J. Devlin and R. S. Worldwide, "Residential energy consumption survey," *US Department of Energy, Tech. Rep*, 2001.

[13] V. Trovato, S. H. Tindemans, and G. Strbac, "Leaky storage model for optimal multi-service allocation of thermostatic loads," *IET Generation, Transmission & Distribution*, vol. 10, no. 3, pp. 585–593, 2016.

[14] SmartGrids, ETP, "Strategic deployment document for Europe's electricity networks of the future," *European Technology Platform SmartGrids. Brussels*, 2008.

[15] T. Bigler, G. Gaderer, P. Loschmidt, and T. Sauter, "Smartfridge: Demand side management for the device level," in *Emerging Technologies & Factory Automation (ETFA), 2011 IEEE 16th Conference on*, pp. 1–8, IEEE, 2011.

[16] Y. Ota, H. Taniguchi, T. Nakajima, K. M. Liyanage, K. Shimizu, T. Masuta, J. Baba, and A. Yokoyama, "Effect of autonomous distributed vehicle-to-grid (v2g) on power system frequency control," in *Industrial and Information Systems (ICIIS), 2010 International Conference on*, pp. 481–485, IEEE, 2010.

[17] M. A. Ortega-Vazquez, "Optimal scheduling of electric vehicle charging and vehicle-to-grid services at household level including battery degradation and price uncertainty," *IET Generation, Transmission & Distribution*, vol. 8, no. 6, pp. 1007–1016, 2014.

[18] R. Moreno, R. Moreira, and G. Strbac, "A milp model for optimising multi-service portfolios of distributed energy storage," *Applied Energy*, vol. 137, pp. 554–566, 2015.

[19] S. Nistor, J. Wu, M. Sooriyabandara, and J. Ekanayake, "Capability of smart appliances to provide reserve services," *Applied Energy*, vol. 138, pp. 590–597, 2015.

[20] P. Mahat, Z. Chen, and B. Bak-Jensen, "Underfrequency load shedding for an islanded distribution system with distributed generators," *IEEE Transactions on Power Delivery*, vol. 25, no. 2, pp. 911–918, 2010.

[21] R. D'hulst, W. Labeeuw, B. Beusen, S. Claessens, G. Deconinck, and K. Vanthournout, "Demand response flexibility and flexibility potential of residential smart appliances: Experiences from large pilot test in belgium," *Applied Energy*, vol. 155, pp. 79–90, 2015.

[22] V. Trovato, I. M. Sanz, B. Chaudhuri, and G. Strbac, "Advanced control

of thermostatic loads for rapid frequency response in great britain," *IEEE Transactions on Power Systems*, vol. 32, no. 3, pp. 2106–2117, 2017.

[23] C. Le Floch, E. Kara, and S. Moura, "PDE modeling and control of electric vehicle fleets for ancillary services: A discrete charging case," *IEEE Transactions on Smart Grid*, 2016.

[24] Electricity North West, "CLASS capability report for trial scenarios." (Online): https://www.enwl.co.uk/innovation/class/learning-and-key-documents/class-trials/, 2014. Accessed March 2018.

[25] A. Kleidaras, A. E. Kiprakis, and J. S. Thompson, "Human in the loop heterogeneous modelling of thermostatically controlled loads for demand side management studies," *Energy*, vol. 145, pp. 754–769, 2018.

[26] V. Trovato, S. Tindemans, and G. Strbac, "Understanding the aggregate flexibility of thermostatically controlled loads," in *PowerTech, 2017 IEEE Manchester*, pp. 1–6, IEEE, 2017.

[27] L. C. Totu, J. Leth, and R. Wisniewski, "Control for large scale demand response of thermostatic loads.," in *ACC*, pp. 5023–5028, 2013.

[28] D. S. Callaway, "Tapping the energy storage potential in electric loads to deliver load following and regulation, with application to wind energy," *Energy Conversion and Management*, vol. 50, no. 5, pp. 1389–1400, 2009.

[29] S. Koch, J. L. Mathieu, and D. S. Callaway, "Modeling and control of aggregated heterogeneous thermostatically controlled loads for ancillary services," in *Proc. PSCC*, pp. 1–7, Citeseer, 2011.

[30] J. A. Short, D. G. Infield, and L. L. Freris, "Stabilization of grid frequency through dynamic demand control," *IEEE Transactions on power systems*, vol. 22, no. 3, pp. 1284–1293, 2007.

[31] T. M. Keep, F. E. Sifuentes, D. M. Auslander, and D. S. Callaway, "Using load switches to control aggregated electricity demand for load following and regulation," in *Power and Energy Society General Meeting, 2011 IEEE*, pp. 1–7, IEEE, 2011.

[32] E. Vrettos, S. Koch, and G. Andersson, "Load frequency control by aggregations of thermally stratified electric water heaters," in *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*, pp. 1–8, IEEE, 2012.

[33] M. Juelsgaard, L. C. Totu, S. E. Shafiei, R. Wisniewski, and J. Stoustrup, "Control structures for smart grid balancing," in *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES*, pp. 1–5, IEEE, 2013.

[34] L. C. Totu and R. Wisniewski, "Demand response of thermostatic loads by optimized switching-fraction broadcast," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 9956–9961, 2014.

[35] L. C. Totu, R. Wisniewski, and J. Leth, "Modeling populations of thermostatic loads with switching rate actuation," *arXiv preprint arXiv:1411.2864*, 2014.

[36] V. Trovato, S. H. Tindemans, and G. Strbac, "Designing effective frequency response patterns for flexible thermostatic loads," in *Environment and Electrical Engineering (EEEIC), 2015 IEEE 15th International Conference on*, pp. 1003–1008, IEEE, 2015.

[37] S. H. Tindemans, V. Trovato, and G. Strbac, "Decentralized control of thermostatic loads for flexible demand response," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 5, pp. 1685–1700, 2015.

[38] V. Trovato and G. Strbac, "Multiple services allocation for flexible thermostatic loads," in *CIRED Workshop 2016*, pp. 1–4, IET, 2016.

[39] K. Vinther, T. Green, E. Shafiei, L. C. Totu, R. Izadi-Zamanabadi, and T. G. Hovgaard, "Control strategies and challenges for utilizing supermarket refrigeration systems in a smart energy context," in *Control Applications (CCA), 2016 IEEE Conference on*, pp. 593–598, IEEE, 2016.

[40] J. L. Mathieu and D. S. Callaway, "State estimation and control of heterogeneous thermostatically controlled loads for load following," in *2012 45th Hawaii International Conference on System Sciences*, pp. 2002–2011, IEEE, 2012.

[41] G. S. Ledva, Z. Du, L. Balzano, and J. L. Mathieu, "Disaggregating load by type from distribution system measurements in real time," in *Energy Markets and Responsive Grids*, pp. 413–437, Springer, 2018.

[42] G. S. Ledva, L. Balzano, and J. L. Mathieu, "Real-time energy disaggregation of a distribution feeder's demand using online learning," *IEEE Transactions on Power Systems*, 2018.

[43] E. Vrettos, A. Witzig, R. Kurmann, S. Koch, and G. Andersson, "Maximizing local pv utilization using small-scale batteries and flexible thermal loads," *EU PVSEC*, 2013.

[44] E. Vrettos, F. Oldewurtel, F. Zhu, and G. Andersson, "Robust provision of frequency reserves by office building aggregations," *IFAC Proceedings Volumes*, vol. 47, no. 3, pp. 12068–12073, 2014.

[45] G. S. Ledva, E. Vrettos, S. Mastellone, G. Andersson, and J. L. Mathieu, "Applying networked estimation and control algorithms to address communication bandwidth limitations and latencies in demand response," in *System Sciences (HICSS), 2015 48th Hawaii International Conference on*, pp. 2645–2654, IEEE, 2015.

[46] C. Ziras, E. Vrettos, and G. Andersson, "Primary frequency control with

refrigerators under startup dynamics and lockout constraints," in *Power & Energy Society General Meeting, 2015 IEEE*, pp. 1–5, IEEE, 2015.

[47] C. Ziras, E. Vrettos, and Y. Shi, "Controllability and stability of primary frequency control from thermostatic loads with delays," *Journal of Modern Power Systems and Clean Energy*, vol. 5, no. 1, pp. 43–54, 2017.

[48] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, and S. McLaughlin, "Development of low-voltage load models for the residential load sector," *IEEE Transactions on Power Systems*, vol. 29, no. 5, pp. 2180–2188, 2014.

[49] G. Tsagarakis, A. J. Collin, and A. E. Kiprakis, "Modelling the electrical loads of uk residential energy users," in *2012 47th International Universities Power Engineering Conference (UPEC)*, pp. 1–6, 2012.

[50] A. J. Collin, G. Tsagarakis, A. E. Kiprakis, and S. McLaughlin, "Multi-scale electrical load modelling for demand-side management," in *2012 3rd IEEE PES Innovative Smart Grid Technologies Europe (ISGT Europe)*, pp. 1–8, 2012.

[51] R. Stamminger, G. Broil, C. Pakula, H. Jungbecker, M. Braun, I. Rüdenauer, and C. Wendker, "Synergy potential of smart appliances," *Report of the Smart-A project*, 2008.

[52] DECC, "Department of energy & climate change, energy consumption in the UK." (Online): https://www.gov.uk/government/collections/energy-consumption-in-the-uk.

[53] National Grid ESO, "Firm frequency response balancing servive, test guidance for providers." (Online): https://www.nationalgrideso.com/document/148811/download, 2012. Accessed August 2020.

[54] J. L. Mathieu, S. Koch, and D. S. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," *IEEE Transactions on Power Systems*, vol. 28, no. 1, pp. 430–440, 2013.

[55] E. Vrettos, J. L. Mathieu, and G. Andersson, "Control of thermostatic loads using moving horizon estimation of individual load states," in *Power Systems Computation Conference (PSCC), 2014*, pp. 1–7, IEEE, 2014.

[56] SmartMeters, "Technical information types of smart meter." (Online):https://www.smartme.co.uk/technical.html. Accessed August 2020.

[57] Wikipedia, "Wemo." (Online): https://en.wikipedia.org/wiki/WEMO. Accessed August 2020.

[58] ThingsBoard, "Smart energy monitoring, data visualization and energy efficiency analysis." (Online): https://thingsboard.io/smart-energy/. Accessed August 2020.

[59] J. Palmer, N. Terry, and T. Kane, "Further analysis of the household electricity survey-early findings: Demand side management," *Department of Energy and Climate Change (DECC): London, UK*, 2013.

[60] T.-H. Chang, M. Alizadeh, and A. Scaglione, "Real-time power balancing via decentralized coordinated home energy scheduling," *Smart Grid, IEEE Transactions on*, vol. 4, no. 3, pp. 1490–1504, 2013.

[61] M. H. Albadi and E. El-Saadany, "A summary of demand response in electricity markets," *Electric power systems research*, vol. 78, no. 11, pp. 1989–1996, 2008.

[62] A. Subramanian, M. J. Garcia, D. S. Callaway, K. Poolla, and P. Varaiya, "Real-time scheduling of distributed resources," *Smart Grid, IEEE Transactions on*, vol. 4, no. 4, pp. 2122–2130, 2013.

[63] R. Carmichael, J. Schofield, M. Woolf, M. Bilton, R. Ozaki, and G. Strbac, "Residential consumer attitudes to time-varying pricing," *Report A2 for the "Low Carbon London" LCNF project: Imperial College London*, 2014.

[64] A. Subramanian, M. Garcia, A. Domínguez-García, D. Callaway, K. Poolla, and P. Varaiya, "Real-time scheduling of deferrable electric loads," in *American Control Conference (ACC), 2012*, pp. 3643–3650, IEEE, 2012.

[65] G. Strbac, "Demand side management: Benefits and challenges," *Energy policy*, vol. 36, no. 12, pp. 4419–4426, 2008.

[66] A. Faruqui, D. Harris, and R. Hledik, "Unlocking the 53 billion savings from smart meters in the eu: How increasing the adoption of dynamic tariffs could make or break the eu's smart grid investment," *Energy Policy*, vol. 38, no. 10, pp. 6222–6231, 2010.

[67] P. Du and N. Lu, "Appliance commitment for household load scheduling," *IEEE transactions on Smart Grid*, vol. 2, no. 2, pp. 411–419, 2011.

[68] A. Kleidaras and A. Kiprakis, "A roadmap for domestic load modelling for large-scale demand management within smart grids," in *Wireless and Satellite Systems*, pp. 33–47, Springer, 2015.

[69] L. C. Totu, R. Wisniewski, and J. Leth, "Demand response of a tcl population using switching-rate actuation," *IEEE Transactions on Control Systems Technology*, 2016.

