Demand Side Management in the Smart Grid

Costanzo, Giuseppe Tommaso; Bindner, Henrik W.

Publication date:
2015

Document Version
Publisher's PDF, also known as Version of record

Ph.D. Thesis
Doctor of Philosophy

DTU Electrical Engineering
Department of Electrical Engineering

Demand Side Management in the Smart Grid

A Direct Load Control approach

Giuseppe Tommaso Costanzo (s120966)

Roskilde 2015

DTU Elektro
Department of Electrical Engineering
CEE - Center for Electric Energy and Power
Technical University of Denmark

Frederiksborgvej, 399
Building 776
4000 Roskilde, Denmark
Phone +45 4525 3500
evmn@elektro.dtu.dk
www.cee.elektro.dtu.dk

Preface

This thesis is prepared at the department of Electrical Engineering at the Technical University of Denmark in fulfilment of the requirements for acquiring the degree of Doctor of Philosophy.

Roskilde, October 30$^{\text{th}}$, 2015

Giuseppe Tommaso Costanzo (s120966)

Acknowledgements

Writing a Ph.D was for me a process of great personal growth and network development, accompanied by enrichment of knowledge and skills. For that, and everything else that led me writing this thesis, I am deeply grateful to my supervisor, Henrik W. Bindner, whose advice was precious to me in countless occasions, and whose passion for research motivated me through this enriching and challenging adventure.

Research meant for me not only spending much time alone, but also sharing thoughts and findings with colleagues from my department and from other institutions. This was possible thanks to the funding I received from the iPower consortium, part of the Danish Strategic Platform for Innovation and Research (SPIR), which not only covered my salary, but also many conference trips, summer schools, meetings, and an external research period, which I had the privilege to spend in two institutions.

For such enriching and stimulating experience at EPFL I am grateful to Prof. Jean-Yves Le Boudec, Prof. Mario Paolone, and their research groups. For inviting me to VITO and introducing me to the charming field of Reinforcement Learning, I am deeply grateful to Bert J. Claessens from VITO Energy/Energy Ville.

In the last three years spent at DTU campus Risø I enjoyed a dynamic, stimulating, and joyful working environment. I also enjoyed countless cakes, as the Danish tradition prescribes. For all this I thank all of my office mates, especially those that first convinced me to start a Ph.D. at DTU: Fabrizio, Anna, and Oliver. Thank you Esteban and Jacopo, for being true friends and contributing substantially to make me feel at home in Copenhagen. Thank you Eva, for always motivating everyone at office with your everlasting good humour.

I thank my family, which always believed in me and though me how to live and pursue my dreams with perseverance and optimism. Finally, I am particularly grateful to my girlfriend Marzia, the woman that enlightens my days, that shared with me the happy and stressful periods during the graduate school, without the support of which I couldn't reach this far.

To Marzia.

Summary

The global effort towards a decarbonised energy sector led to an increased share of renewable energies in the energy mix of most industrialised countries. Production from renewable resources as wind and solar is intrinsically intermittent, and often installed at medium and low voltage grid. This, together with the growing urbanisation, the increasing popularity of electricity-based climate control systems and electricity-based private transportation, challenges the distribution systems operators to run the low voltage grids with small safety margins.

Depending on the customer and the specific application, a portion of the electricity demand (or production) is flexible, and can eventually be used to support the system stability. Such practice is called Demand-side Management (DSM), and flexible units are called Distributed Energy Resources (DERs), if they only produce energy, or Demand Side Resources (DSRs), in the case they can also (or only) consume. This research investigates a specific control approach to DSM, called Direct Load Control (DLC). DLC is based on the assumption that the controlled units react in a predictable way to received control signals, and notify a failure if their operation is compromised.

In this research DLC is investigated at unit level and at aggregation level by means of model-based predictive control and model-free online control. Predictive control allows maintaining the flexibility of DSRs longer than online control, and allows pursuing quality of service (QoS) objectives towards both customers and SOs. However, it needs a model of the controlled units to optimise their operation over time, and field data to monitor the units' operation. On the other hand, the online control here investigated takes myopic decision based on field data, without the need of unit model. Therefore it does not guarantee optimality in the long run.

Two approaches are investigated within predictive control: the first is based on grey-box modelling and quadratic Model Predictive Control (MPC), while the second is based on Artificial Intelligence modelling and gradient-free optimisation. This research contributes to the first approach with a novel coordination scheme for a cluster of DSRs, which innovates with respect to other solutions proposed in literature by allowing independent design of the units local controllers and the use of heterogeneous control signals, such as continuous, integer or binary. The second approach foresees the application of self-learning models to predictive control. In this context, the contribution of this research is twofold: first it proposes the application of ensemble methods and decision trees to modelling of climate control systems, then it provides experimental validation of a combined Reinforcement Learning/in-domain

knowledge approach for DSM application.

The second approach to DSM in based on model-free online control. The proposed controller uses only field measurements and a generic domain knowledge to abstract the unit flexibility for DSM applications. It allows real-time control of power flow in single DSRs and, since there is no operation planning involved, decisions are taken online on the basis of field measurements. In this context, a novel approach is proposed to control thermostatically-controlled loads (TCLs) within a framework for control of micro grids by means of explicit power set points.

This thesis is submitted as a collection of the following papers:

- Paper A - Grey-box Modeling for System Identification of Household Refrigerators: a Step Toward Smart Appliances. 4th IEEE Youth International Conference on Energy, 2013. DOI: 10.1109/IYCE.2013.6604197. Authors: Giuseppe Tommaso Costanzo, Peder Bacher, Fabrizio Sossan, Mattia Marinelli, and Henrik Madsen.

- Paper B - A Coordination Scheme for Distributed Model Predictive Control: Integration of Flexible DERs. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2013. DOI: 10.1109/ISGTEurope.2013.6695474. Authors: Giuseppe Tommaso Costanzo, Oliver Gehrke, Esteban Morales Bondy, Fabrizio Sossan, Jacopo Parvizi, Henrik Madsen, and Henrik W. Bindner.

- Paper C - Performance Assessment of Aggregation Control Services for Demand Response. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2014. DOI: 10.1109/ISGTEurope.2014.7028779. Authors: Daniel Esteban Morales Bondy, Giuseppe Tommaso Costanzo, Kai Heussen, and Henrik W. Bindner.

- Paper D - Experimental analysis of data-driven control for a building heating system. Submitted to Elsevier Sustainable Energy, Grids and Networks (SEGAN) on July 14, 2015. Available at http://arxiv.org/abs/1507.03638. Authors: Giuseppe Tommaso Costanzo, Sandro Iacovella, Frederik Ruelens, Tim Leurs, and Bert J. Claessens.

- Paper E - Electric Space Heating Scheduling for Real-time Explicit Power Control in Active Distribution Networks. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2014. DOI:10.1109/ISGTEurope.2014.7028798. Authors: Giuseppe Tommaso Costanzo, Andrey Bernstein, Lorenzo Reyes Chamorro, Henrik W. Bindner, Jean-Yves Leboudec, and Mario Paolone.

Three other publications are relevant to this thesis, but not directly related to, therefore they are mentioned below but not included in the thesis:

- Paper F - An Experimental Study on Load-Peak Shaving in Smart Homes by Means of Online Admission Control. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2012. DOI: 10.1109/ISGTEurope.2012.6465658. Authors: Giuseppe Tommaso Costanzo, Anna Magdalena Kosek, Guchuan Zhu, Luca Ferrarini, Miguel F. Anjos, and Gilles Savard.

- Paper G - A System Architecture for Autonomous Demand Side Load Management in Smart Buildings. IEEE Transactions on Smart Grid, Vol. 3, no. 4, December 2012. DOI: 10.1109/TSG.2012.2217358. Authors: Giuseppe Tommaso Costanzo, Guchuan Zhu, Miguel F. Anjos, and Gilles Savard.
- Paper H - Power Admission Control With Predictive Thermal Management in Smart Buildings. IEEE Transactions on Industrial Electronics, Vol. 62, no. 4, April 2015. DOI: 10.1109/TIE.2014.2387091. Authors: Jianguo Yao, Giuseppe Tommaso Costanzo, Guchuan Zhu, and Bin Wen.

Resumé

Den globale indsats for en kuldstoffri energisektor har ført til en øget andel af vedvarende energi i energimixet hos størstedelen af de industrialiserede lande. Vanskelighederne ved at balancere forbrug og produktion samt ved at overholde transmissions og distributionsnettets effektbegrænsninger tvinger driften af elnettet ud i kritiske områder for at bevare en pålidelig elforsyning af høj kvalitet.

For at imødekomme balanceringen og for at afhjælpe eventuelle flaskehalse i transmission- eller distributionsnettet kan fleksibelt forbrug- og produktion forskydes tidsmæssigt eller geografisk. Dette koncept kaldes for Demand-side Management (DSM). Dette forskningsprojekt undersøger en specifik reguleringstilgang til DSM kaldet Direct Load Control (DLC) som metode til at aktivere forbrugsfleksibilitet (Demand-side ressources, DSRs, på engelsk). Reguleringstilgangen består af to forskningsakser: prædiktiv styring og online styring.

To tilgange undersøges indenfor prædiktiv regulering: den første er baseret på greyboks modellering og kvadratisk modelprædiktiv regulering (model predictive control, MPC, på engelsk), mens den anden er baseret på kunstig intelligens modellering og gradient-fri optimering. Denne forskning bidrager til den første tilgang med en ny koordineringsmetode for en klynge af DSR, som innoverer på andre løsninger analyseret i litteraturen ved at tillade den uafhængige design af de lokale enheders styringsalgoritmer, og kan håndtere heterogene typer af styringssignaler: kontinuert, diskret og binært. Den anden tilgang omhandler anvendelsen af selv-lærende modeller for prædiktiv regulering, med anvendelse af ensemble metoder og beslutningstræer af termisk modellering af bygninger. En viderbygning til sidstenævnte er også præsenteret, denne bruger Extreme Learning Machines (ELMs) for regressioner, sammen med Batch Reinforcement Learning (BRL) for beregning af styringsbeslutninger.

Den anden forskningsakse handler om model-fri online regulering, som kun har behov for feltmålinger og en generelt domæneviden for at abstrahere enhedens fleksibilitet og anvende den til DSM. Da der ikke er nogen operationsplanlægning involveret, og fleksibiliteten afhænger af enhedernes tilstand og modtagne effektsetpunkter, er beslutninger taget i reel tid og den opdaterer fleksibilitetsinformation og effektsetpunkter beregnet til hver tidsskridt. I den forstand bidrager denne forskning med abstrahering af fleksibiliteten af termostatstyrede enheder (thermostatically-controlled loads, TCLs, på engelsk) og brugen af denne abstrahering indenfor et framework for mikro-grid regulering ved brug af eksplicite effektsetpunkter.

Contents

CHAPTER 1

Introduction

"Research is to transform money into knowledge. Innovation is to transform knowledge into money."

Geoff Nicholson, 3M, the father of post-it notes.

In the past twenty years the costs of wind and solar energy has been sensibly decreasing, allowing these two technologies to play a major role in the decarbonisation of the electricity sector. This goal has gained increasing importance on governments' agendas in order to reduce the environmental impact of human activities. At the same time, technologies as electric climate control systems (heat pumps, HVACS, and resistive space heating) and electric vehicles have gained popularity in many countries, especially in northern Europe. As result, increased uncertainty in energy production and energy consumption puts the System Operators (SOs) to face already challenges in order to maintain the electricity provision stable and of appropriate quality. In this context the Smart Grids, as the evolution of the actual electricity networks, promise to help tacking a number of these challenges.

The Smart Grid concept comes from the merge of an IT and control infrastructure with a traditional electric network, which functionality and reliability is enhanced by diverse data analysis and control techniques. Among the different applications of Smart Grids, this research focuses on intelligent energy consumption, and spans from modelling of demand-side resources (DSRs), design of control strategies for flexibility aggregation and operation, to field implementation and demonstration.

As it will emerge in the following chapter, DSRs are so diverse in type and capabilities that controlling their flexibility brings value to the System Operators and the end users. However, depending on the resources location, type, Demand Side Management scheme, and the required Quality of Service (QoS), some control schemes emerge to be more convenient than others. This research explores the main advantages and disadvantages of different techniques proposed in literature and aims at putting forward new perspectives and solutions to overcome some of the technical challenges that were found open.

Research scopes and thesis structure

This research explores three directions: classical model-based predictive control using grey-box physical modelling, Machine Learning model-based and model-free predictive control, and model-free real-time control of DSRs.

Chapter 2 provides an introduction to the topic of Smart Grids, from the political effort to reduce the CO_2 emissions, to the technical challenges associated to high share of renewable energies in the power mix. Being this research mainly focused on Direct Load Control, Chapter 3 provides an insight on such scheme and provides the reader more specific references.

Chapter 4 is dedicated to model-based predictive control, and starts with introducing the DSRs modelling approach. In this context, the first contribution consists in the thermal modelling of residential refrigerators. The second part of the chapter presents two predictive control schemes for DLC: Centralised Model Predictive Control (CMPC) and Distributed Model Predictive Control (DMPC). While the main advantage of CMPC is the high quality of service provided to the end users and the SO, this approach scales very bad for increasing number of controlled units and it requires centralised knowledge of units models. DMPC algorithms, on the other hand, scale much better since they are based on distributed and parallel optimisation. Two famous implementations of distributed optimisation require the local units controller to respect a specific mathematical formalism. In this context, this research introduces an innovative approach for distributed model predictive control, called Sequential DMPC (SDMPC), which allows an independent design of units controllers to the specific units' needs. This algorithm achieves units coordination by means of central information sharing without the need of an intelligent or mathematically complex coordination; it can be applied to heterogeneous types of units, and it handles continuous and discrete control signals. Performances of SDMPC and CMPC architectures are compared and assessed with respect to the quality of service for end users and SOs. However, grey-box system modelling often requires significant time and effort, on top of in-domain system knowledge.

Chapter 5 presents an application of Artificial Intelligence to system modelling in the view of limiting as much as possible the modelling effort and cost. In this context, the innovation brought by this research consists in applying Extremely Randomized Trees (ERT) as AI technique to building thermal modelling and the use of Dynamic Programming for computing the optimal closed-loop control policy. Even though ERT benefits of high simplicity, experimental tests enlighten some limits of this technique, such as poor scalability, moderate regression performances, and relatively long training period. To overcome these issues, in the second part of Ch. 5 we introduce the Model-assisted Batch Reinforcement Learning (MABRL). This alternative approach aims at tackling the problem of regression performance and scalability by substituting the ERT with Extreme Learning Machines (ELMs), and the problem of large learning time by combining standard Fitted Q-iteration (FQi) with the support of ELMs.

In Chapter 6 we finally focus on model-free real-time control of DSRs. A system for real-time control of active distribution networks by means of explicit direct power

references, called COMMELEC [Ber+14a], is presented. This system, developed during a visiting research period at EPFL (Switzerland), generated promising results in simulations, and contributes in the same direction by abstracting the flexibility of a generic thermal load and enabling it to control within COMMELEC.

Chapter 7 summarises the topics presented and the contributions, it outlines the final considerations and presents a perspective on future research. Finally, the two appendices include the papers collection and some insights on selected topics.

CHAPTER 2

Background

This chapter outlines the background information relevant to this research. We start by presenting the main drivers to the international effort in reducing CO_2 emissions and decarbonising the power sector; then, our focus shifts towards the challenges of a decarbonised electricity sector. These challenges can be tackled with the help of Smart Grids which, among a number of advanced services, offer the control of a part of the energy consumption. This technology is called Demand Side Management (DSM), and it is the core topic of this research. This chapter includes a classification of control architectures for DSM programs, and introduces some relevant DSM research projects.

2.1 Governmental strategic plans and main drivers for clean energy

United Nations. On September 23, 2014, leaders from more than 40 countries, 30 cities and several corporations met at United Nations in New York for the Climate Summit 2014 in order to agree on a large-scale joint effort to cut the greenhouse gas emissions and transversally improve the efficiency in sectors as: transportation, lighting, appliances, buildings and district energy systems. As outcome, the Global Energy Efficiency Accelerator Platform was launched in order to double the global rate of increase in energy efficiency by 2030. Indeed, the main driver to such interest in efficiency is the increasing energy demand in both developing and already industrialised countries.

United States of America. In the United States, the energy efficiency became somehow popular in the political agenda in 1975 with the Energy Policy and Conservation Act, which focused more on security of energy supply rather than energy efficiency. The Act was signed into law by President Gerald Ford on December 22nd, 1975, and stated diverse provisions among which: the institution of the strategic petroleum reserve, the issue of the Corporate Average Fuel Reduction (CAFE) regulations, and the Energy Conservation Program for Consumer Products Other Than Automobiles.

After refusing to ratify the Kyoto Protocol in 1997, United States took a step forward to energy efficiency with the Energy Independence and Security Act of 2007.

This Act innovated the US energy policy in many areas, including: vehicle fuel economy, production of biofuels, and improved energy saving standards for appliances and private and public buildings. Few years later, in 2009, the first aggressive step in favour of clean energy was taken, and on June 26th the American Clean Energy and Security Act was passed by the U.S. House of Representatives.

European Union The beginning of a common European energy policy was motivated as well by energy security issues. S. Lansdorf reports that initially the energy policies were exclusively defined at national level [Sus11], until September 17th, 1974, when the Council of European Communities adopted the *Resolution concerning a new energy policy strategy for the Community*, which recognised the urgent need for a common energy policy and prepared joint targets for energy producers, consumers and markets [The75]. The international commitment for climate protection expressed in the *Earth Summit* in Rio, 1992, and in the ratification of the *Kyoto protocol* in 1997 was followed by the EU with the *Treaty of Amsterdam*, 1999, and the *Treaty of Nice*, 2003, which prepared the background for the first properly-called *EU Energy Action Plan*, issued in March 2007. This plan laid down three core points: sustainability, security of supply, and competitiveness, together with quantifiable objectives for reduction of emissions, renewable energies, and energy efficiency: the *20/20/20 targets* (20% reduction of greenhouse gas emissions below 1990 levels, 20% of EU consumption matched by renewable resources, and 20% reduction of primary energy use compared with projected values). In December 2013 the European Climate Foundation (ECF) published the report *Roadmap 2050* [Eur13], which points out two considerations: 1) in order to reduce the greenhouse gas emissions from 80% to 95% by 2050 the actual trend has to be maintained at similar peace at least until 2030, and 2) a low-carbon power sector is economically affordable provided an effective and collaborative sharing of cross-border resources enabled by: cross-border infrastructures, active demand response, disinvestment in high carbon assets and investment in low carbon assets.

Denmark A European pioneer in the decarbonisation of human activities is Denmark, which requires all non ETS[1]sectors to participate in the total reduction of GHGs emissions and intends to become independent from fossil fuels by 2050. As part of this, oil for heating applications and coal are to be phased out by 2030, and electricity and heating supply must be entirely covered by renewables by 2035 [The13]. The recent *Energy Agreement*, published on March 22nd, 2012, secures a broad political commitment to energy savings and promotion of renewables: wind, biogas and biomass. Particularly relevant to this research is the commitment of the Danish government to ensure a share of around 70% of renewables in the energy mix by the year 2020, 70% of which being wind power.

[1]Emissions Trading System (EU ETS), such as transport (except aviation and international maritime shipping), buildings, agriculture and waste.

2.2 Increase of energy demand and decarbonisation of the power sector

In the report *World Energy Outlook 2011*, the International Energy Association (IEA) estimates an annual growth of world energy demand of 2.4% between 2009 and 2035, mainly driven by economic and population growth [IEA11]; China and India would cover half of this increase (Fig. 2.1). According to that, the energy production will also increase and the energy mix in the coming decades will be heavily influenced by technology development and deployment.

In the same report, IEA prepared three scenarios for the deployment of new technologies, which reflect three different levels of effort in the decarbonisation of the energy sector. The *New Policies Scenario* foresees a limited deployment of modern technologies (e.g. electric vehicles, and carbon capture and storage), while the other two scenarios, *450 Scenario* and *Current Policies Scenario*, relate respectively to: significant measures towards clean energy, and no additional measures with respect to 2010. Figure 2.2 shows the projection of the average global energy mix in 2035 with respect to the three scenarios, while Fig. 2.3 shows the foreseen share of different renewable resources for the new policies scenario.

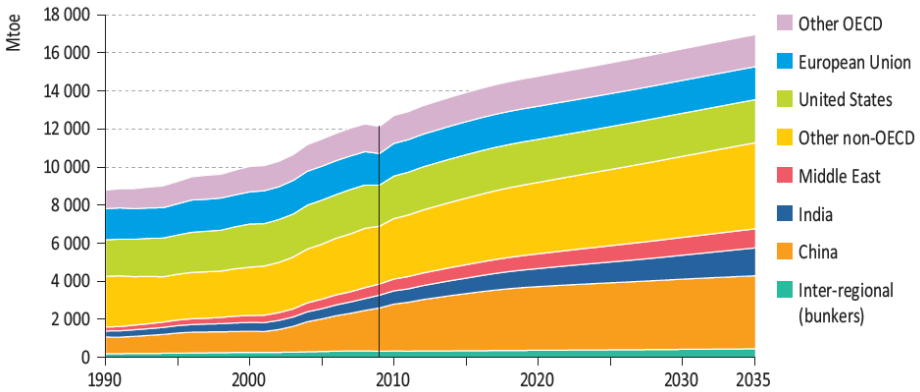


Figure 2.1: World Energy Demand in the New Policies Scenario.[2]

[2]Based on IEA data from World Energy Outlook 2011 © OECD/IEA 2011, IEA Publishing; modified by G.T. Costanzo. Licence: http://www.iea.org/t&c/termsandconditions/

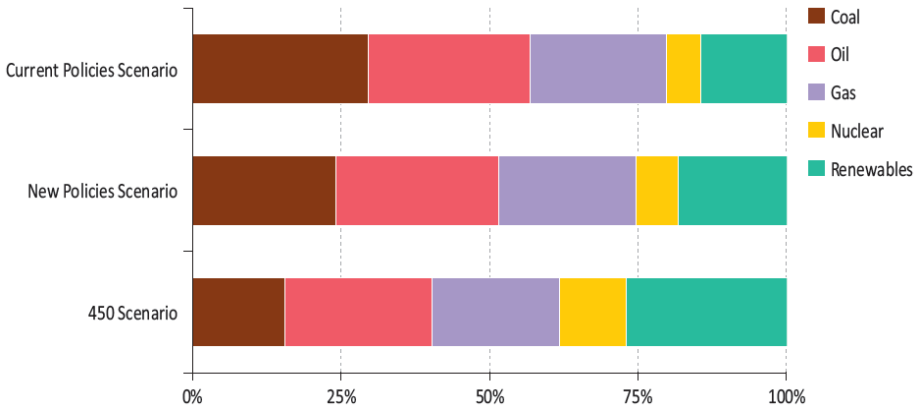


Figure 2.2: Share of energy sources in world primary energy demand by scenario, 2035.[3]

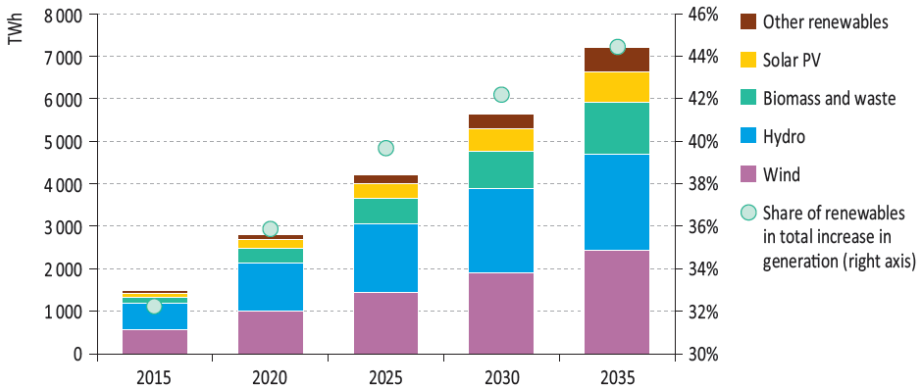


Figure 2.3: Incremental global renewable-based electricity generation relative to 2009 by technology in the New Policies Scenario.[2]

2.3 Challenges related to a decarbonised electric energy sector

Due to the an increased share of renewable energies in the power mix, wind and solar in particular, Transmission System Operators (TSOs) and Distribution System Operators (DSOs) are already facing the following challenges [Nor13]:

[3]Based on IEA data from World Energy Outlook 2011 © OECD/IEA 2011, IEA Publishing; modified by G.T. Costanzo. Licence: http://www.iea.org/t&c/termsandconditions/

- angle stability: high fluctuating production from renewables mines the ability of TSOs to maintain system synchronism and balance the torque of synchronous machines. System frequency has to be maintained between ±1% of the nominal value, which in Europe is 50Hz;
- voltage stability: high penetration of PV in the distribution grid mines the ability of the DSOs to maintain the line voltage around ±10% the nominal value;
- congestion management: the activation of balancing resources in some parts of the network, as well as excess of consumption due to high amount of Electric Vehicles being charged, may cause grid overload, with consequent blackouts or brownouts.