[70] National Grid, "Short term operation reserve." (Online): http://www2.nationalgrid.com/UK/Services/Balancing-services/Reserve-services/Short-Term-Operating-Reserve/STOR-Runway/. Accessed March 2018.

[71] National Grid, "Demand turn up." (Online):

https://www.nationalgrid.com/uk/electricity/balancing-services/reserve-services/demand-turn. Accessed March 2018.

[72] S. Kawachi, H. Hagiwara, J. Baba, K. Furukawa, E. Shimoda, and S. Numata, "Modeling and simulation of heat pump air conditioning unit intending energy capacity reduction of energy storage system in microgrid," in *Power Electronics and Applications (EPE 2011), Proceedings of the 2011-14th European Conference on*, pp. 1–9, IEEE, 2011.

[73] M. U. Kajgaard, J. Mogensen, A. Wittendorff, A. T. Veress, and B. Biegel, "Model predictive control of domestic heat pump," in *American Control Conference (ACC), 2013*, IEEE, 2013.

[74] G. Papaefthymiou, B. Hasche, and C. Nabe, "Potential of heat pumps for demand side management and wind power integration in the german electricity market," *IEEE Transactions on Sustainable Energy*, vol. 3, no. 4, pp. 636–642, 2012.

[75] O. Malik and P. Havel, "Active demand-side management system to facilitate integration of res in low-voltage distribution networks," *Sustainable Energy, IEEE Transactions on*, vol. 5, no. 2, pp. 673–681, 2014.

[76] N. Lu and D. P. Chassin, "A state-queueing model of thermostatically controlled appliances," *Power Systems, IEEE Transactions on*, vol. 19, no. 3, pp. 1666–1673, 2004.

[77] T. Masuta and A. Yokoyama, "Supplementary load frequency control by use of a number of both electric vehicles and heat pump water heaters," *Smart Grid, IEEE Transactions on*, vol. 3, no. 3, pp. 1253–1262, 2012.

[78] M. Starke, D. Letto, N. Alkadi, R. George, B. Johnson, K. Dowling, and S. Khan, "Demand-side response from industrial loads," in *2013 NSTI Nanotechnology Conference and Expo*, vol. 2, p. 46, 2013.

[79] Y.-J. Kim, J. L. Kirtley, and L. K. Norford, "Variable speed heat pump design for frequency regulation through direct load control," in *T&D Conference and Exposition, 2014 IEEE PES*, pp. 1–5, IEEE, 2014.

[80] National Grid, "Demand side response." (Online): https://www.nationalgrid.com/uk/electricity/balancing-services/demand-side-response-dsr. Accessed March 2018.

[81] DECC, "Department of energy and climate change statistics." (Online): https://www.gov.uk/government/uploads/system/uploads/attachment/data/file/337649/chapter_5.pdf. Accessed March 2018.

[82] B. Stephen and S. J. Galloway, "Domestic load characterization through smart meter advance stratification," *Smart Grid, IEEE Transactions on*, vol. 3, no. 3, pp. 1571–1572, 2012.

[83] S. Ramos, J. Duarte, J. Soares, Z. Vale, and F. J. Duarte, "Typical load profiles in the smart grid context—a clustering methods comparison," in *Power and Energy Society General Meeting, 2012 IEEE*, pp. 1–8, IEEE, 2012.

[84] R. Li, C. Gu, F. Li, G. Shaddick, and M. Dale, "Development of low voltage network templates—part i: Substation clustering and classification," *IEEE Transactions on Power Systems*, vol. 30, no. 6, pp. 3036–3044, 2015.

[85] Office for National Statistics, "2011 census: Population and household estimates for the United Kingdom." (Online): https://cy.ons.gov.uk/peoplepopulationandcommunity/populationandmig-/ration/populationestimates/datasets/2011censuspopulationandhouseholde-/stimatesfortheunitedkingdom. Accessed March 2018.

[86] M. A. Zehir and M. Bagriyanik, "Demand side management by controlling refrigerators and its effects on consumers," *Energy Conversion and Management*, vol. 64, pp. 238–244, 2012.

[87] R. Kemna, M. van Elburg, W. Li, and R. van Holsteijn, "Eco-design of water heaters," *Draft reports of tasks*, pp. 1–5, 2007.

[88] Z. Yu, L. Jia, M. C. Murphy-Hoye, A. Pratt, and L. Tong, "Modeling and stochastic control for home energy management," *IEEE Transactions on Smart Grid*, vol. 4, no. 4, pp. 2244–2255, 2013.

[89] P. Siano, "Demand response and smart grids—a survey," *Renewable and Sustainable Energy Reviews*, vol. 30, pp. 461–478, 2014.

[90] I. H. Gil, B. Hayes, A. Collin, and S. Z. Djokic, "Distribution network equivalents for reliability analysis. part 2: storage and demand-side resources," in *IEEE/PES Innovative Smart Grid Technologies Europe*, 2013.

[91] D. Niyato, Q. Dong, P. Wang, and E. Hossain, "Optimizations of power consumption and supply in the smart grid: Analysis of the impact of data communication reliability," *IEEE Transactions on Smart Grid*, vol. 4, no. 1, pp. 21–35, 2013.

[92] Nang Luong, "Application of virtual power plant model in intelligent grid control." (Online): http://nangluongvietnam.vn/news/vn/timkiem&q=virtual%20. Accessed March 2018.

[93] J. Mathieu, M. Dyson, D. Callaway, and A. Rosenfeld, "Using residential electric loads for fast demand response: The potential resource and revenues, the costs, and policy recommendations," in *ACEEE Summer Study on Energy Efficiency in Buildings*, pp. 189–203, Citeseer, 2012.

[94] H. Niska, "Extracting controllable heating loads from aggregated smart meter data using clustering and predictive modelling," in *Intelligent Sensors, Sensor Networks and Information Processing, 2013 IEEE Eighth International Conference on*, pp. 368–373, IEEE, 2013.

[95] Y.-H. Lin and M.-S. Tsai, "Non-intrusive load monitoring by novel neuro-fuzzy classification considering uncertainties," *Smart Grid, IEEE Transactions on*, vol. 5, no. 5, pp. 2376–2384, 2014.

[96] E. Vrettos, C. Ziras, and G. Andersson, "Fast and reliable primary frequency reserves from refrigerators with decentralized stochastic control," *IEEE Transactions on Power Systems*, vol. 32, no. 4, pp. 2924–2941, 2017.

[97] R. E. Mortensen and K. P. Haggerty, "A stochastic computer model for heating and cooling loads," *Power Systems, IEEE Transactions on*, vol. 3, no. 3, pp. 1213–1219, 1988.

[98] S. El-Ferik and R. P. Malhame, "Identification of alternating renewal electric load models from energy measurements," *IEEE Transactions on Automatic Control*, vol. 39, no. 6, pp. 1184–1196, 1994.

[99] H. Masjuki, R. Saidur, I. Choudhury, T. Mahlia, A. Ghani, and M. Maleque, "The applicability of iso household refrigerator–freezer energy test specifications in malaysia," *Energy*, vol. 26, no. 7, pp. 723–737, 2001.

[100] M. Hasanuzzaman, R. Saidur, and H. Masjuki, "Effects of operating variables on heat transfer and energy consumption of a household refrigerator-freezer during closed door operation," *Energy*, vol. 34, no. 2, pp. 196–198, 2009.

[101] C. J. Hermes and C. Melo, "Assessment of the energy performance of household refrigerators via dynamic simulation," *Applied Thermal Engineering*, vol. 29, no. 5-6, pp. 1153–1165, 2009.

[102] O. Laguerre and D. Flick, "Heat transfer by natural convection in domestic refrigerators," *Journal of Food Engineering*, vol. 62, no. 1, pp. 79–88, 2004.

[103] O. Laguerre, S. Benamara, and D. Flick, "Numerical simulation of simultaneous heat and moisture transfer in a domestic refrigerator," *International journal of refrigeration*, vol. 33, no. 7, pp. 1425–1433, 2010.

[104] C. H. Wai, M. Beaudin, H. Zareipour, A. Schellenberg, and N. Lu, "Cooling devices in demand response: A comparison of control methods," *Smart Grid, IEEE Transactions on*, vol. 6, no. 1, pp. 249–260, 2015.

[105] C. Perfumo, E. Kofman, J. H. Braslavsky, and J. K. Ward, "Load management: Model-based control of aggregate power for populations of thermostatically controlled loads," *Energy Conversion and Management*, vol. 55, pp. 36–48, 2012.

[106] T. Kane, S. K. Firth, K. J. Lomas, D. Allinson, and K. Irvine, "Variation of indoor temperatures and heating practices in uk dwellings," 2011.

[107] D. S. Callaway and I. A. Hiskens, "Achieving controllability of electric loads," *Proceedings of the IEEE*, vol. 99, no. 1, pp. 184–199, 2011.

[108] E. De Cian, E. Lanzi, and R. Roson, "The impact of temperature change on energy demand: a dynamic panel analysis," 2007.

[109] J. Leth, R. Wisniewski, J. Rasmussen, and H. Schioler, "Stochastic analysis of synchronization in a supermarket refrigeration system," *arXiv preprint arXiv:1501.06006*, 2015.

[110] R. Wisniewski, J. Leth, and J. G. Rasmussen, "Analysis of synchronization in a supermarket refrigeration system," *Control Theory and Technology*, vol. 12, no. 2, pp. 154–162, 2014.

[111] F. R. Kschischang, B. J. Frey, and H.-A. Loeliger, "Factor graphs and the sum-product algorithm," *IEEE Transactions on information theory*, vol. 47, no. 2, pp. 498–519, 2001.

[112] M. Cosovic and D. Vukobratovic, "Distributed gauss-newton method for state estimation using belief propagation," *IEEE Transactions on Power Systems*, 2018.

[113] S. E. Z. Soudjani and A. Abate, "Aggregation and control of populations of thermostatically controlled loads by formal abstractions," *IEEE Transactions on Control Systems Technology*, vol. 23, no. 3, pp. 975–990, 2014.

[114] D. Wang, S. Parkinson, W. Miao, H. Jia, C. Crawford, and N. Djilali, "Hierarchical market integration of responsive loads as spinning reserve," *Applied energy*, vol. 104, pp. 229–238, 2013.

[115] N. Ruiz, I. Cobelo, and J. Oyarzabal, "A direct load control model for virtual power plant management," *IEEE Transactions on Power Systems*, vol. 24, no. 2, pp. 959–966, 2009.

[116] J. L. Mathieu, "Modeling, analysis, and control of demand response resources," 2012.

[117] S. Bashash and H. K. Fathy, "Modeling and control insights into demand-side energy management through setpoint control of thermostatic loads," in *American Control Conference (ACC), 2011*, pp. 4546–4553, IEEE, 2011.

[118] K. Kalsi, M. Elizondo, J. Fuller, S. Lu, and D. Chassin, "Development and validation of aggregated models for thermostatic controlled loads with demand response," in *2012 45th Hawaii International Conference on System Sciences*, pp. 1959–1966, IEEE, 2012.

[119] L. Yao and H.-R. Lu, "A two-way direct control of central air-conditioning load via the internet," *Power Delivery, IEEE Transactions on*, vol. 24, no. 1, pp. 240–248, 2009.

[120] H.-A. Loeliger, J. Dauwels, J. Hu, S. Korl, L. Ping, and F. R. Kschischang,

"The factor graph approach to model-based signal processing," *Proceedings of the IEEE*, vol. 95, no. 6, pp. 1295–1322, 2007.

[121] A. Kleidaras, M. Cosovic, D. Vukobratovic, and A. Kiprakis, "Demand response for thermostatically controlled loads using belief propagation," *Innovative Smart Grid Technologies (ISGT), 2017 IEEE International Conference*, pp. 1–6, 09 2017.

[122] H.-A. Loeliger, "An introduction to factor graphs," *IEEE Signal Processing Magazine*, vol. 21, no. 1, pp. 28–41, 2004.

[123] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Security & Privacy*, no. 3, pp. 75–77, 2009.

[124] E. McKenna, I. Richardson, and M. Thomson, "Smart meter data: Balancing consumer privacy concerns with legitimate applications," *Energy Policy*, vol. 41, pp. 807–814, 2012.

[125] C. Efthymiou and G. Kalogridis, "Smart grid privacy via anonymization of smart metering data," in *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*, pp. 238–243, IEEE, 2010.

[126] A. Rial and G. Danezis, "Privacy-preserving smart metering," in *Proceedings of the 10th annual ACM workshop on Privacy in the electronic society*, pp. 49–60, ACM, 2011.