Yet, grid reinforcements can't be disregarded since the best production sites may be far from the centres of demand. Therefore, costs related to grid reinforcements are highly dependent on the region, and in Europe they account approximately for the 25% of the total transmission investments in the period between 2011 and 2035. Integration costs vary also according to the nature of the renewable source (e.g. biomass plants can be built closer to the consumption sites, with lower integration costs), see Fig. 2.4.

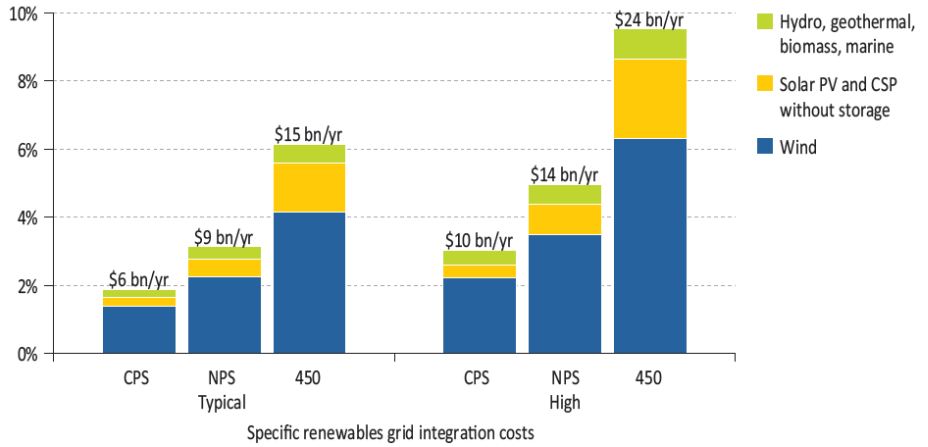


Figure 2.4: Renewables grid integration costs as a share of global transmission and distribution investment costs in the New Policies Scenario by integration cost, 2011-2035. [4]

Wind and solar provide, by nature, fluctuating production. Given that the balance between production, consumption, import and export of energy must be guaranteed

[4]Based on IEA data from World Energy Outlook 2011 © OECD/IEA 2011, IEA Publishing; modified by G.T. Costanzo. Licence: http://www.iea.org/t&c/termsandconditions/

at any time in a power grid, the increase of wind and solar share beyond a certain limit requires the shut down of traditional power plants and the increase of balancing reserves (which consist of plants that are enabled to rump up or down the electricity production very fast). Also curtailing PV and wind production may be necessary, with a consequent increased costs of balancing reserves. Furthermore, high penetration of renewables has profound impact on energy markets, which have to integrate new stakeholders (from large wind farms owners to small PV installations owners). The Smart Energy Demand Coalition (SEDC) points out that government legislation has major role in the market context, by facilitating the integration between energy markets and ancillary services markets [Coa14], and by reducing the market barriers to small players.

2.4 The Smart Grid and Demand Response

Active interaction between electricity producers and consumers improves efficiency, system reliability and lowers the cost of electricity supply. This and many others are the promises of the *Smart Grid*, an electricity grid that *"integrates all the branched customers' and producers' actions to distribute electric energy efficiently, sustainably, at low operating costs and safely"* [Bel09], and to do so it *"uses the information network to enhance the functioning of the electricity grid."* [PM09].

Demand Side Management (DSM) and Demand Response (DR) are two leading applications of Smart Grid technologies and are two sides of the same coin. DSM refers to the practice of modifying the energy consumption at demand side, typically low voltage grid, upon different contingencies in the power system, e.g. production volatility and grid congestion. On the other hand, DR refers to the change of users electric consumption in response to external incentives, e.g. energy price volatility, contracted payments, or low-carbon performance.

This research focuses on the control perspective of energy distribution, therefore on Demand Side Management. In this context, the System Operator can benefit from flexible consumption or production from Demand Side Resources (DSRs). DSRs are devices connected to the distribution grig that consume or produce energy, and which operation can be controlled in different ways. Note the difference between DSRs and Distributed Energy Resources (DERs); these latter refer exclusively to units that produce energy.

DSRs are crucial to the Smart Grid, because they provide services to both the end customers and the power system operators (SO): the customer perceives a value from the operation of a DSR (e.g. the dishwasher) which, at the same time, offers a service to the SO by shifting the washing cycle off from peak hours. PV installations equipped with an inverter enabled for curtailing the power production, HVACs equipped with a controller that schedules the building heating on a price basis, EVs that can compensate voltage drops in LV feeders by adjusting the charging schedule, all of these are examples of DSRs. End users can be flexible in the required service level, and DSRs can be flexible in the operation modes. All considered, DSM can

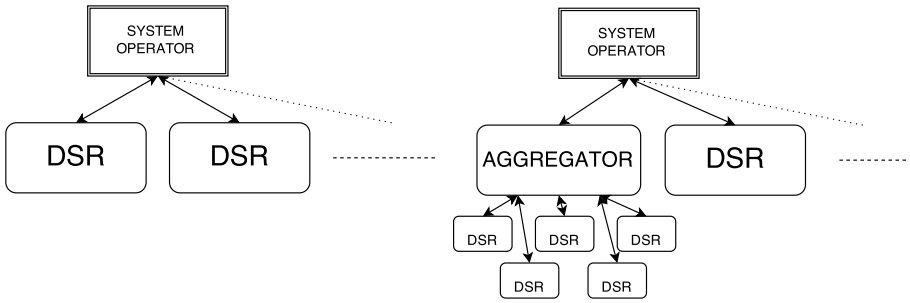


Figure 2.5: Demand Side Management without aggregation (left) and with aggregation services (right)

create considerable value to both SOs and end customers.

Note that, depending on the regulatory barriers, many DSRs may not be large enough to provide ancillary services directly to the System Operator. Therefore aggregators are needed between the DSRs and the SO. Each aggregator represents one or several pools of DSRs upon the SO and it manages the DSRs' flexibility.

Aggregators are therefore crucial to large scale deployment of DSM. The DSRs can activate their flexibility either occasionally or on regular basis depending on the contract with the aggregator and the type of DSM program, they can react to changes of electricity price, associated CO_2 emissions, or alternative signals and forms of compensation.

System Operators can benefit from flexible demand, whether they are vertically-integrated or not. In the case of vertically-integrated SO the interest in DSM is concentrated in one entity, whereas in liberalised energy markets the value of DR is perceived differently by different actors:

- Transmission System Operators (TSO): energy reserves of various type (primary, secondary);
- Distribution System Operators (DSO): congestion management, feed-in management, voltage support;
- Balance Responsible Parties (BRP): arbitrage, energy costs and penalties minimisation (portfolio optimisation);
- Aggregators: local DR management (deal with customers comfort), flexibility trade in markets (portfolio optimisation);
- Markets: intra-day balancing services, day-ahead capacity allocation (bulk energy consumption/production).

The first DR programs were investigated in 1979, when the interest was on the communication infrastructure for Automated Meter Reading (AMR), Automated Distribution Systems (ADS), and communication networks for direct load management

[Rus80]. In 1984 appeared the first field tests of DR based on dynamic electricity pricing, in which customers were billed for their consumption depending on the Time of Use (TOU) [HN84].

Today's research mainstream within DR is toward the automation of Demand Response, so to enable DERs to react in real-time to external signals issued by the SO or the aggregator. The interaction between DERs and other entities in the Smart Grid, e.g. other controllers, utilities or energy traders, can be characterised in terms of communication type and information exchange [Kos+13]. In this context, the DR can be classified by external behaviour (i.e. reaction to external signals) [Kos+13]: Indirect Load Control (ILC), Direct Load Control (DLC), Transactional Load Control (TLC), Autonomous Load Control (ALC) (Fig. 2.6).

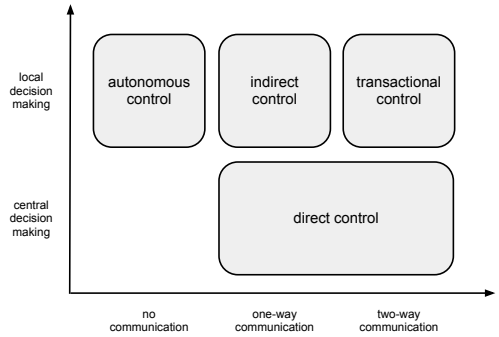


Figure 2.6: DER control matrix for demand side management. [5]

This classification can be used for comparing different DR systems, it can help in the definition of controller specifications, and it puts forward a point of view for controllers standardisation.

Direct Load Control (DLC). It is a technique based on issuing specific commands to controllable DSRs. The decision on each DSR operation is taken by an external controller that embeds the main intelligence and has a knowledge of the DSR status. A direct controller uses two-way or one-way communication to exchange information with a DSR . In case of one-way communication a DSR is obliged to follow control commands. The only deviation from expected behaviour can occur due to safety issues or unit failure. Vice versa, with bidirectional communication each DSR is obliged to acknowledge the received command. A DSR informs the external controller whether the control command can be executed or not, communicating also its status. Commands sent by the direct controller may vary depending on the DSR type and the communication protocol.

Indirect Load Control (ILC). It is a scheme for managing DSRs by issuing signals which may or may not affect the units operation. Indirect Load Control uses one-way communication, and the interested DSRs are not obliged to react to the

[5] Authorised reproduction from [Kos+13].

control signal or send any feedback. Thanks to that, this scheme is easily scalable and cheap to deploy. However, an aggregated model of the response of the DSR population is needed in order to appropriately design the ILC control strategy. In fact, it has been shown that price-based indirect control can cause consumption kick-back effects (e.g. see [SBN14]).

Transactional Load Control (TLC). This scheme is also called *market-based control* [KVM10], and it refers to a control strategy based on negotiations in a bid-based market. In transactional control DSRs are competing for one or more resources on an equilibrium market with the use of bids. After the end of the transaction DSRs can optimise their production or consumption with use of the equilibrium value determined by the market. The main goal of the transactional control is to distribute resources efficiently by taking into account correlated needs of different DSRs [Cle96]. A DER independently decides about the bidding amount and autonomously plan the control accordingly to the winning bid.

Autonomous Load Control (ALC). This scheme is based on measurements of frequency and/or voltage that are performed at unit level, in a way that each single unit can participate in frequency or voltage regulation independently. Due to the use of local measurements, this kind of control is characterised by a fast response and it does not need a communication infrastructure or higher-level control entities. DSR participating to DSM via Autonomous Load Control do not need any supervision, they work autonomously, and they can provide valuable services to the power system.

Demand Response allows reducing investments for grid reinforcements and peak generation by controlling the consumption as balancing resource for wind and solar generation, it reduces the need for spinning reserves (plants that are continuously running in order to supply power on short notice), bringing therefore financial and environmental benefits to the society.

Yet, benefits of DR are not limited to the system level. The Smart Energy Demand Coalition (SEDC) estimates that in 2013 business and private customers in U.S.A. earned over 2.2 billions of Dollars in direct revenues from trading flexibility in Demand Response programs [Coa14]. In the same report, SEDC concludes that comparable benefits could be achieved also in Europe, with consequent beneficial impact on local economies.

2.5 Some leading Demand Response research project

The European Commission, in its directive 2012/27/EU, grants strong support to DR by issuing specific directives [6]:

- Integration with the markets: *"Member States shall ensure that national energy regulatory authorities encourage demand side resources, such as demand*

[6] Directive 2012/27/EU of the European Parliament and of the Council of 25 October 2012, Art. 15.8

response, to participate alongside supply in wholesale and retail markets."

- Non discriminatory treatment: *"Subject to technical constraints inherent in managing networks, Member States shall ensure that transmission system operators and distribution system operators, in meeting requirements for balancing and ancillary services, treat demand response providers, including aggregators, in a non-discriminatory manner, on the basis of their technical capabilities."*

On the same line, the U.S. Federal Energy Regulatory Commission (FERC) issued on March 15, 2011 a final rule addressing the compensation for Demand Response in the organised wholesale energy markets, i.e. day-ahead and real-time energy markets.[7] Then, FERC and the Department of Energy (DOE) have jointly submitted in July 2011 to the Congress the *"Implementation Proposal for The National Action Plan on Demand Response"*, which objective is to *"implement the National Action Plan to achieve the nation's demand response potential."* [8]

As a result of joint effort from public and private players to study and understand the consequences of the evolution of power systems into Smart Grids, several research project have been granted in both E.U. and U.S.A., among which we find:

- DISPOWER (E.U.-2001/2005). This was the first project at E.U. level (2001-2005). Its main focus was on studying and supporting the transition of the European power system to a more market-based structure with efficient and safe integration of Renewable Energy Resources. The DISPOWER project produced a series of tools for planning, operating and forecasting distributed generation, as well as recommendations for European interconnection standards [DSS06].

- INTELLIGRID (U.S.-2004). Interoperability is a primary research theme within the IntelliGrid program. Directed by the Electric Power Research Institute (EPRI), this project is highly focused on integrating electronic devices and smart control together with the latest communication technologies in the operation and management of the distribution network [Ins13].

- FENIX (E.U.-2005/2009). The scope of this project is to conceptualise, design, and demonstrate a technical architecture and commercial framework to enable DER-based systems become the solution for the future cost efficient, secure and sustainable EU electricity supply system [Wel+09]. Being its main application toward Virtual Power Plants (VPPs)[9], FENIX defines de-facto an architecture

[7] U.S. Federal Energy Regulatory Commission, Docket No. RM10-17-000; Order No. 745; March 15, 2011.

[8] Report to Congress prepared by the staff of the Federal Energy Regulatory Commission (FERC) and the Department of Energy: *Implementation Proposal for the National Action Plan on Demand Response*; July, 2011.

[9] A VPP is a flexible representation of a portfolio of smaller generators and demands. It creates a single operation profile from a cluster of units and it is characterised by a set of parameters usually associated with a traditional transmission connected generator, such as scheduled output, ramp rates,

for aggregators, which are based on Direct Load Control toward DERs, but that operate with Transactional Load Control toward markets.

- ADDRESS (E.U.-2008/2012). Being focused mainly on active demand, this project aims at providing technical solutions to participation of small customers in energy retail and in services markets. The ADDRESS architecture is inspired by FENIX, and customers participation can be on voluntary basis or upon contract [Pee+09]. Between ADDRESS aggregators and customers there is an ILC scheme, whereas at customer level the control scheme is DLC.

- Grid4EU (E.U.-2011/2015). This project is led by a group of European DSOs, and aims at testing innovative system concepts and technologies to help removing some of the barriers to the smart grids deployment: technical, economic, social, environmental or regulatory barriers. It focuses on how DSOs can dynamically manage electricity supply and demand, which is crucial for integration of large amounts of renewable energy, and empowers consumers to become active participants in their energy choices.

- EcoGridEU (D.K.-2011/2015). This project's approach is based on real-time energy market, where DSRs respond to variable electricity prices. It extends the current electricity market to a shorter time horizon and to smaller assets, allowing small-scale DSRs to directly participate in the market. This approach improves the system operation by enabling to DR resources that would otherwise have been left unused.

- iPower (D.K.-2011/2015). This project studies different challenges related to the Smart Grid, from residential and industrial demand response to grid stability, user acceptance, and energy markets. The control schemes mainly investigated are DLC, ILC, and TLC, with focus on modelling, control, and proof of concept via demonstrations. Being a broad project, its main goals are : 1) describe the flexibility potential for relevant DSRs and propose control and aggregation schemes from small to large size units; 2) assess the value of flexibility for DSOs and TSOs, and the entity of investments to access flexibility; 3) demonstrate the investigated solutions and promote innovation; 4) develop and demonstrate flexibility services and products for TSOs and DSOs; 5) provide an in-depth evaluation of social and market value of mobilising flexibility.

The presented projects have in common a traversal approach to Smart Grids, which includes socio-economic evaluations and analysis of regulations. It seems that regulatory barriers are among the most difficult to overcome, and they are perceived as first-priority non-technical difficulty to overcome [Coa14]. For an extensive survey on European Smart Grid projects, the interested reader can refer to [Gio+12; THL10],

voltage regulation capability and reserves. Furthermore, as the VPP also incorporates controllable demands, parameters such as demand price elasticity and load recovery patterns are also used for characterisation of the VPP [SML08].

whereas for an insight on Demand Response projects and potentials in the United States, [Ele09; CGK10] are good starting points. Below, some projects are presented and classified according to the control schemes: DLC, ILC, TLC, ALC.

Projects within Direct Load Control

DLC can often, but not exclusively, be related to a Virtual Power Plant (VPP), which is a flexible representation of a group of responsive DERs [SML08], and represents a pool of flexible units in a single portfolio of services to the power system. An example of a VPP using direct control is Power Hub [Vin11; AS], a research project developed and implemented by Dong Energy[10]. PowerHub is a VPP that controls consumption and production units using load and flexibility prediction tools, and it provides an interface to use the VPP in connection to commercial ancillary services and energy markets. In the United States, numerous successful DR programs have been deployed using Direct Load Control. At glance, Pacific Gas&Energy and Lawrence Berkeley National Laboratory launched in California the SmartAC Ancillary Services Pilot project, designed to assess the feasibility of providing spinning and non-spinning reserves to the California Independent System Operator's (CAISO's) Ancillary Services (AS) market [SBM09] by controlling air conditioners.

Projects within Indirect Load Control

Price-responsive controllers for power systems were proposed in [Sch+80], and have been a subject of different research and demonstration projects: [Ham+07], [NSN11], [LA05], EcoGrid EU [Jor+11]. An overview of the challenges of indirect control and related technical considerations is presented in [Pin+12], while state of the art solutions for ILC are presented in [SBN14], together with model-based price signal evaluations, forecast and hardware-in-the-loop simulations.

Projects within Transactional Load Control

PowerMatcher [Bli+10b] is an example of a demand response architecture using price-based transactional control scheme. A multi-agent platform is used to integrate and manage small and medium-sized DERs in the low voltage grid [Bli+10b], which is clustered into logical trees. Within each cluster runs a local energy pool-market, on which the DSRs can place their bids. After all bids are collected, the market responsible agent (Auctioneer agent) searches the equilibrium price. This is broadcasted to the DSRs, who allocate their production or consumption according to their bid. The PowerMatcher concept was demonstrated in the PowerMatching City, consisting in 22 Dutch homes located in a suburb of Hoogkerk (Netherlands) [Bli+10a]. Another example of the price-based transactional control architecture is investigated in the Olympic Peninsula Project [Ham+07], designed, implemented and conducted by The

[10] www.dongenergy.dk

Pacific Northwest National Lab [Ham+07]. An experimental testbed consisting of residential, commercial and municipal loads, and distributed generators was designed to deal with transmission and distribution congestion, load shifting, and peak shaving. The set-up for bidding was a local 5 minutes-based market.

Projects within Autonomous Load Control

Autonomously-controlled power system services, such as frequency and voltage support, have been a subject of research in various scientific papers, among which we find: [PL06], which presents an autonomous frequency controller for microgrids, [KGB12], which introduces a combined frequency and voltage approach, and [Dou+11], which investigates autonomous frequency control using refrigerators. ALC controllers differ from each others depending on the units for which they have been designed; however they have shown to suffer from system inflexibility, e.g. the *50.2Hz problem* with PVs in the German grid, and unforeseen units emerging behaviour.

Open issues

Each control scheme has its advantages and drawbacks. However, from this review emerges that the characteristic of controllers and aggregators that mostly affects the QoS is the predictive behaviour. A controller able to predict units reaction to specific signals and forecasted external conditions, can perform optimal resource allocation and ensure better QoS. However, predictive controllers and aggregation schemes are sensible to model accuracy and quality of forecasted external contingencies, so that they tend to be expensive to deploy. Predictive DSRs controllers and aggregators can be based on DLC, ILC, or TLC. On the other hand, non-predictive controllers are cheaper and do not require units modelling. However, they suffer from non-predictable units behaviour, and kick-back effects upon flexibility activation have been shown in [SBN14]. In this study, Sossan showed that the consumption of a cluster of TCLs can increase beyond a safety level after a load shading period. Where ALC schemes also suffer from kick-back effect [Dou+11], they are even cheaper than ILC solutions and proved to be provide valuable ancillary services to the SOs, in particular voltage support.

CHAPTER 3

Direct Load Control

This chapter aims at providing an introduction to Direct Load Control, a technique based on issuing specific commands to controllable Distributed Energy Resources. First, a taxonomy for DER flexibility, called BBB (Buckets, Batteries and Bakeries), is introduced to classify the technical aspects of units flexibility: power, energy, and running time, then quantitative considerations about flexibility are outlined on the basis of previous work found in literature. Therefore, an overview on DSO services that flexible DSRs can provide, with a particular focus on DLC-controlled units and DLC-based aggregators closes the chapter.

3.1 A taxonomy for flexibility: the BBB model

As the real-time balance between production and consumption has become increasingly challenging due to high penetration of renewables, the concept of flexibility has been gaining importance and popularity. Yet, no formal definition of flexibility exists. Petersen et al. introduced in [Pet+13] a way to define flexibility with respect to the unit ability to deviate its controlled operation from normal (non-controlled) operation. The BBB (Buckets, Batteries and Bakeries) taxonomy, depicted in Fig. 3.1, is based on the following operation constraints, which can be found in most of practical systems: 1) power capacity, 2) energy capacity, 3) energy level at specific deadline, 4) minimum runtime.

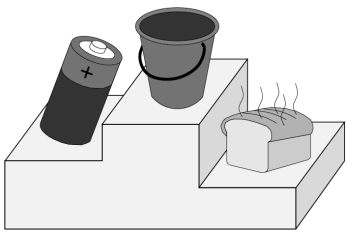


Figure 3.1: Buckets, Batteries and Bakeries (BBB) is a taxonomy for modelling flexibility in Smart Grids.[1]

[1] Authorised reproduction from [Pet+13].

The *Bucket*, being a power-constrained integrator, is an energy storage described by the following model [Ped+14]:

$$x_{k+1} = a_d x_k + T_s u_k \tag{3.1}$$
$$E_{min} \leq x_k \leq E_{max} \tag{3.2}$$
$$P_{min} \leq u_k \leq P_{max} \tag{3.3}$$

where k is the time step, T_s is the discretisation step, x is the energy level, u is the power, and a_d is the drain rate. Therefore, the Bucket is suitable to describe thermal loads whose temperature must be managed within a given comfort range, and whose dominant dynamics can be approximated with a first-order model.

The *Battery* model is analogous to Eq. 3.1-3.3, with the additional constraint that the energy level at a specified time has to have a specified level [Pet+13]:

$$x_N = E_N \ , \tag{3.4}$$

i.e. the battery needs to be charged or, depending on the application, discharged to a specific energy level E_N by a given time N. The Battery model is therefore suitable to describe Electric Vehicles and, in general, electrical storage.