[127] V. Trovato, F. Teng, and G. Strbac, "Role and benefits of flexible thermostatically controlled loads in future low-carbon systems," *IEEE Transactions on Smart Grid*, vol. 9, no. 5, pp. 5067–5079, 2018.

Appendix

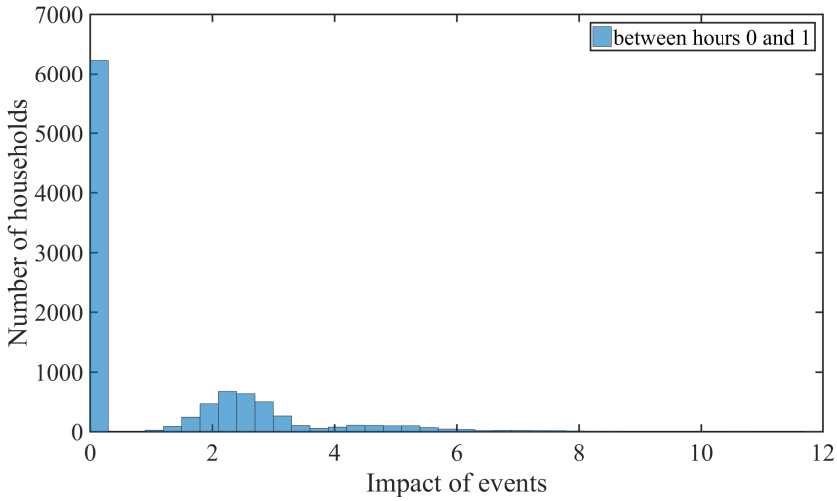


Figure 7.1: Human interaction events: distribution between 0am and 1am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

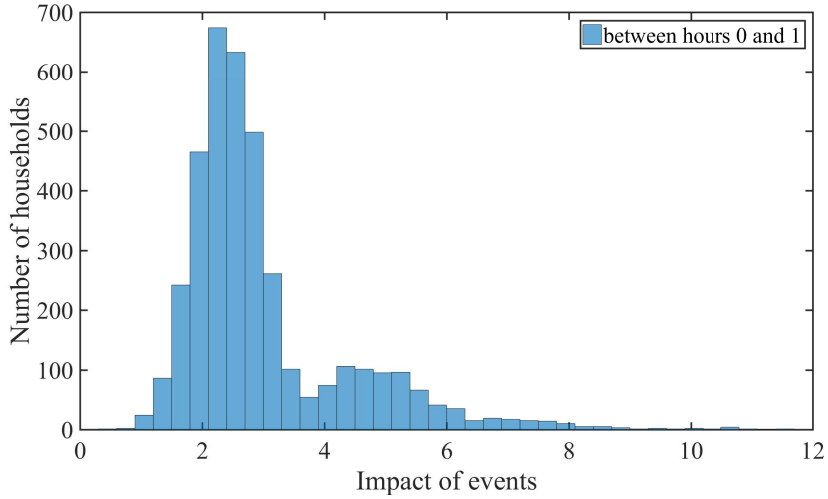


Figure 7.2: Human interaction events: distribution between 0am and 1am for 10,000 households, "zoom" to non-zero values.

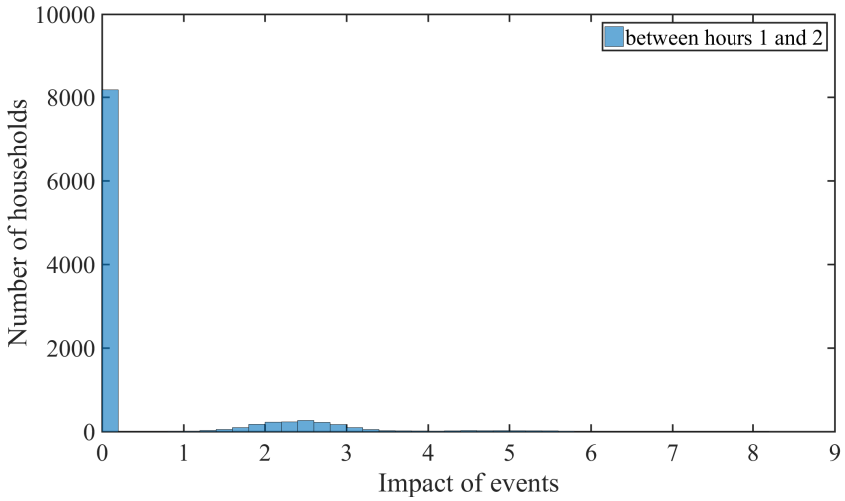


Figure 7.3: Human interaction events: distribution between 1am and 2am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

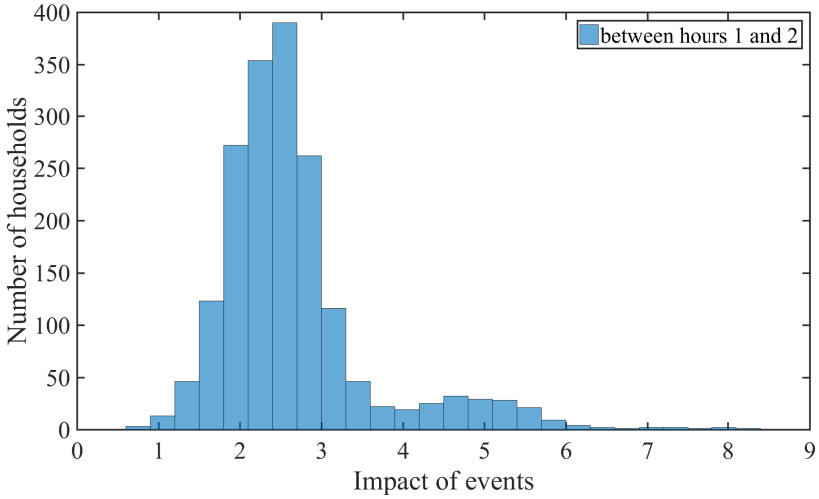


Figure 7.4: Human interaction events: distribution between 1am and 2am for 10,000 households, "zoom" to non-zero values.

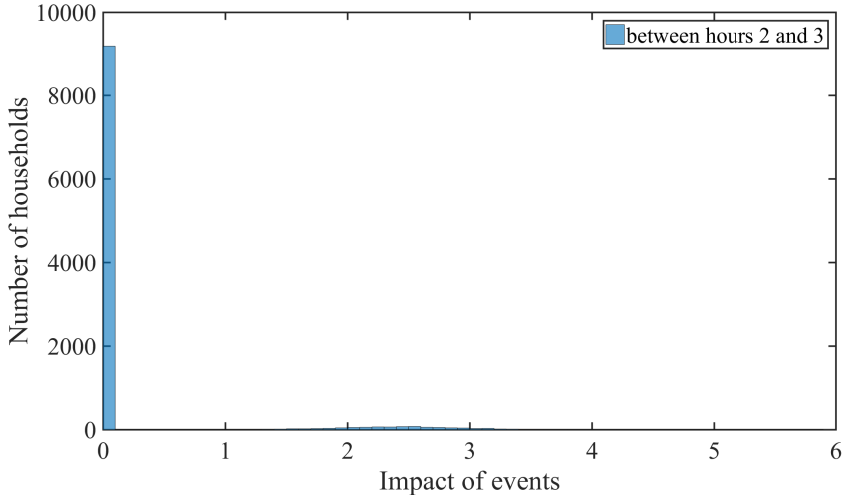


Figure 7.5: Human interaction events: distribution between 2am and 3am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

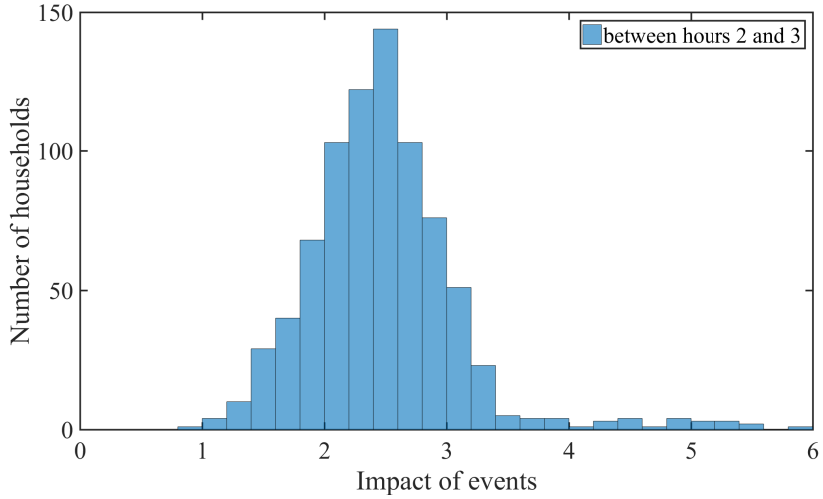


Figure 7.6: Human interaction events: distribution between 2am and 3am for 10,000 households, "zoom" to non-zero values.

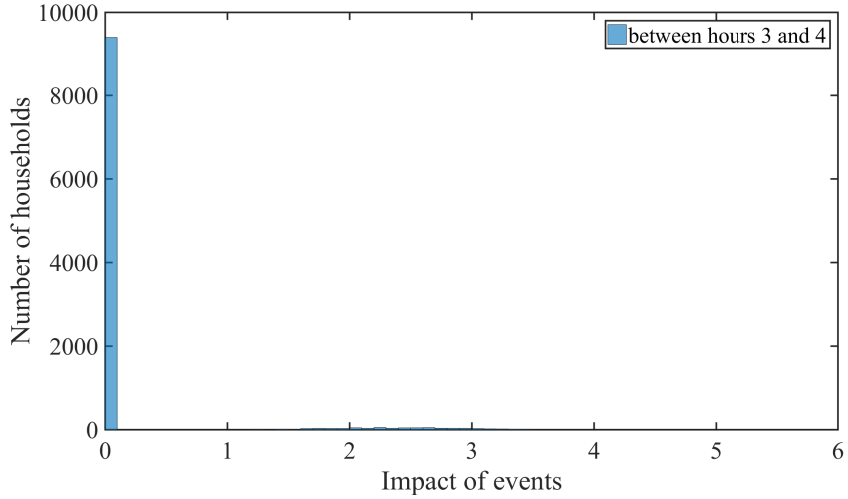


Figure 7.7: Human interaction events: distribution between 3am and 4am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

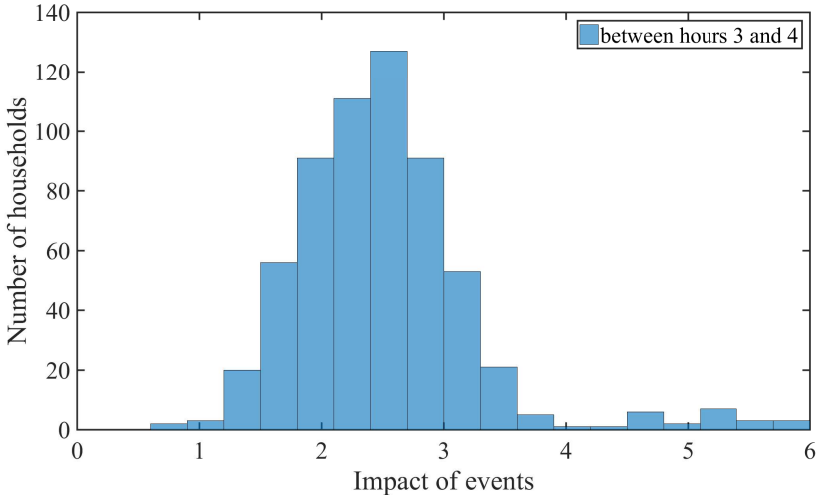


Figure 7.8: Human interaction events: distribution between 3am and 4am for 10,000 households, "zoom" to non-zero values.

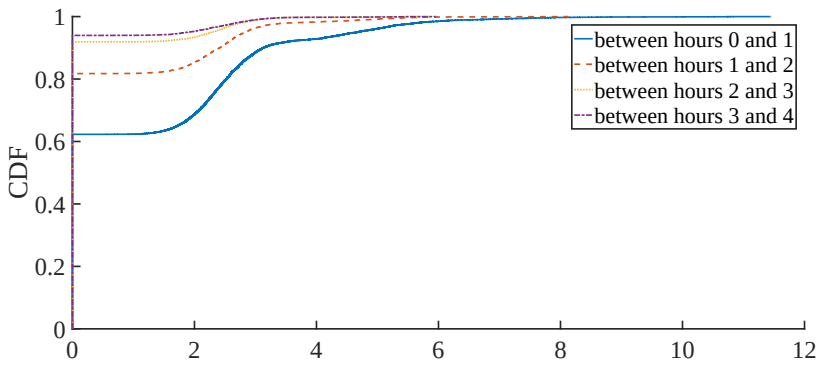


Figure 7.9: Cumulative distribution between midnight and 4am.