Finally, according to [Pet+13], the *Bakery* model extends the Battery with the constraint that the process has to run in one continuous stretch at constant power consumption. Here we extend this concept by characterising the process power consumption with a predefined profile:

$$u_k = P_k \ , \ P_k \in \mathcal{P} \tag{3.5}$$

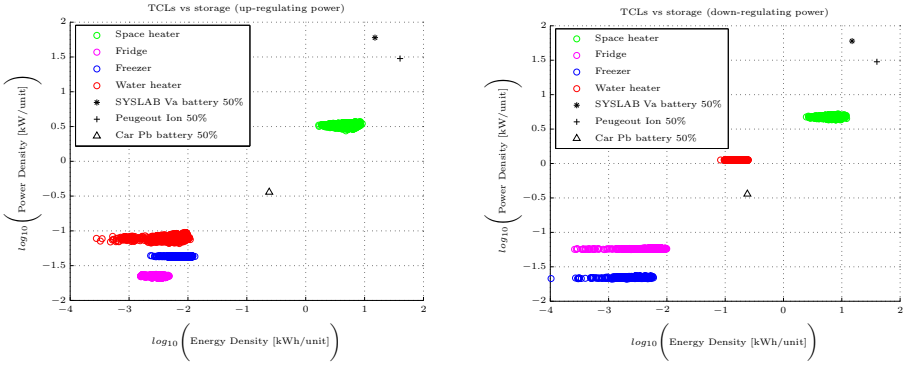
in a way that the Bakery model is suitable to describe the operation of some household appliances, such as dishwashers, washing machines, tumble dryers, microwave ovens, etc. This taxonomy gives a qualitative classification of flexible loads, from the most flexible type, the Bucket, to the least flexible, the Bakery. According to the BBB taxonomy, one can directly relate the load types described in [Cos+12b] as: *Regular*, *Burst* and *Baseline*, to respectively: Bucket, Bakery, and Bakery.

3.2 Flexibility potential

In [SBN14], Sossan proposes a quantitative study on flexibility contribution to up-regulating and down-regulating power of a population of DSRs via montecarlo simulation. Up-regulating power is provided when the production increases or consumption decreases, vice-versa down-regulating power implies a decrease in power production or an increase in power consumption. The units considered in [SBN14] are: electric space heating systems, fridges, freezers, water heaters, liquid vanadium batteries, EV lithium-ion batteries and EV lead batteries. This study analyses the units energy that can be shifted with respect to their baseline consumption.

Figure 3.2 shows the energy and the power that units can provide for up-regulation and down-regulation. As expected, it is shown that electric space heating systems are

the Thermostatic-Controlled Loads (TCLs) that have the highest power and energy flexibility, since the thermal capacity of buildings is usually higher than the one of refrigerators or water boilers.



((a)) ((b))

Figure 3.2: Different DSRs flexibility in case of up-regulating (a) and down-regulating (b) power provision. [2]

Interesting facts are reported in [SBN14]: first, TCLs energy flexibility is comparable to the one of EV Li-ion batteries, but it is dramatically affected by external conditions, e.g. by temperature. Seasonal effects on heating systems are quite strong. Second, batteries can contribute with power surge or power injection, whereas TCLs can only contribute with power surge or no-consumption. Third, TCLs whose thermostatic cycle is significantly asymmetric (e.g. electric water heaters) have consistently lower contribution to up regulation than to down regulation. This is due to the fact that these loads are most the time in off state. Figure 3.2 shows that the flexibility of fridges and freezers is one order of magnitude smaller than the flexibility of lead-acid batteries. Moreover, activating TCLs flexibility causes a synchronisation of their state, so that a kick back effect can be expected after activating their flexibility. Having explored a way to classify flexibility in this section, we present in the next section different modes of activating flexibility.

3.3 Flexibilitx operation modes

A direct controller can use two-way or one-way unicast communication to exchange information with DSRs. In case of one-way communication a DSR is obliged to follow the control commands and the only deviation from expected behaviour can occur due to violation of operational limits or failures. Conversely, in case of bidirectional

[2] Authorised reproduction from [SBN14].

communication each DER is obliged to acknowledge the received command and inform the external controller whether the received command can be executed or not. Commands sent by the external controller may vary depending on the DER type and communication protocol. Gehrke et. al. presented in [GI10] four direct load control schemes: deferred consumption, delta consumption, scheduled operation and direct power control (Fig. 3.3).

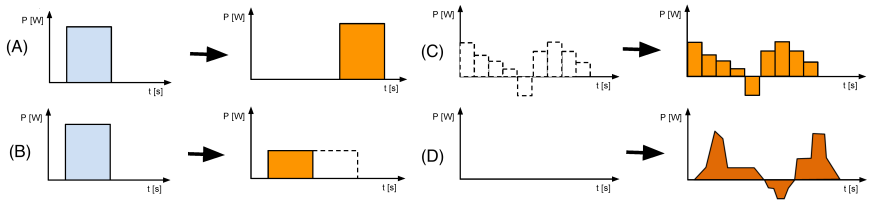


Figure 3.3: Operation schemes for DER direct control: (A) Deferred operation, (B) Delta operation, (C) Scheduled operation, (D) Direct power control.[3]

Load control schemes can be extended to cover both consumption and production, and they are defined as follows:

Deferred operation The consumption or production of a certain amount of energy is shifted in time (Fig. 3.3 (A)). Neither the amount of consumed or produced power, nor the duration of the operation is modified. The signal that is received by the DSR is of type Δt, which represents a required delay of operation.

Delta operation The amount of energy consumed or produced by the controlled DSR is decreased or increased by an offset of ΔP, called power difference (Fig. 3.3 (B)). Decrease in consumption might result in an increased operation time, e.g. thermal loads.

Scheduled operation The DSR is provided with an operation schedule s, consisting of time series of power set points and time stamps, where $s = \{(t_i, P_i)\}, i \in N$, (Fig. 3.3 (C)).

Direct power control At run-time the DSR is provided with a power set point, P (Fig. 3.3 (D)).

Each DSR is equipped with a local controller for low-level hardware control, which is responsible of performing the necessary actions in order to meet the requirements set by the external DR controller, i.e. the operation modes. Then, an aggregator, by using an optimised mix of the direct control schemes presented above, can operate a

[3] Authorised reproduction from [GI10].

unit or a cluster of units in order to provide services to the DSO[Nor+13], which are presented in the next section.

3.4 Use of flexibility: DR services to the SOs

DSOs' needs for flexibility are different from TSOs' ones. The TSO, being in general a monopoly, has the responsibility of maintaining the balance between production and consumption (frequency regulation), and to do so, it runs several markets for ancillary services. These services are generally provided by central and decentralised power plants (typically coal, oil or gas fired). As some plants are being shut down to accommodate more energy from renewables, ancillary services are predicted to increase in cost in the near future, so that TSOs may be interested in flexible loads as cheap ancillary services providers. On the other hand, DSOs have two duties: first guarantee power supply to customers, and second within $\pm 10\%$ of the nominal voltage. Therefore, DSOs have to avoid Blackouts (disconnection of feeders due to line overload) and Brownouts (voltage instability due to high local consumption or production). With this respect, DSOs may be interested in buying flexibility if it represents a cheaper alternative to grid reinforcements (install larger MV/LV transformers and higher-capacity cables); clearly, the break even point changes from feeder to feeder depending on the specific average load. As most of the electricity demand occurs at distribution level, DR has larger potential for DSO services.

This section emphasises the technical motivations for this research. An overview is provided on specific technical solutions for DSO services using DR, and their market implementation, which are more extensively presented in [Nor13]. First we introduce the major problems related to increased consumption and production that are affecting DSOs, and which services are needed to cope with them. Then, we introduce some Demand Response products that can be used to deliver the needed services.

Note the **distinction** between DR services and DR products: a **DR service** is intended as the way the flexibility is provided from one or multiple DSRs to the system operator, while a **DR product** is the way the flexibility is requested to DSRs in order to implement a service. For example: a service consists in limiting the load on a HV/MV transformer serving a 10kV feeder for a given time window, while a product allowing such service consists in shifting the consumed power of controllable loads served by that feeder. In this context, we see that a service is chosen depending on the specific DSO contingency, while the product is chosen depending on the specific type of unit (some units can be either off or operating at full power, while some others allow regulating the consumed power). Furthermore, recalling Ch. 3.3, a specific product requested to a unit or to a cluster of units can be implemented using a specific operation mode, or a combination of them.

3.4.1 DSO contingencies

DSOs manage the distribution grid, which is composed by: 10kV feeders, MV/LV substations, and 0.4kV feeders. Increased load in 0.4kV feeders reflects to an increased

load in 10kV feeders, and both phenomena, even though can be managed separately, influence each other. The critical contingencies presented below are analysed with respect to the 10kV distribution grid; same considerations can be extended to the 0.4kV feeders. Nordentoft in [Nor13] presents contingencies and services for DSOs related to load, and voltage.

1) Regular growth of electricity consumption. This case describes the increase of electricity consumption from regular appliances, as air conditioners and standard household appliances, and new appliances, as electric vehicles, and heat pumps or electric space heaters, above the capacity limit of 70% (Fig. 3.4).

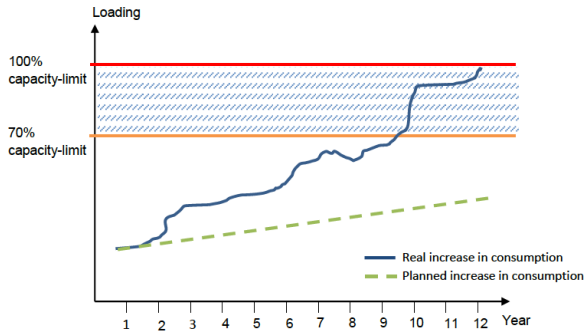


Figure 3.4: Increased load above the capacity limit.[4]

Considering that, as rule of thumb, the 70% capacity limit is used to determine whether a feeder needs upgrades or not, the number of hours in which the capacity limit is violated determines the viability of DR with respect to grid reinforcements. As usually the limit is passed only in few hours a day (Fig. 3.5), and not in all seasons, the DSO is likely to request flexibility to shift some load away from peak hours.

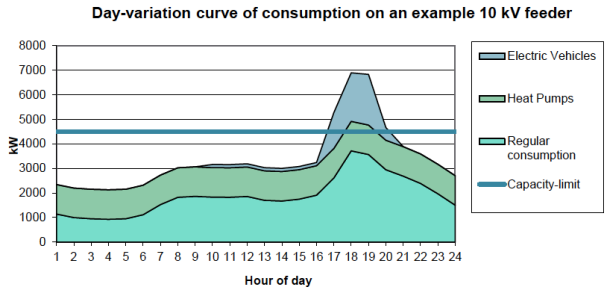


Figure 3.5: EVs and HPs load in peak hours.[4]

[4] Authorised reproduction from [Nor+13].

2) Active use of reserve capacity. This case occurs when a feeder uses its 30% safety capacity to supply another feeder that gets disconnected from the MV after a fault (Fig. 3.6).

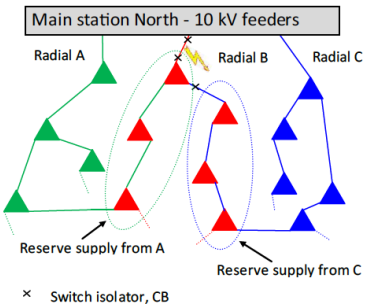


Figure 3.6: Example of a cable error close to the main station on feeder B. The sub grid of feeder B is split in two sub grids, which are supplied from neighbour sub grids A and C.[5]

The DSO may buy flexibility from interruptible customers in order to make sure that backup lines don't get overloaded in this emergency operation (Fig. 3.7). If this load reduction is absolutely reliable, DSOs can use the reserve capacity even up to 100% of line capacities.

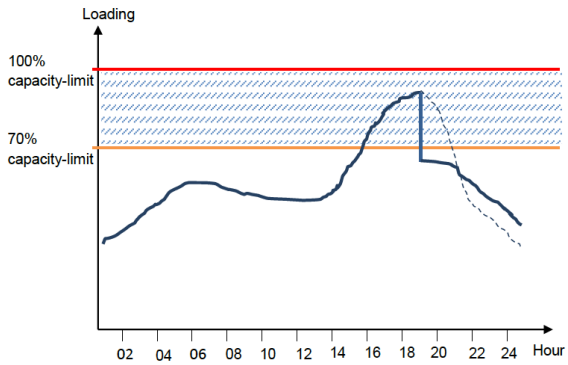


Figure 3.7: Example of load reduction in case of emergency operation.[5]

3) Congestion due to activation of regulating power. In some occasions, 10kV feeders may face congestions due to the activation of regulating power from the TSOs even if the regular load in the feeder was below the capacity limit (Fig. 3.8). If regulating power is needed to balance the power system, then such reserve should be activated.