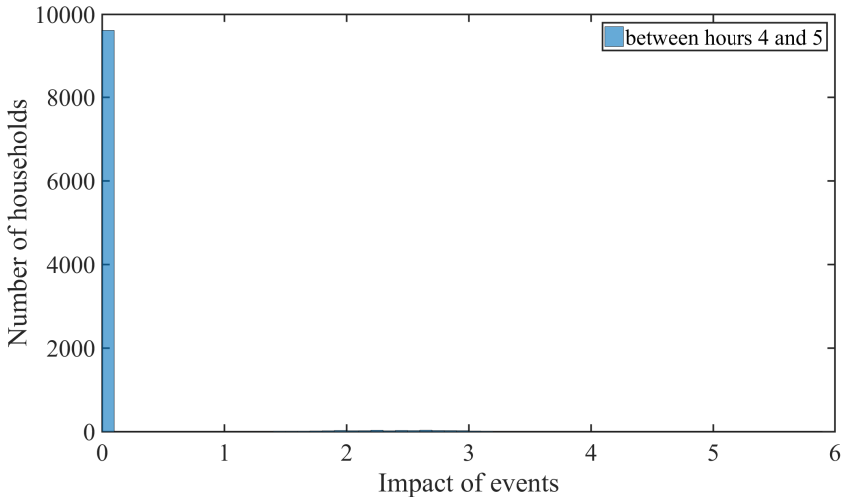


Figure 7.10: Human interaction events: distribution between 4am and 5am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

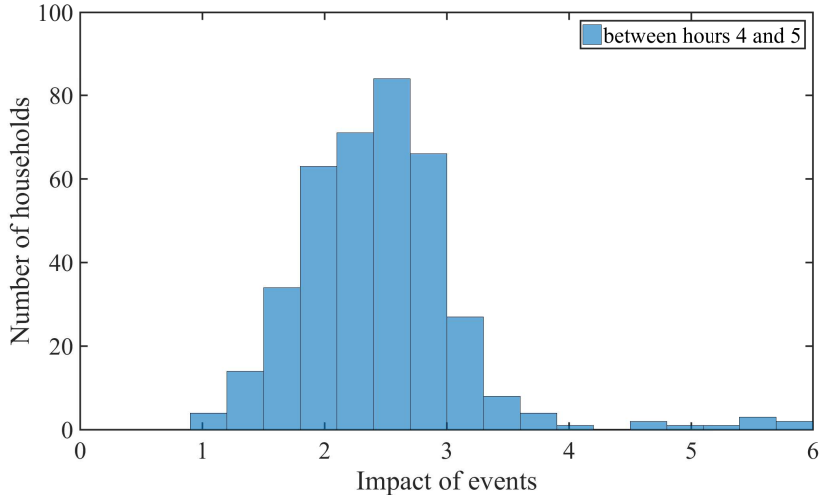


Figure 7.11: Human interaction events: distribution between 4am and 5am for 10,000 households, "zoom" to non-zero values.

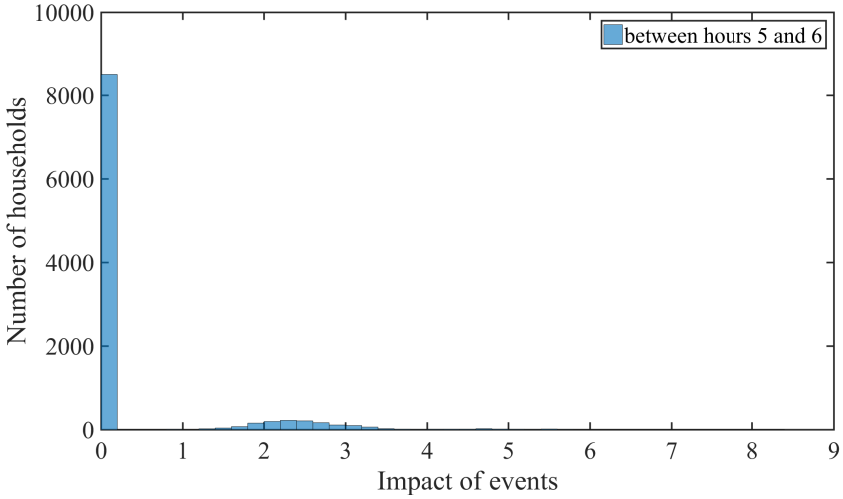


Figure 7.12: Human interaction events: distribution between 5am and 6am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

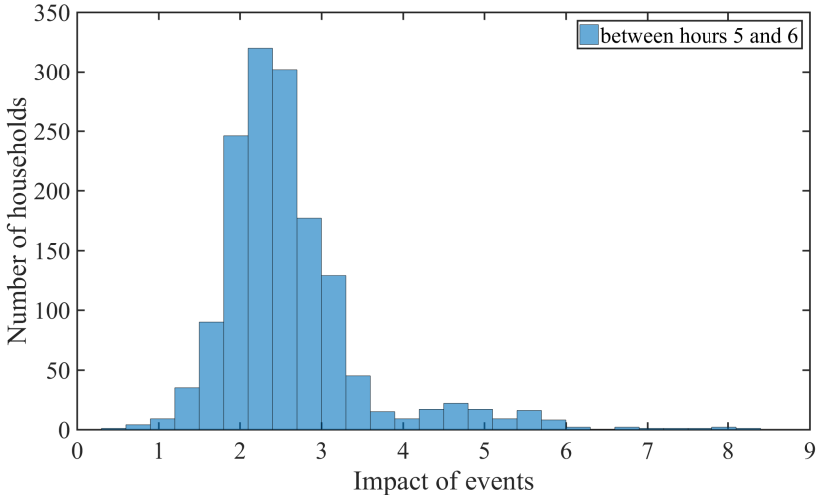


Figure 7.13: Human interaction events: distribution between 5am and 6am for 10,000 households, "zoom" to non-zero values.

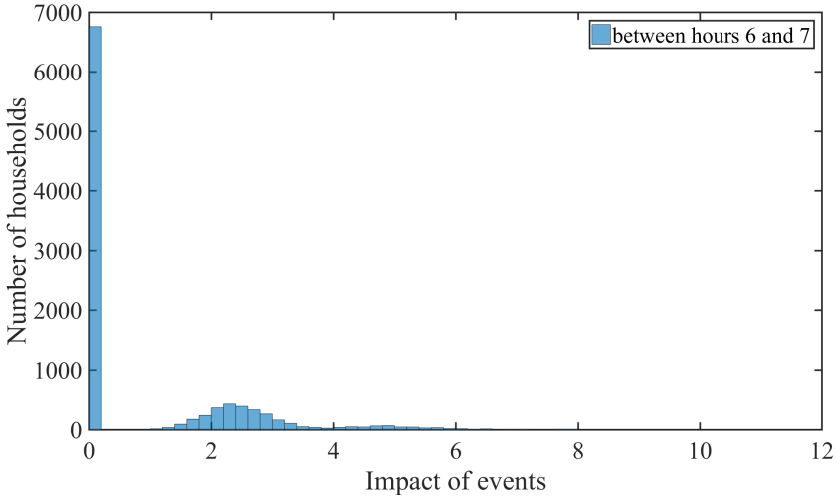


Figure 7.14: Human interaction events: distribution between 6am and 7am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

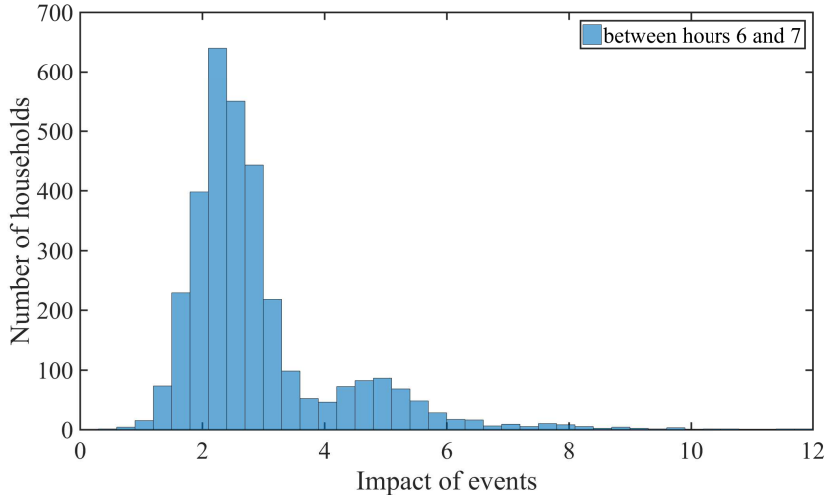


Figure 7.15: Human interaction events: distribution between 6am and 7am for 10,000 households, "zoom" to non-zero values.

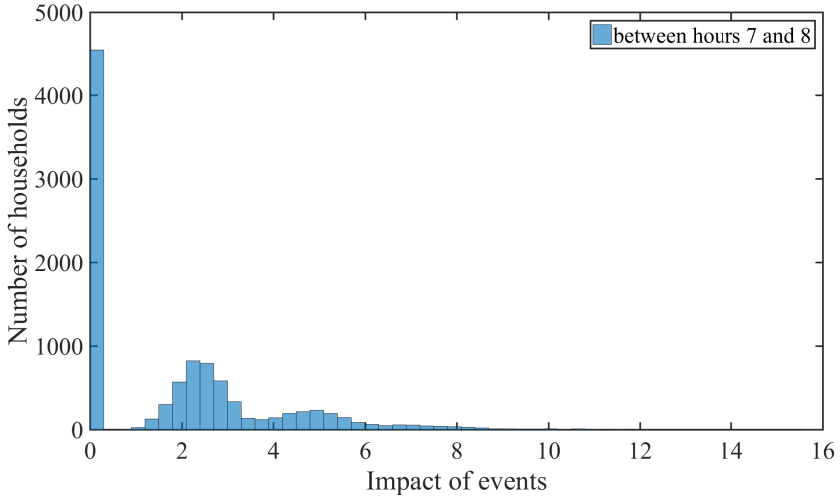


Figure 7.16: Human interaction events: distribution between 7am and 8am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

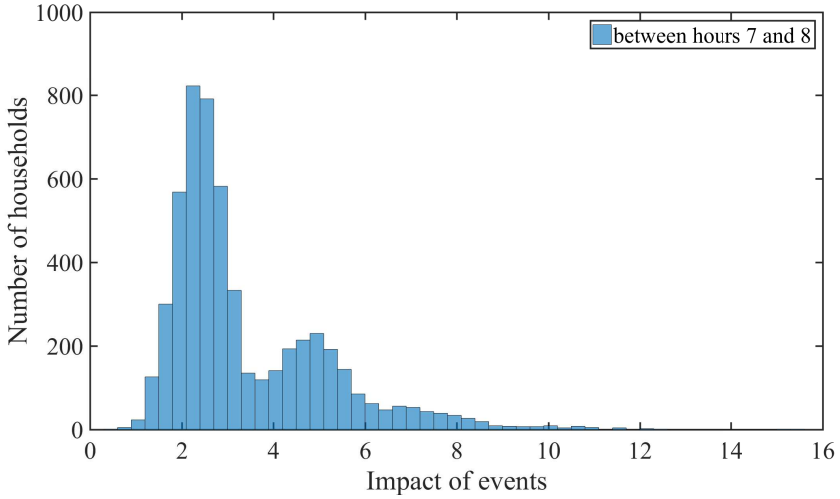


Figure 7.17: Human interaction events: distribution between 7am and 8am for 10,000 households, "zoom" to non-zero values.

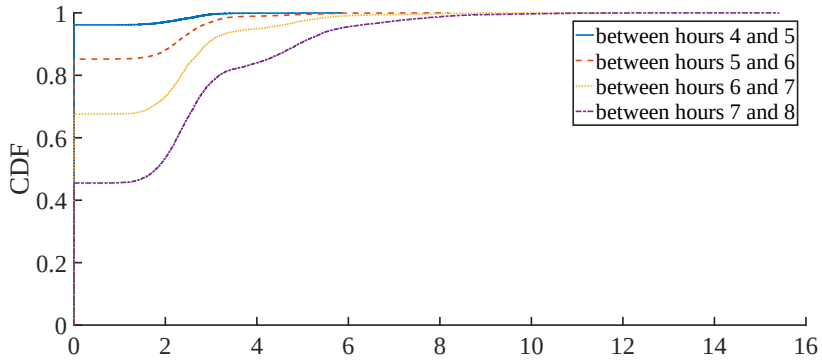


Figure 7.18: Cumulative distribution between 5am and 8am.

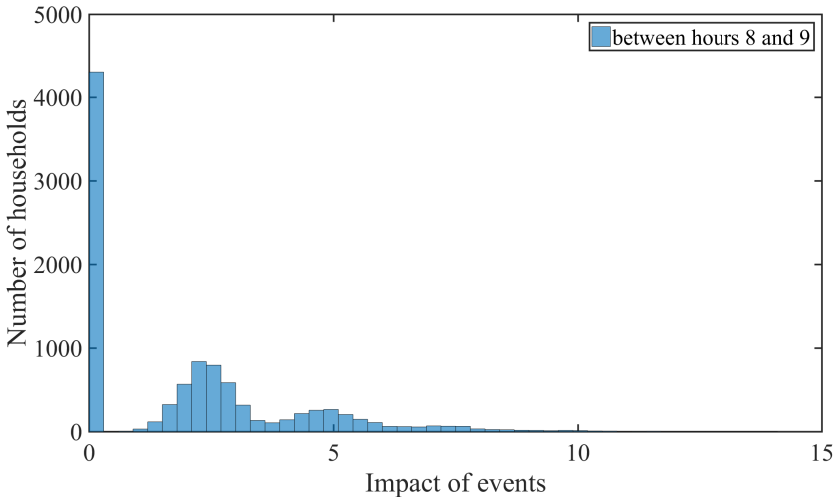


Figure 7.19: Human interaction events: distribution between 8am and 9am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

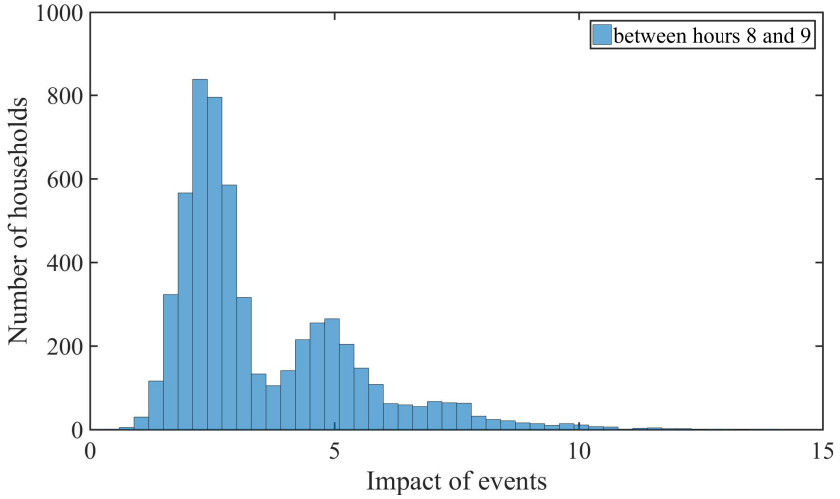


Figure 7.20: Human interaction events: distribution between 8am and 9am for 10,000 households, "zoom" to non-zero values.

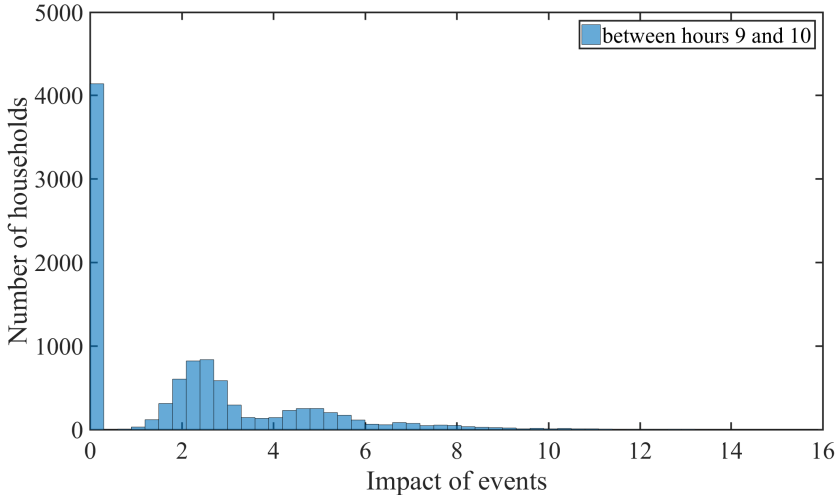


Figure 7.21: Human interaction events: distribution between 9am and 10am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

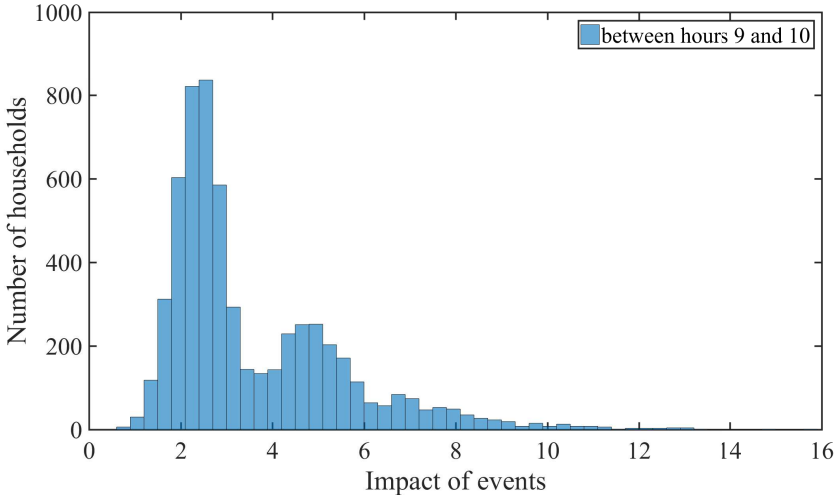


Figure 7.22: Human interaction events: distribution between 9am and 10am for 10,000 households, "zoom" to non-zero values.

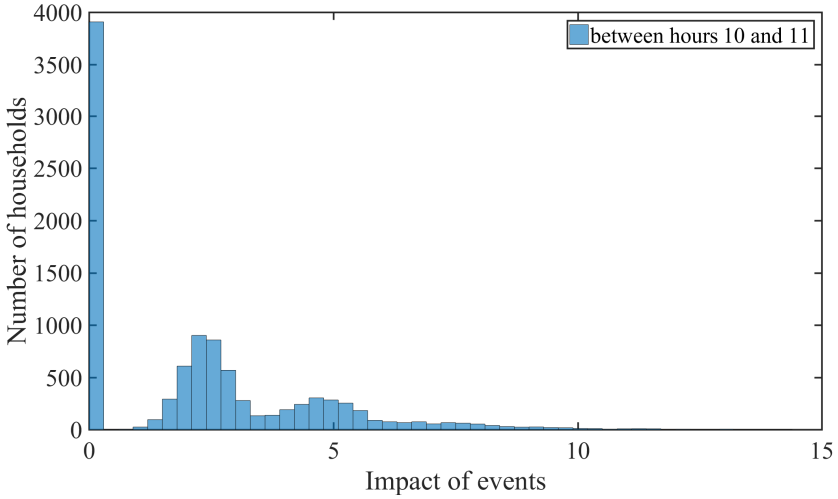


Figure 7.23: Human interaction events: distribution between 10am and 11am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

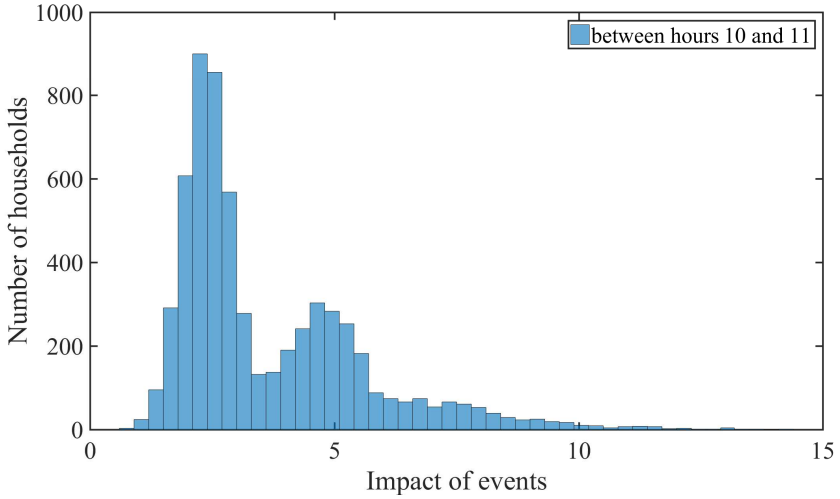


Figure 7.24: Human interaction events: distribution between 10am and 11am for 10,000 households, "zoom" to non-zero values.

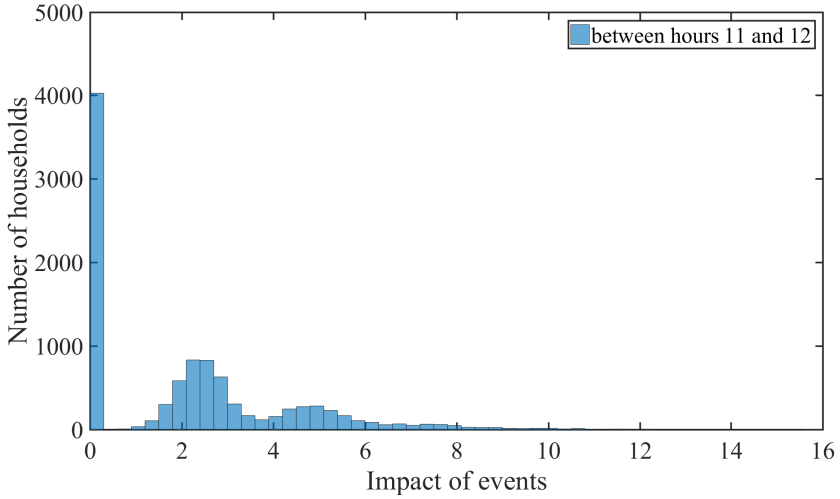


Figure 7.25: Human interaction events: distribution between 11am and 12pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

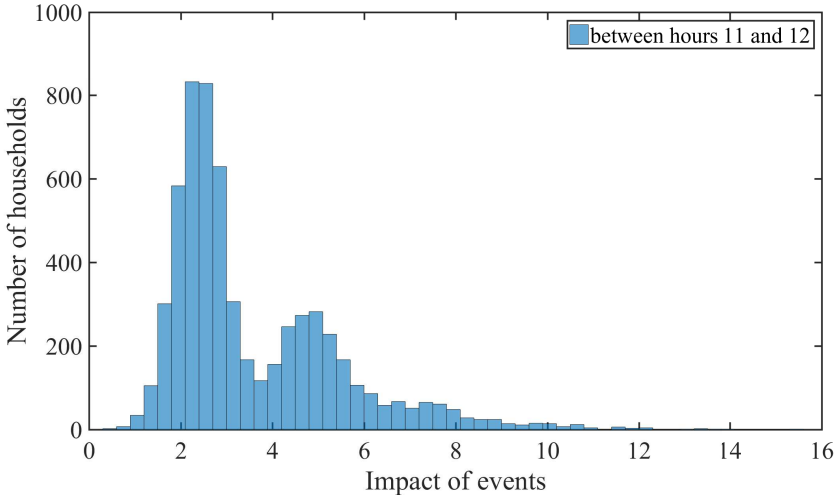


Figure 7.26: Human interaction events: distribution between 11am and 12pm for 10,000 households, "zoom" to non-zero values.

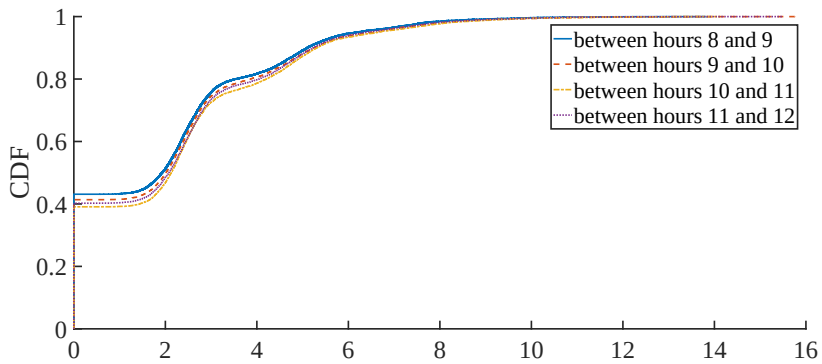


Figure 7.27: Cumulative distribution between 9am and 12pm.