[5]Authorised reproduction from [Nor+13].

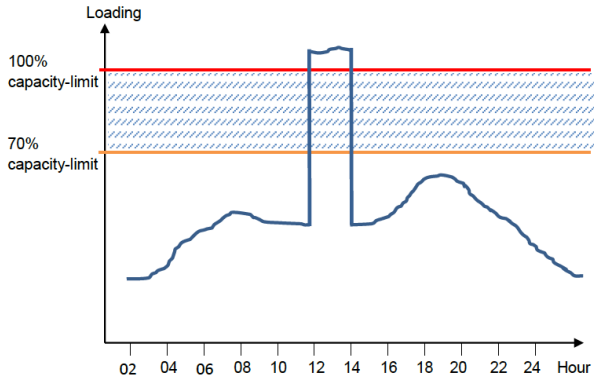


Figure 3.8: Example of congestion due to activation of regulating power by the TSO.[6]

This situation is likely to happen in high wind penetration scenario: instead of selling power at negative price to nearby countries, large water boilers in district heating systems are activated. Here the DSO is interested in coordinating the activation of regulating power in several feeders in order to keep the total load with respect to the TSO unchanged. This is an application of Virtual Power Plants.

4) High power flow caused by low energy prices. In a scenario where some large customers are allowed to bid in the day-ahead market, it may happen that the total load in a feeder exceeds the operational limits due to low energy prices. In fact, nowadays day-ahead markets rarely take into account grid constraints using nodal prices. Therefore, the DSO may be interested in limiting the total load in some feeders by activating the available flexibility.

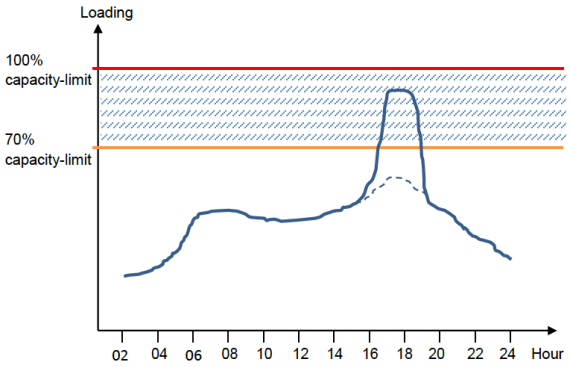


Figure 3.9: Line congestion due to low electricity price.[6]

[6] Authorised reproduction from [Nor+13].

5) Other LV contingencies: voltage instability and VAR compensation. The increasing penetration of distributed generation can put DSOs to face challenges related to voltage level at the 0.4kV distribution feeders. Abnormal power surge due to EV charging or power injection from PVs causes the voltage to drop or rise outside the predefined ±10% limits.

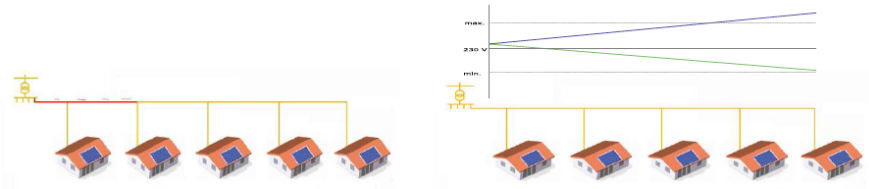


Figure 3.10: Line congestion and voltage instability.[7]

Note that voltage instability and grid congestion may also occur independently: with high consumption only at one end of the feeder the line voltage experiences a drop, while the MV-LV substation is not overloaded. Vice-versa, with high consumption at the beginning of the feeder the voltage is stable but the substation transformer may be overloaded. Demand Response can help the SO to cope with such challenges by offering the services discussed below, regardless of the value associated to trading flexible consumption in energy markets.

Reactive power compensation at national level could also become an issue due to increased local generation and its associated power converters. Daily variations of active consumption can cause variations on the consumed reactive power, so that DSOs may be interested in DR services for VAR compensation that are cheap than classic VAR compensation means (shunt reactors and capacitors, synchronous compensator (STATCOM) or Static Var Compensator (SVC)), e.g. using decentralised plants and DFIG turbines which, by controlling the magnetisation of the generators, can provide or absorb reactive power to the grid.

3.4.2 Demand Response products

This section presents some Demand Response products that can be combined in order to deliver the services presented in the previous section. Recall the DR product is the way the flexibility is requested to DSRs in order to implement a service. Then the DSRs can operate in different modes (Ch. 3.3).

1) Powercut Planned and Powercut Urgent. Powercut planned is a product used for handling predictable load peaks by sending a request to specific units to switch off during the peak hours. Even though this product does not ensure that

[7] Authorised reproduction from [Nor+13].

the capacity limit is respected at all time, if the non-controllable load is somehow predictable and the amount of flexible power is sufficient the DSO should be able to operate within the safe power band.

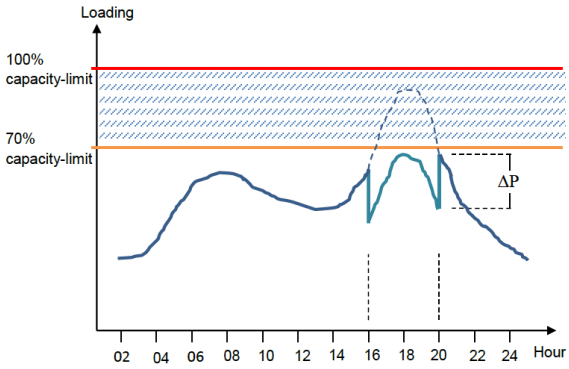


Figure 3.11: DR product: Powercut.[8]

In a similar way, the Powercut urgent is delivered on event basis instead of schedule basis, when the DSO requests it as a critical situation arises unexpectedly.

2) Power Reserve. This service consists in shutting down a portion of the load in a feeder on request, enabling the DSO to use the reserve capacity, normally used for supporting faulted feeders, for normal everyday operation. However, the DSO would rarely request this product, as it is needed only when the feeder load is above the capacity limit and a neighbour feeder has to be supplied after a fault.

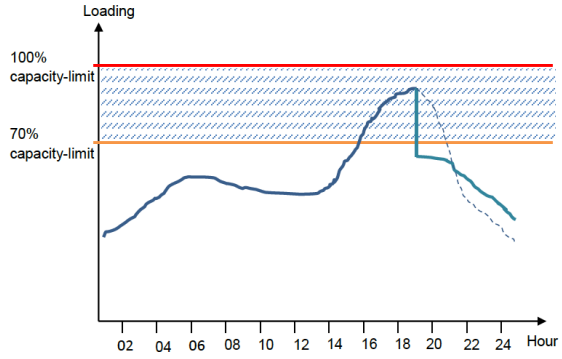


Figure 3.12: DR product: Power Reserve.[8]

[8] Authorised reproduction from [Nor+13].

The difference with the Powercut product stays in the duration of the activation. In Power Reserve there is no predefined activation time.

3) Powercap. This product guarantees that the power flow in the contracted feeder stays within a given limit. The aggregator or a single DSR is notified by the DSO with the maximum load allowed in the feeder and, by using power measurements at the substation, it regulates the consumed power in order to respect the given limit. This product helps in compensating also the consumption of non-controllable loads and, in the case that multiple aggregators jointly deliver this product, an outer control loop can be used to assign the share among the different parties.

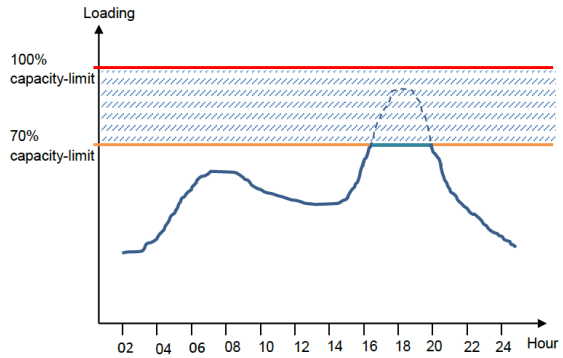


Figure 3.13: DR product: Powermax.[9]

4) Powermax. This product is somehow similar to Powercap, but the DSO does not have any guarantee that the feeder load stay within the capacity limit.

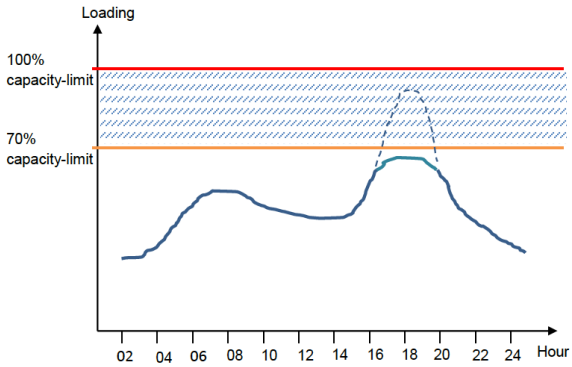


Figure 3.14: DR product: Powermax.[9]

[9]Authorised reproduction from [Nor+13].

With Powermax the aggregator only ensures that his DSR portfolio respects the contracted consumption (or production) limit, therefore no feedback loop with power measurements at the feeder substation is needed. Powermax has similar technical interest as Powercut, and also in this case the DSO should carefully compute the amount of load to ask the Aggregator(s) to shed. However, Powermax can probably be cheaper for the Aggregator since this last can dynamically (and optimally) allocate the load shed among the resources in its portfolio.

3.5 Conclusions

In this chapter we introduced the Direct Load Control as control scheme for Demand Response. The BBB (Bucket, Battery, Bakery) model has been presented as mean to classify flexibility of Demand Side Resources, after which a procedure to quantify flexibility of different types of loads has been reported from literature. Therefore we presented how the DSR can operate under Direct Load Control, by examining different operation modes. Finally the reader, after having been guided through the concept of flexibility, its quantification, and how it can be provided, is introduced to the contingencies of distribution networks for which DR services are valuable, and how these services are requested to DSRs or aggregators.

The next three chapters contain the background information for the respective attached papers, which present in details the control schemes and algorithms that are at the basis of DLC operation modes (deferred operation, delta operation, scheduled consumption, direct power control).

CHAPTER 4

Grey-box modelling of DSRs and Model Predictive Control

"When dealing with scientific or technical problems, it is always possible to distinguish, more or less well, three moments: the first is the summarisation of facts or phenomena under study and the formalisation of the problem with equations; the second phase consists in the manipulation of equations in order to obtain the elements that are accessible in experimental trials; the third phase, finally, is the one in which the elements deducted from theory are compared with experimental results."

Ercole Bottani [1]

This chapter presents an approach to Direct Load Control based on grey-box modelling and Model Predictive Control (MPC). The original contribution of this chapter is twofold: on the modelling side, a grey-box modelling approach is introduced for thermostatically-controlled loads as household refrigerators (Paper A), while on the control side an innovative coordination scheme for distributed MPC, called Sequential Distributed Model Predictive Control (SDMPC), is introduced (Paper B).

[1] Ercole Bottani (1897-1978), professor at Politecnico di Milano. Thanks to his intuition and formation as technician, he taught electrotechnics with a different approach, especially from his colleagues from the physics department. He started from the voltage and current measurements, and from the most simple electric circuits to arrive to the complicated laws of physics. He used to say: *"The Duomo must not be built from the Madonnina!"* (The *Madonnina* is the statue of the Virgin Mary at the top of Milan Cathedral, in Italy.)

4.1 Modelling of Demand Side Resources

Modelling is a procedure that brings together the understanding of a system nature and behaviour with a suitable mathematical treatment. The obtained model should describe the system at the most convenient level of precision, which depends on the purpose for which the model is designed, and suppress unnecessary details when appropriate. Even if it is not possible to define the best tradeoff between accuracy and complexity in universal sense, appropriate mathematical tools, can help in taking the right choices. This latter is known as model validation.