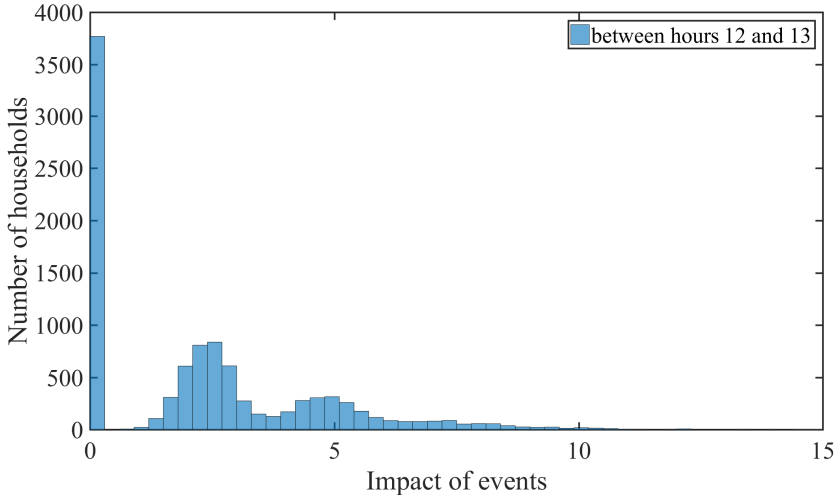


Figure 7.28: Human interaction events: distribution between 12pm and 1pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

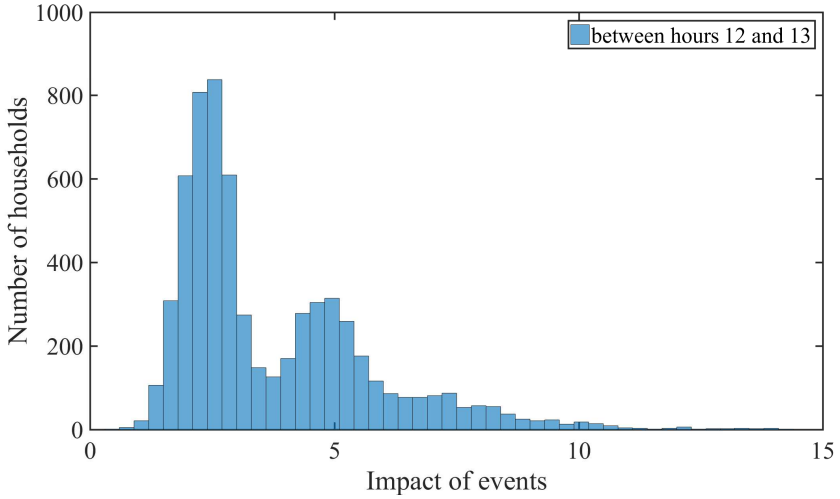


Figure 7.29: Human interaction events: distribution between 12pm and 1pm for 10,000 households, "zoom" to non-zero values.

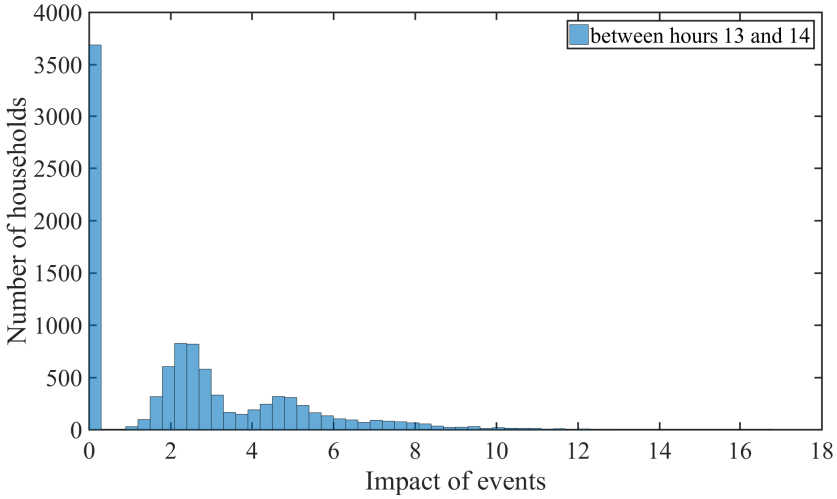


Figure 7.30: Human interaction events: distribution between 1pm and 2pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

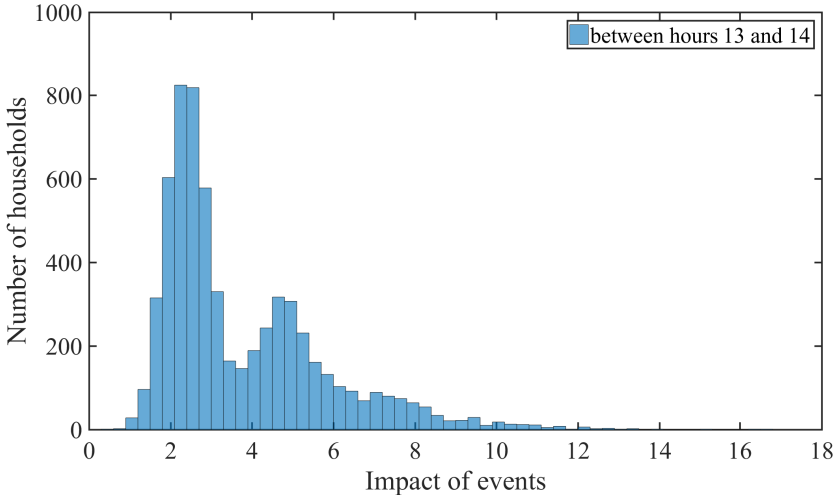


Figure 7.31: Human interaction events: distribution between 1pm and 2pm for 10,000 households, "zoom" to non-zero values.

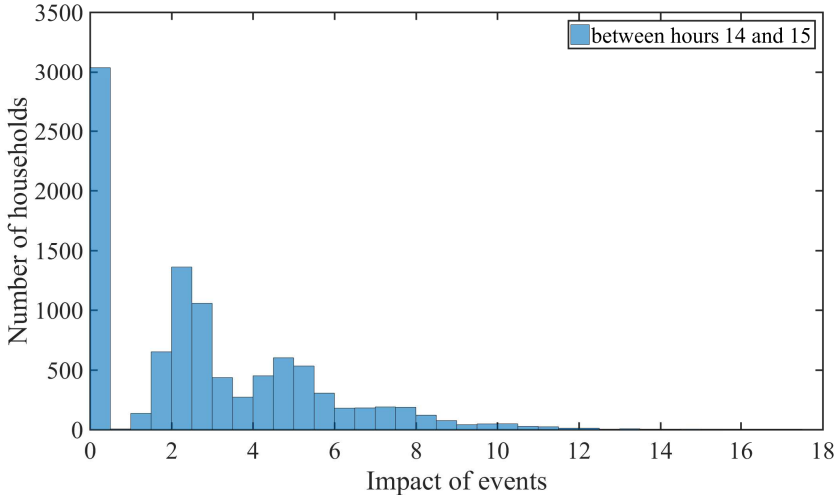


Figure 7.32: Human interaction events: distribution between 2pm and 3pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

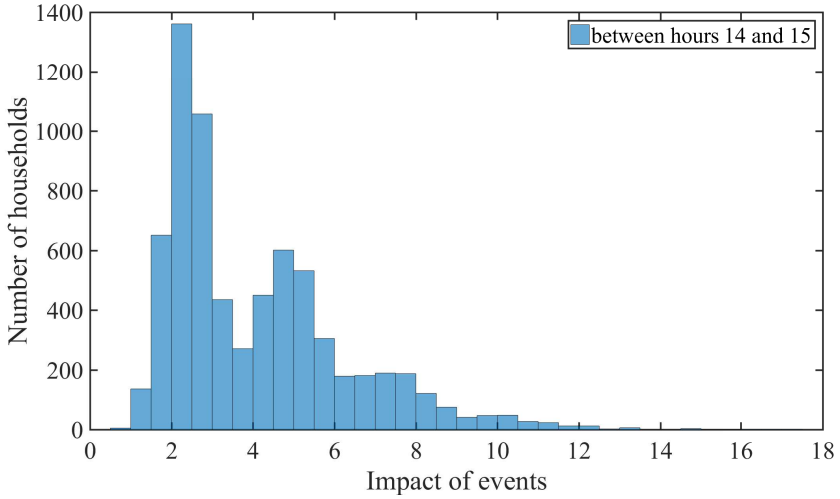


Figure 7.33: Human interaction events: distribution between 2pm and 3pm for 10,000 households, "zoom" to non-zero values.

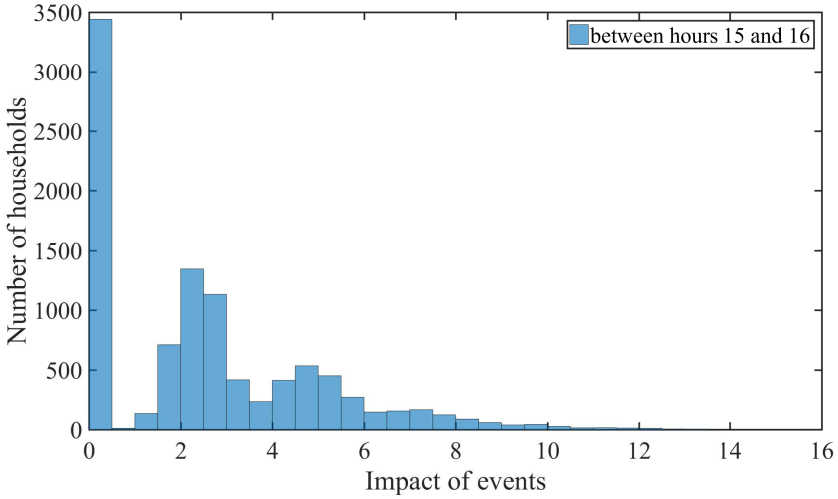


Figure 7.34: Human interaction events: distribution between 3pm and 4pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

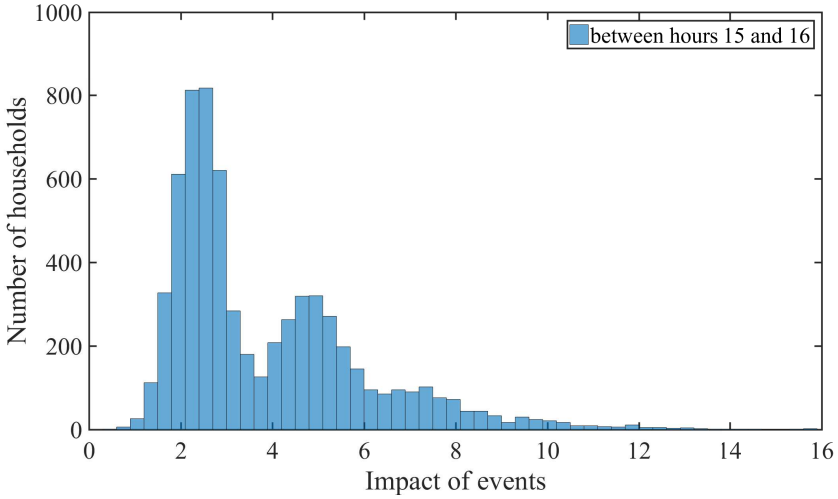


Figure 7.35: Human interaction events: distribution between 3pm and 4pm for 10,000 households, "zoom" to non-zero values.

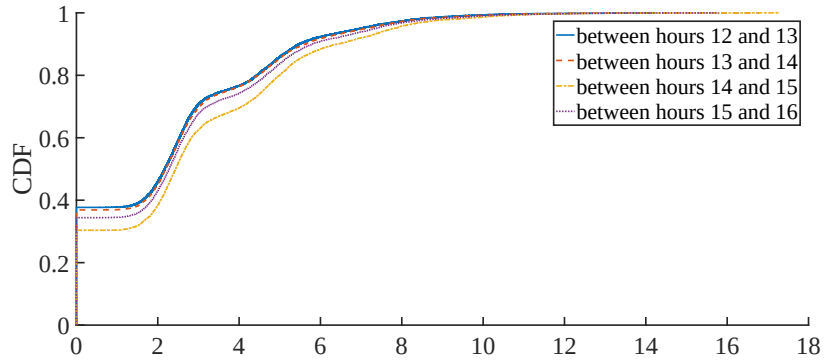


Figure 7.36: Cumulative distribution between 1pm and 4pm.