A deep and extensive knowledge of the system allows setting directly the values of the model parameters, procedure known as **white-box** modelling. However, for the majority of real systems, white-box modelling is too expensive, if not virtually impossible. A different approach, known as **black-box** modelling uses general function approximators to capture the system dynamics and confirm the system behaviour from experimental evidence. Yet this procedure leads to models which structure is hardly related to the structure of the real system, and which parameters often have no physical meaning. Black-box models can be used to analyse the system behaviour in numerical simulation, and often cannot be scaled up or down.

A mixed approach between white-box and black-box modelling, called **grey-box** modelling, allows using prior knowledge to model known aspects of the system, while unknown or less-known parts of the system are identified using empirical data. This chapter introduces the grey-box model identification and its application to certain classes of systems. Recalling the taxonomy for load flexibility presented in section 3.1, systems that are characterised by a thermal process (climate control and refrigerators) are buckets, EVs are batteries, and white good appliances (e.g. dishwasher, tumble dryer) are bakeries. All non-controllable and/or non directly-observable loads represent a disturbance to the system.

Grey-box modelling In general, a grey-box model can consist of sets of coupled first-order linear Stochastic Differential Equations (SDEs) with the following structure:

$$dx_t = (A(\theta)\, x_t + B(\theta)\, u_t)\, dt + \sigma(\theta)\, d\omega_t \ ,$$

where $A(\cdot)$, $B(\cdot)$, and $\sigma(\cdot)$ are linear functions of θ, which is the set of parameters to be identified, and ω is a standard Wiener process. x_t and u_t are the system state and the control input, which are both function of time, while the system output at the sampling time k is given by:

$$y_k = C(\theta)\, x_k + D(\theta)\, u_k + \eta_k \ .$$

If experimental data is available, parameters identification strives at finding the optimal θ so that the model outputs match the observed data. In this research, the method used for parameters identification is the Maximum Likelihood Estimation (MLE), which is introduced in Appendix B.1.

Thermostatically-controlled Loads (TCLs)

This section presents the modelling of TCLs (bucket model in the BBB taxonomy), using Equivalent Thermal Parameters (ETP) models [Zha+12]:

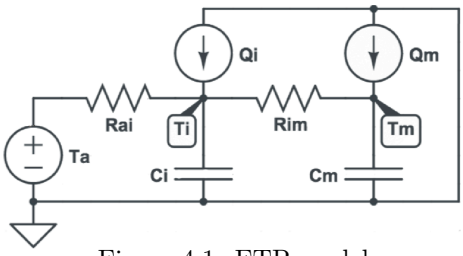


Figure 4.1: ETP model.

$$\begin{aligned} \frac{dT_i}{dt} &= \frac{1}{C_i}\left[\frac{1}{R_{ai}}\left(T_a - T_i\right) + \frac{1}{R_{im}}\left(T_m - T_i\right) + Q_i\right] \\ \frac{dT_m}{dt} &= \frac{1}{C_m}\left[\frac{1}{R_{im}}\left(T_i - T_m\right) + Q_m\right] \end{aligned}, \tag{4.1}$$

where T_a is the TCL external temperature, T_i is the TCL internal temperature (e.g. air temperature for refrigerators or building, and water temperature for water heaters), T_m is the TCL mass temperature (e.g. envelope temperature), and Q_i and Q_m are heat gains of various nature. In the following we present the application of ETP to grey-box modelling of household refrigerators and electric space heating systems.

Household refrigerators

System modelling and parameters identification of a domestic refrigerator is presented in [Cos+13b] (Paper A), which constitutes the contribution of this research in appliances grey-box modelling. The motivation to this study is to provide a simple, ready-to-use and validated lumped parameters model of household refrigerator. The approach consists of forward model selection and validation based on experimental data and statistical testing.

The experimental setup consists of: one household refrigerator of 60 litres capacity with freezer bay and single compressor, one power meter DEIF-MIC2, one ADAM-6024 ADC card, four calibrated temperature sensors TI-LM35, one remotely controlled power outlet. Every second we sample the refrigerator internal temperature in two points, the ambient temperature in two points, and the refrigerator active power consumption. Given the stratification of temperatures in the refrigeration chamber, two sensors are used in order to provide the average internal temperature, analogously it is determined the external ambient temperature. Figure 4.2 shows a schematic representation of a common refrigeration system for household applications, which simplified model is expressed in Eq. 4.2.

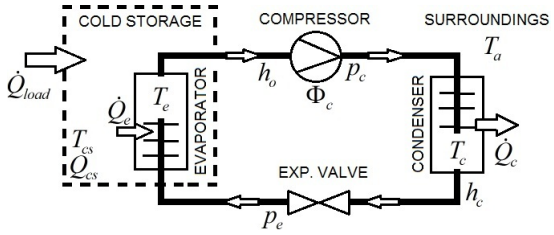


Figure 4.2: Single stage vapour-compression refrigeration system.

$$\frac{dQ_{cs}(t)}{dt} = \dot{Q}_{load}(t) - \dot{Q}_e(t) \ , \tag{4.2}$$

where:

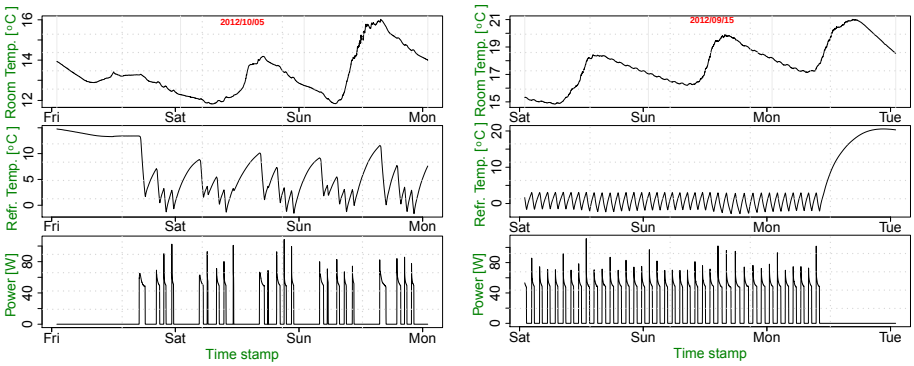
$$\begin{aligned} dQ_{cs} &= m_{cs} c_{cs} dT_{cs} \\ \dot{Q}_{load} &= UA_{cs} \left(T_a - T_{cs}\right) \\ \dot{Q}_e &= \dot{m}_r \left[h_o \left(p_e\right) - h_c \left(p_c\right)\right] \approx COP \cdot \Phi_c \\ \dot{m}_r &= N_c \alpha \rho_r \left(p_e\right) \end{aligned} \tag{4.3}$$

In Eq. 4.3 m_{cs} is the cold storage mass, and c_{cs} is its specific heat capacity. h_o and h_c are the evaporation and condensation enthalpies at the evaporation and condensation pressures, respectively p_e and p_c. UA_{cs} is the overall transmittance coefficient from the refrigeration chamber to the ambient, and $\dot{m}_r$ is the refrigerant mass flow rate. COP is the overall coefficient of performance, here defined as the ratio between $\dot{Q}_e$, the thermal power extracted at evaporator side, and Φ_c, the refrigerator electrical consumption.

Three different models of increasing complexity are designed, identified and validated under the hypotheses of: homogeneous materials, linear cooling cycle with constant COP, and neglecting the freezer compartment. System identification is performed using CTSM [Dan13], which is based on maximum likelihood estimation, and forward model selection and validation is based on experimental data (Figure 4.3) and statistical testing.

Model T_i. Here the refrigeration chamber is represented with a thermal mass, C_i, while the envelope (insulation) is modelled as a pure thermal resistance, R_{ia}. The compressor has a direct refrigeration effect, so that it is modelled as a current generator. This model is a single state stochastic state space model:

$$\begin{aligned} dT_i &= \left[\tfrac{1}{C_i R_{ia}} \left(T_a - T_i\right) - \tfrac{1}{C_i} A_c \Phi_c\right] dt + \sigma_1 dw \\ y_{t_k} &= T_{i,t_k} + e_{t_k}, \quad e_{t_k} \sim N(0, \sigma_e^2) \end{aligned} \ , \tag{4.4}$$



((a)) Identification data set. ((b)) Validation data set.

Figure 4.3: Refrigerator operation: ambient temperature, internal temperature and electrical power consumption.

where A_c is the cycle COP and w is a standard Wiener process independent from the residual e_{t_k}. T_a is the ambient temperature, T_i is the refrigeration chamber temperature and Φ_c is the compressor active power consumption.

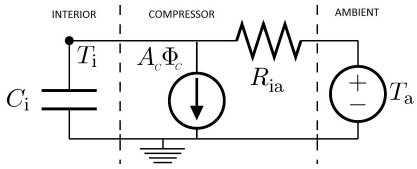


Figure 4.4: Refrigerator preliminary model (electrical equivalent): T_i.

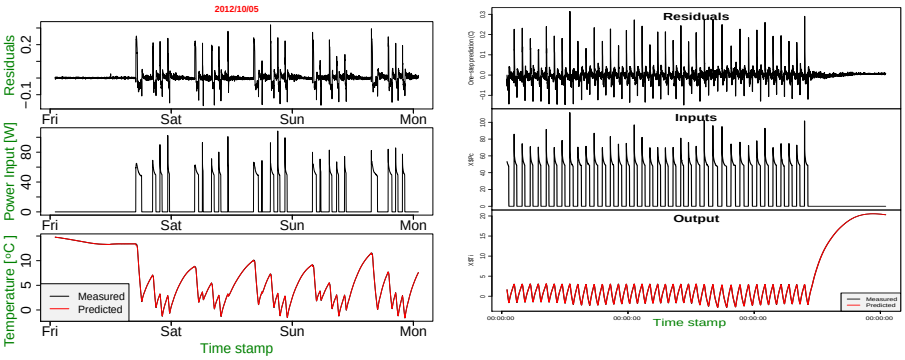
Performing system identification, the parameters in Table 4.5 are obtained. One can observe that, for some of them the estimated variance has the same order of magnitude of the estimated value, which tells us that the uncertainty of this model is very high. Note that a prior estimation of the parameters has to be provided, in order to avoid the estimate to converge to local minima. An overall estimate of parameters is carried in [Cos+13b], based on engineering considerations and calculations, to initialise the MLE estimation in CTSM.

Model T_i is simulated and the one-step prediction error is computed on both identification and validation datasets, as depicted in Fig. 4.6. The one-step prediction is obtained by initialising the model with the experimental data and predicting the internal temperature for the next time step. Then the prediction error is evaluated by comparing the prediction with the actual measured temperature.

If the prediction error resembles to a *white noise*, the model is adequate and cannot further improved. The residual (error) analysis consists in testing whether

Table 4.5: Identified parameters: T_i model

PARAMETER	VALUE	STD. DEV.
$R_{ia}\left[\frac{^oC}{W}\right]$	1.4749	2.5617
$C_i\left[\frac{J}{K}\right]$	$8.9374 \cdot 10^3$	$1.5481 \cdot 10^4$
$T_i(0)\,[^oC]$	14.774	$2.9795 \cdot 10^{-2}$
A_c	0.58092	1.0075
$exp(\sigma_1)$	-5.4552	$1.2511 \cdot 10^{-2}$
$exp(\sigma_e)$	-24.332	75.437
$Loglikelihood$	7995.168	



((a)) Identification data set. ((b)) Validation data set.

Figure 4.6: Model T_i one step prediction.

the one-step prediction error can be assumed as white noise [Mad08]. Observing the prediction error in Fig. 4.6(a) it is possible to depict non-stationarities and identify outliers due to erroneous measurements. The autocorrelation of residuals tells us how each error is correlated to the other errors. Figure 4.7 shows the residuals analysis with respect the identification data set. The first graph on the left presents the auto correlation function (ACF) of residuals, the graph in the middle the periodogram, and the graph on the right the cumulated periodogram. The bands in the ACF and cumulated periodogram represent the 2σ intervals, which represent 95% confidence interval.

High correlation of residuals at low values of lags indicates that the model is not correct at low frequencies, and slow dynamics are not well caught. Moreover, the confidence bands in the cumulated periodogram are largely violated, hence model T_i is discarded.