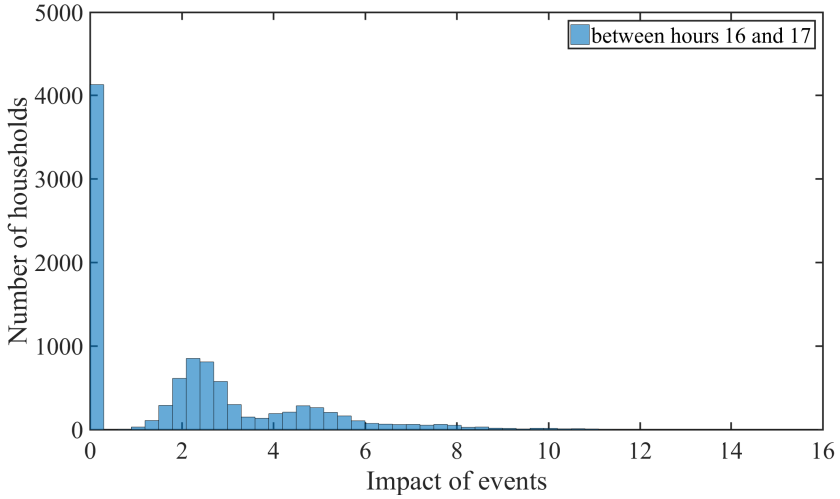


Figure 7.37: Human interaction events: distribution between 4pm and 5pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

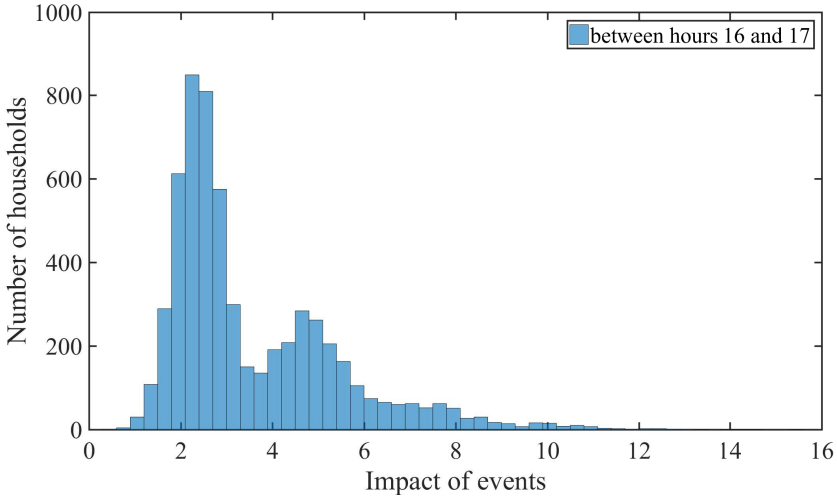


Figure 7.38: Human interaction events: distribution between 4pm and 5pm for 10,000 households, "zoom" to non-zero values.

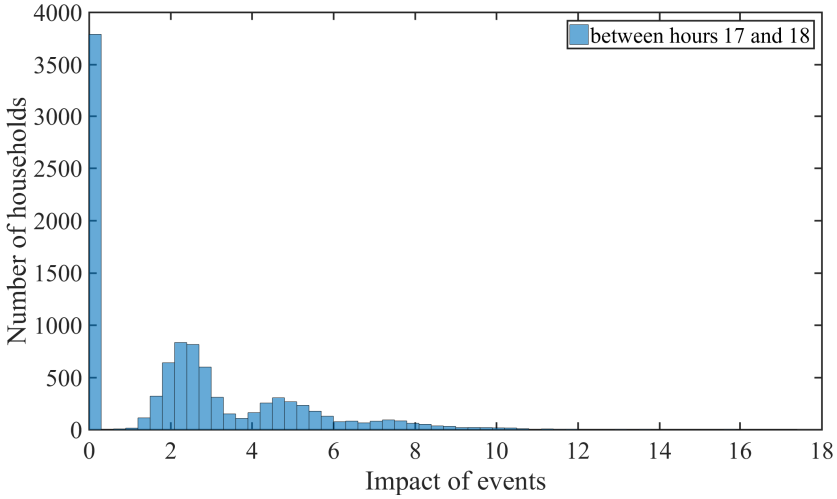


Figure 7.39: Human interaction events: distribution between 5pm and 6pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

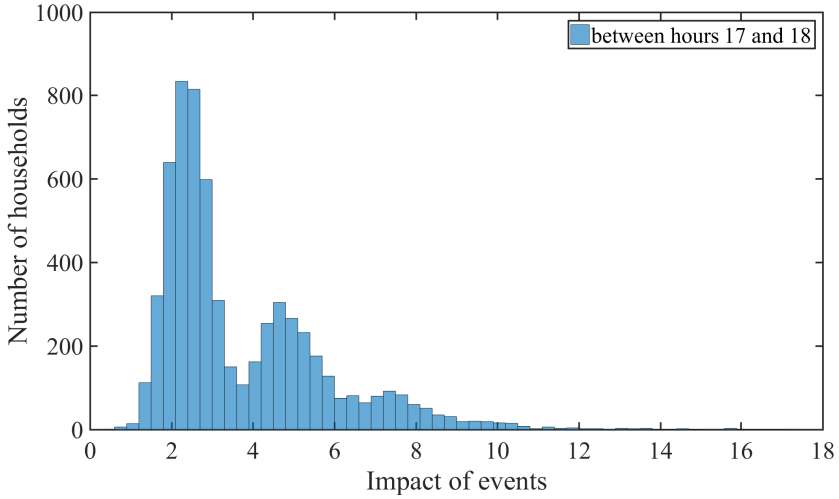


Figure 7.40: Human interaction events: distribution between 5pm and 6pm for 10,000 households, "zoom" to non-zero values.

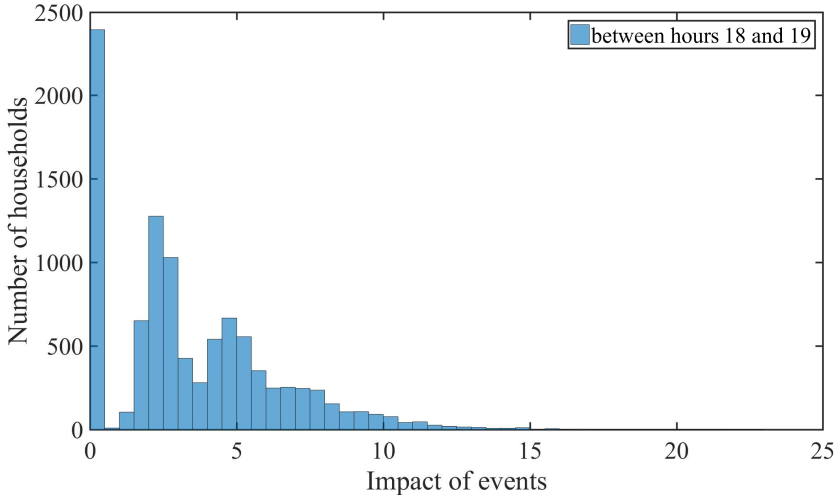


Figure 7.41: Human interaction events: distribution between 6pm and 7pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

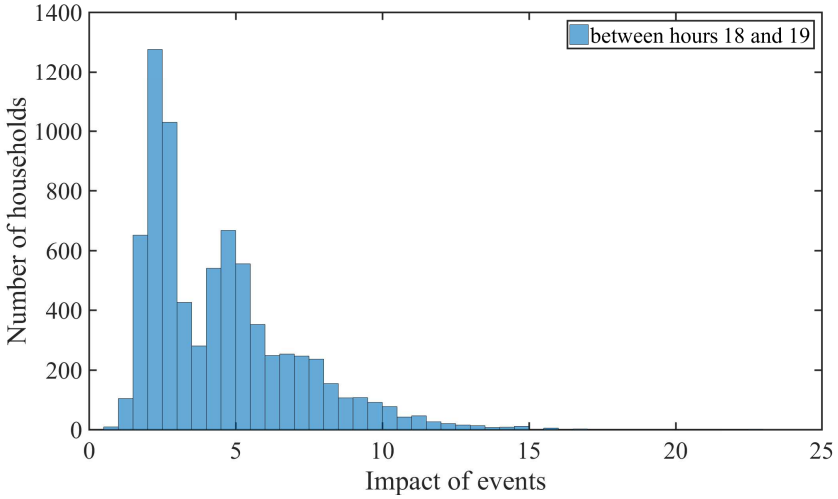


Figure 7.42: Human interaction events: distribution between 6pm and 7pm for 10,000 households, "zoom" to non-zero values.

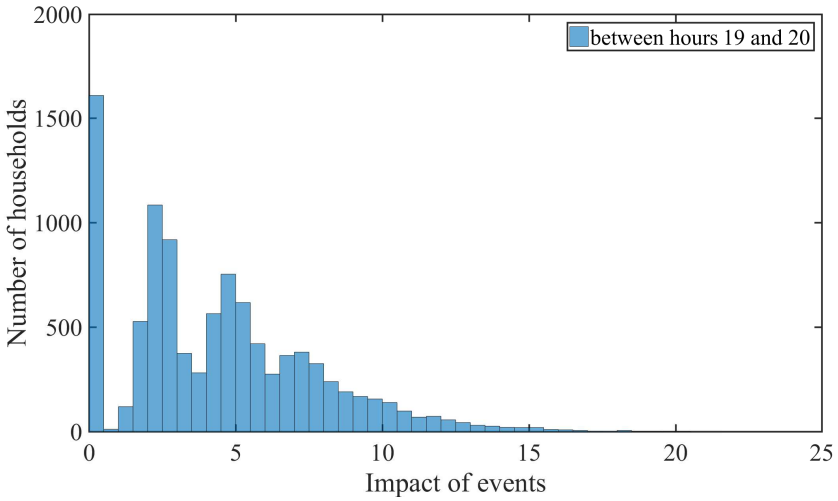


Figure 7.43: Human interaction events: distribution between 7pm and 8pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

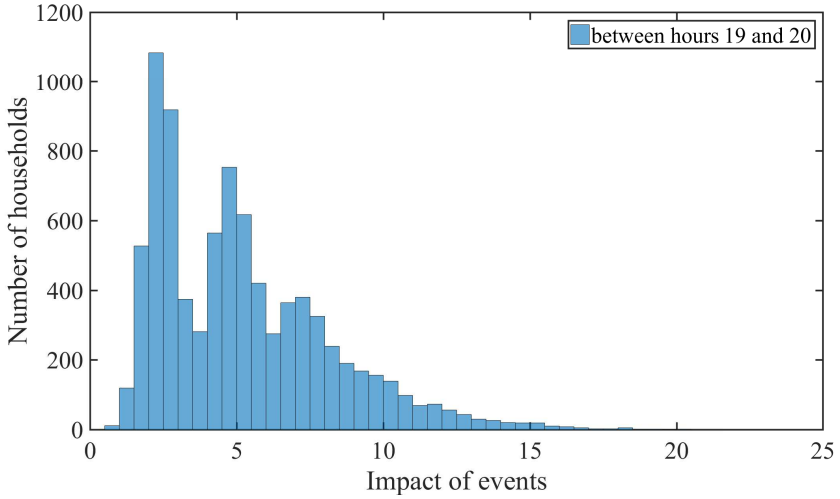


Figure 7.44: Human interaction events: distribution between 7pm and 8pm for 10,000 households, "zoom" to non-zero values.

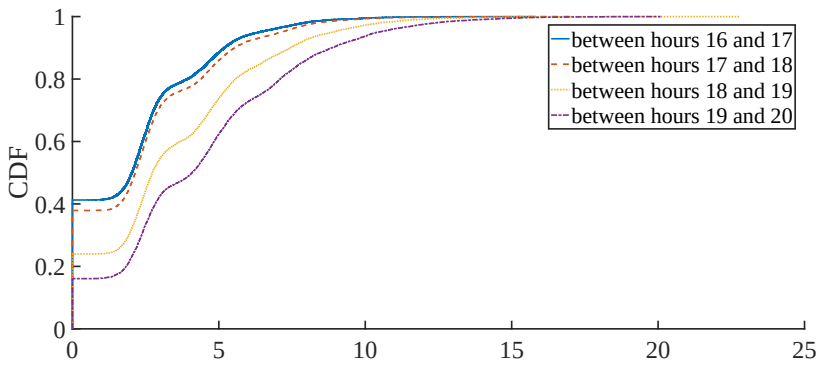


Figure 7.45: Cumulative distribution between 5pm and 8pm.

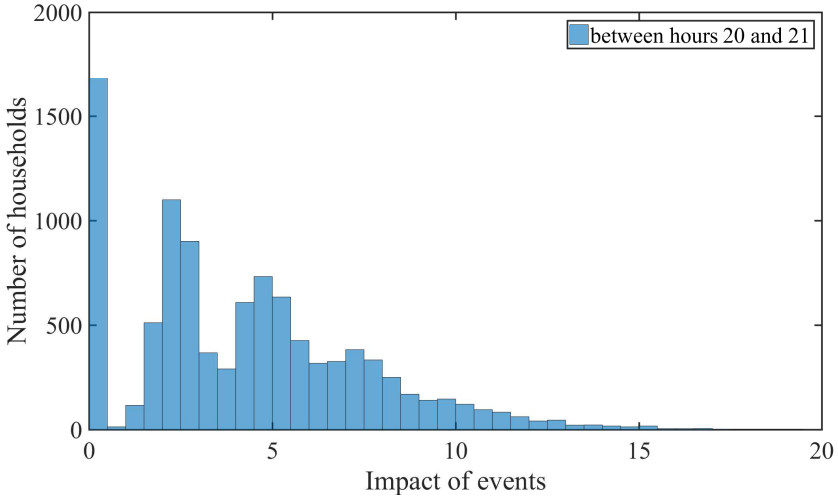


Figure 7.46: Human interaction events: distribution between 8pm and 9pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

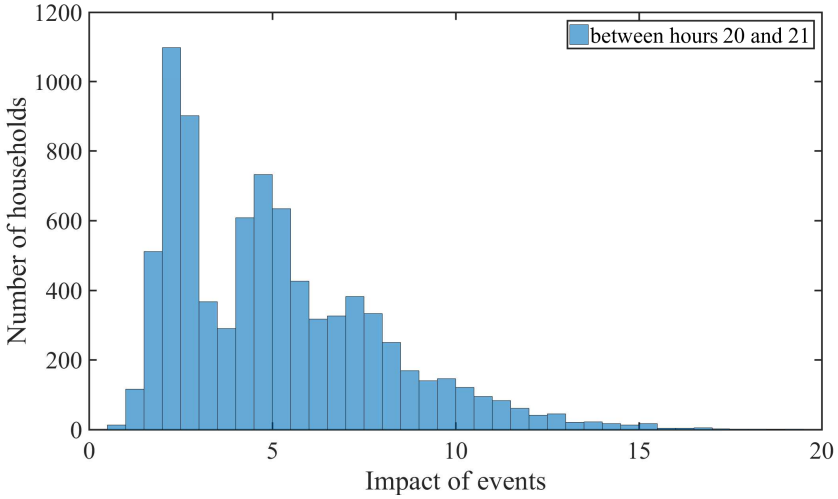


Figure 7.47: Human interaction events: distribution between 8pm and 9pm for 10,000 households, "zoom" to non-zero values.

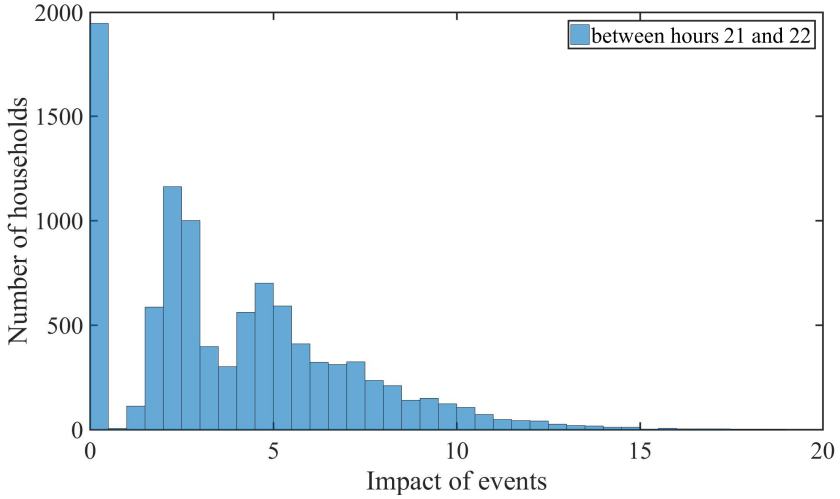


Figure 7.48: Human interaction events: distribution between 9pm and 10pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

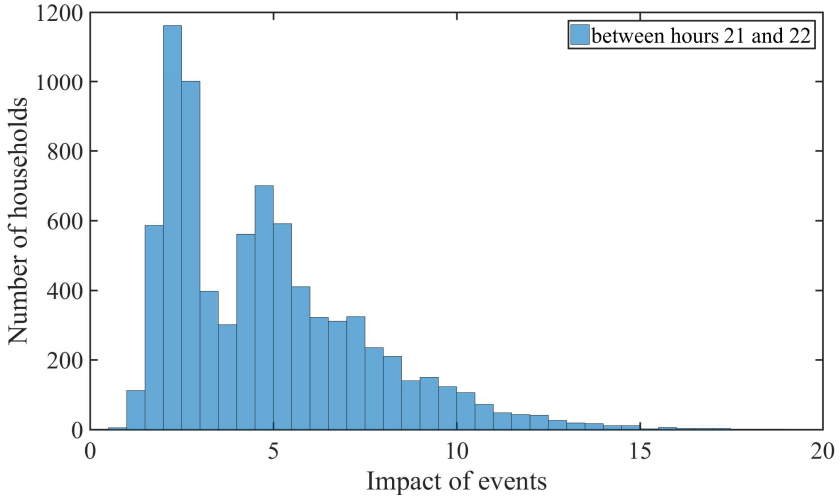


Figure 7.49: Human interaction events: distribution between 9pm and 10pm for 10,000 households, "zoom" to non-zero values.

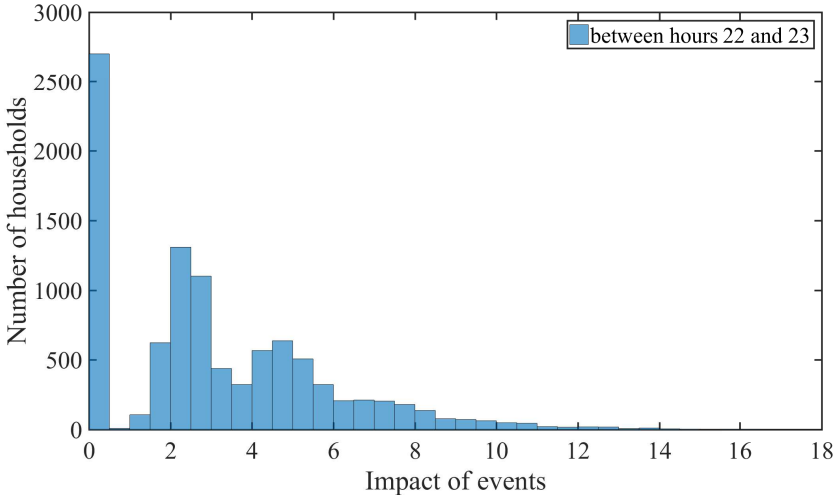


Figure 7.50: Human interaction events: distribution between 10pm and 11pm for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

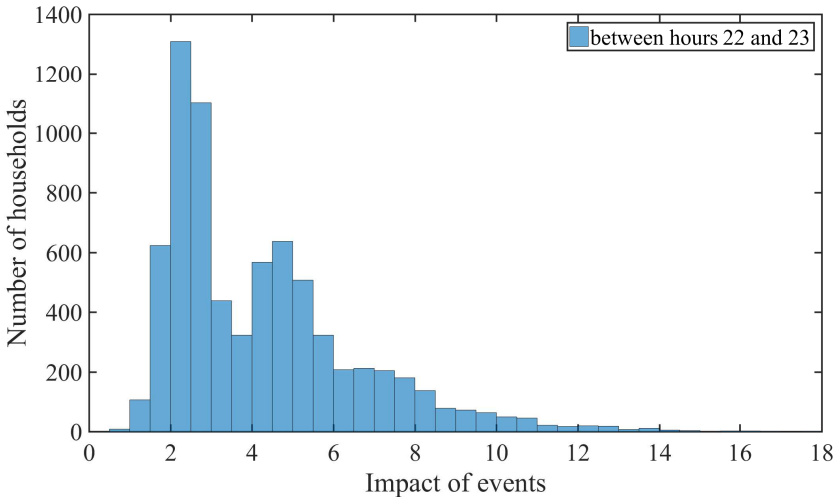


Figure 7.51: Human interaction events: distribution between 10pm and 11pm for 10,000 households, "zoom" to non-zero values.

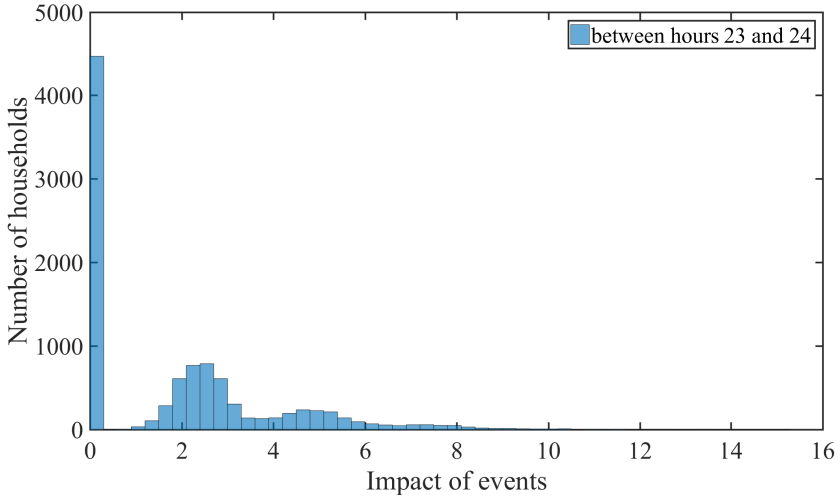


Figure 7.52: Human interaction events: distribution between 11pm and 12am for 10,000 households. Impact 1 equals to "door open for 12 sec at an angle of 90^o".

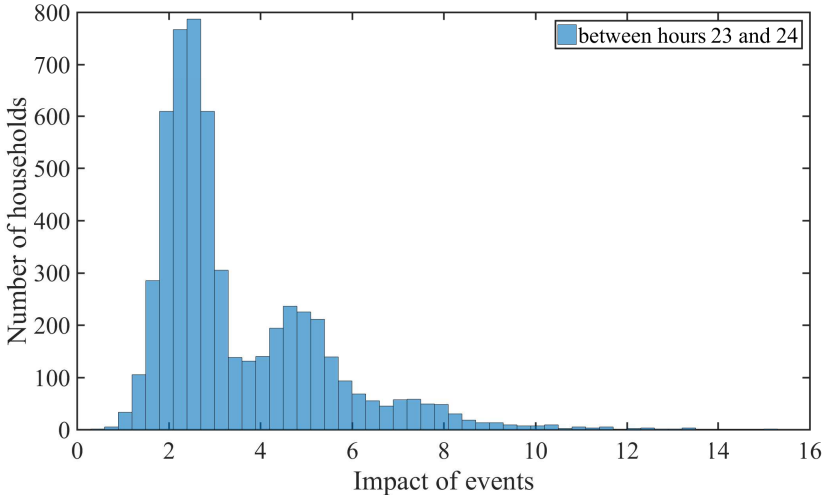


Figure 7.53: Human interaction events: distribution between 11pm and 12am for 10,000 households, "zoom" to non-zero values.

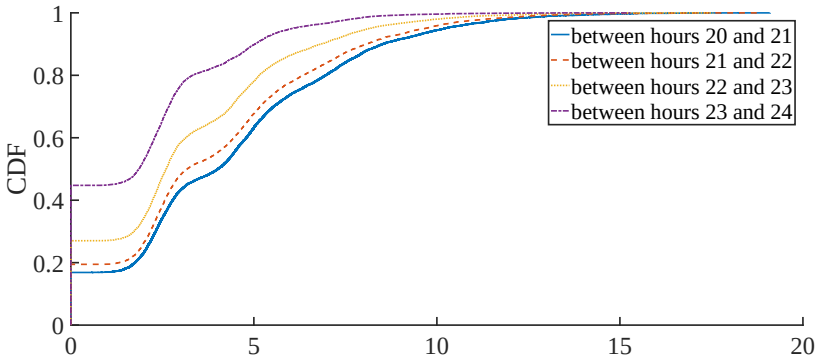


Figure 7.54: Cumulative distribution between 9pm and 12am.