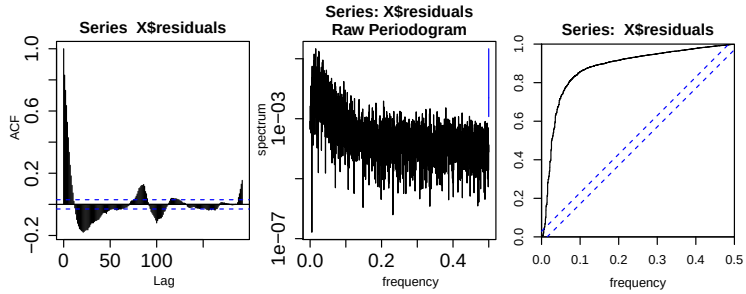


Figure 4.7: Model residuals analysis: T_i, identification data set.

Model T_iT_{evap}. This model extends the previous one by accounting for the heat transfer between the refrigeration chamber and the evaporator with an additional state for the evaporator temperature, T_e:

$$\begin{aligned} dT_i &= \left[\frac{1}{C_i R_{ia}}\left(T_a - T_i\right) + \frac{1}{C_i R_{ei}}\left(T_e - T_i\right)\right] dt + \sigma_1 dw_1 \\ dT_e &= \left[\frac{1}{C_{evap} R_{ei}}\left(T_i - T_e\right) - \frac{1}{C_{evap}} A_c \Phi_c\right] dt + \sigma_2 dw_2 \quad , \\ y_{t_k} &= T_{i,t_k} + e_{t_k}, \quad e_{t_k} \sim N(0, \sigma_e^2) \end{aligned} \tag{4.5}$$

where w_1, w_2 and e_{t_k} are independent stochastic processes. Figure 4.8 shows the ETP circuit associated to model 4.5:

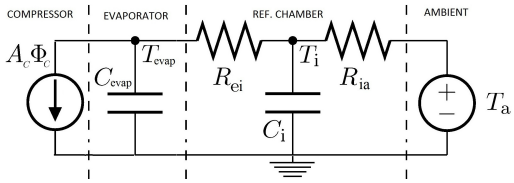


Figure 4.8: Refrigerator model (electrical equivalent): T_iT_{evap}.

System identification leads to the parameters in Table 4.9, while the residual analysis in Fig. 4.10 shows a clear improvement of model T_iT_{evap} compared to T_i and the cumulative periodogram is almost inside the confidence bands. The one-step prediction is no longer presented as it is not possible to appreciate relevant differences with respect to Fig. 4.6. A further extension of model T_iT_{evap} is presented in the next subsection, where the parameters of the refrigerator envelope are accounted.

Table 4.9: Identified parameters: T_iT_{evap} model.

PARAMETER	VALUE	STD. DEV.
$R_{ia}\left[\frac{^oC}{W}\right]$	$9.0188 \cdot 10^{-1}$	$3.5460 \cdot 10^{-2}$
$R_{ei}\left[\frac{^oC}{W}\right]$	$9.0348 \cdot 10^{-1}$	$2.5121 \cdot 10^{-1}$
$C_i\left[\frac{J}{K}\right]$	$1.1600 \cdot 10^{4}$	$1.6529 \cdot 10^{2}$
$C_e\left[\frac{J}{K}\right]$	$3.4342 \cdot 10^{2}$	$9.9157 \cdot 10^{1}$
$T_i(0)\,[^oC]$	14.774	$1.0263 \cdot 10^{-2}$
$T_{evap}(0)\,[^oC]$	16.181	3.6991
A_c	0.8383	$2.6217 \cdot 10^{-2}$
$exp(\sigma_1)$	$-1.7406 \cdot 10^{1}$	$5.6451 \cdot 10^{-2}$
$exp(\sigma_2)$	$-8.9551 \cdot 10^{-1}$	$2.6646 \cdot 10^{-1}$
$exp(\sigma_e)$	$-1.2246 \cdot 10^{1}$	$1.1364 \cdot 10^{-1}$
$Loglikelihood$	12096.4351	

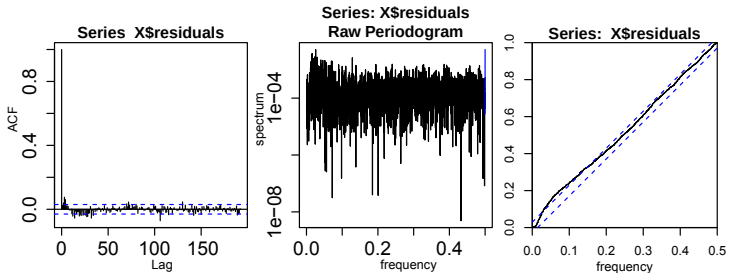


Figure 4.10: Model residuals analysis: T_iT_{evap}, identification data set.

Model $T_iT_{evap}T_e$. Here the T_iT_{evap} model is extended by adding a state to the envelope and separating the envelope thermal resistance in inner resistance, R_{ie}, and outer resistance, R_{ea}:

$$\begin{aligned}
dT_{evap} &= \left[\frac{1}{C_{evap}R_{evi}}\left(T_i - T_{evap}\right) - \frac{1}{C_{evap}}A_c\Phi_c\right]dt + \sigma_1 dw_1 \\
dT_i &= \left[\frac{1}{C_iR_{evi}}\left(T_{evap} - T_i\right) + \frac{1}{C_iR_{ie}}\left(T_e - T_i\right)\right]dt + \sigma_2 dw_2 \\
dT_e &= \left[\frac{1}{C_eR_{ie}}\left(T_i - T_e\right) + \frac{1}{C_eR_{ea}}\left(T_a - T_e\right)\right]dt + \sigma_3 dw_3 \\
y_{t_k} &= T_{i,t_k} + e_{t_k}, \quad e_{t_k} \sim N(0, \sigma_e^2)
\end{aligned} \quad , \tag{4.6}$$

where w_1, w_2, w_3 and e_{t_k} are independent. Follows the electric equivalent model:

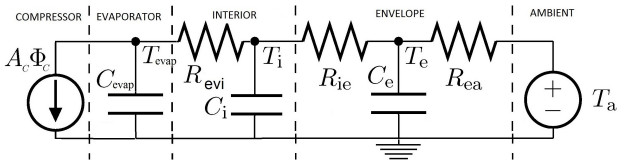


Figure 4.11: Refrigerator model (electrical equivalent): $T_iT_{evap}T_e$.

Table 4.12: Identified parameters: $T_iT_{evap}T_e$ model

PARAMETER	VALUE	STD. DEV.
$R_{ea}\left[\frac{^oC}{W}\right]$	$7.2869 \cdot 10^{-2}$	$1.8571 \cdot 10^{-2}$
$R_{evi}\left[\frac{^oC}{W}\right]$	2.2431	$5.1971 \cdot 10^{-1}$
$R_{ie}\left[\frac{^oC}{W}\right]$	3.7394	1.9380
$C_i\left[\frac{J}{K}\right]$	$4.4245 \cdot 10^{3}$	$2.2810 \cdot 10^{3}$
$C_e\left[\frac{J}{K}\right]$	$1.0755 \cdot 10^{4}$	$2.4514 \cdot 10^{3}$
$C_{evap}\left[\frac{J}{K}\right]$	$1.9177 \cdot 10^{1}$	4.8643
$T_i(0)\,[^oC]$	14.774	$8.6339 \cdot 10^{-3}$
$T_e(0)\,[^oC]$	14.38	6.1042
$T_{evap}(0)\,[^oC]$	18.568	5.5536
A_c	$2.1808 \cdot 10^{-1}$	$1.1258 \cdot 10^{-1}$
$exp(\sigma_1)$	$-1.7661 \cdot 10^{1}$	$1.2498 \cdot 10^{1}$
$exp(\sigma_2)$	$-2.0051 \cdot 10^{1}$	2.4919
$exp(\sigma_3)$	$-6.2477 \cdot 10^{-1}$	$1.0131 \cdot 10^{-1}$
$exp(\sigma_e)$	$-1.1766 \cdot 10^{1}$	$7.9326 \cdot 10^{-2}$
Loglikelihood	12306.517	

Figure 4.13 presents the residuals analysis using the identification data set, where it is depicted that model $T_iT_{evap}T_e$ outperforms in data fitting, and the cumulative periodogram stays almost at all times within the confidence bands. However, performing the residual analysis with respect to the validation data set (Fig.4.14), we see that $T_iT_{evap}T_e$ has poorer fit than T_iT_{evap}, denoting an overfitting. The more complex model has better fit on the training data, but does not generalise well enough with respect to other data sets. Moreover, model T_iT_{evap} has also good residuals properties, a comparable loglikelihood value (12096), and its parameters are closer to the prior estimates (see [Cos+13b]). Hence, model T_iT_{evap} is selected.

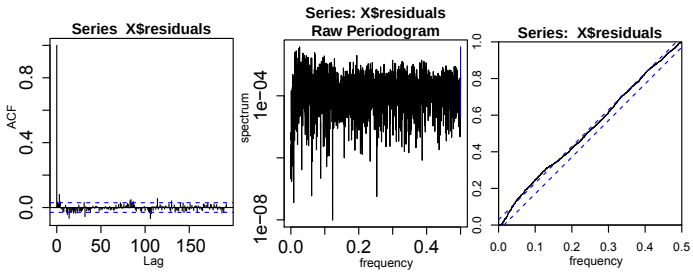
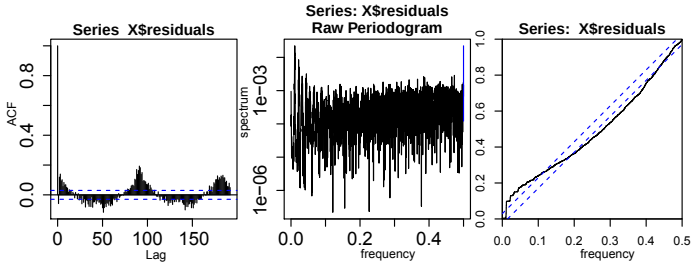
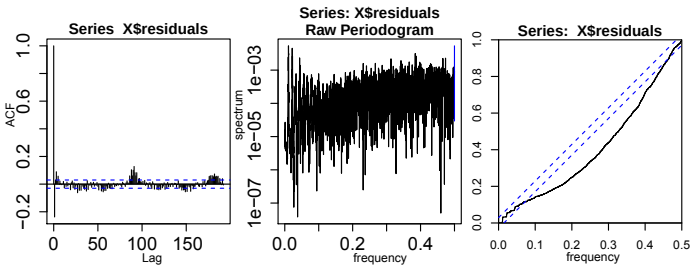


Figure 4.13: Model residuals analysis: $T_i T_{evap} T_e$, identification data set.



((a)) Model $T_i T_{evap}$.



((b)) Model $T_i T_{evap} T_e$.

Figure 4.14: Residual analysis, validation dataset.

This study showed an application of grey-box stochastic modelling for household refrigeration systems. Identified models are simple, reliable and, since they are SDE-based, they can be used for forecasting, linear MPC, and simulation. Thanks to the diffusion terms, model uncertainties are also provided. Being aware of the non-linearities in either the refrigeration process and the thermal exchanges, the proposed model has been extended and revised in [Sos+15], where a non-linear model of the COP of the refrigeration cycle is proposed.

Sossan presents also in [Sos+13] a first order dynamic model of a domestic water heater, which accounts only for the global thermal energy content of the water, and does not describe the thermal stratification of the water inside the tank, as it is instead investigated in [Van+12].

Electric space heating in Power FlexHouse

PowerFlexHouse is a test facility organised as an office building connected to SYSLAB, part of PowerLabDK, a research facility for intelligent, active and distributed power systems at DTU. PowerFlexHouse is electrically heated with 10 space heaters and cooled by 5 air conditioners. Tap water comes from a hot water storage tank and the space is illuminated by 24 fluorescent lamps. There is a small kitchen consisting of a fridge and a coffee machine. Devices in PowerFlexHouse are remotely controlled. The state of the building and appliances is read from various sensors. PowerFlexHouse layout diagram is shown in Fig. 4.15, note that the air conditioning is available only in five rooms out of eight: Room 1, Room 2, Room 3, Room 4 and Room 8 (main hall).

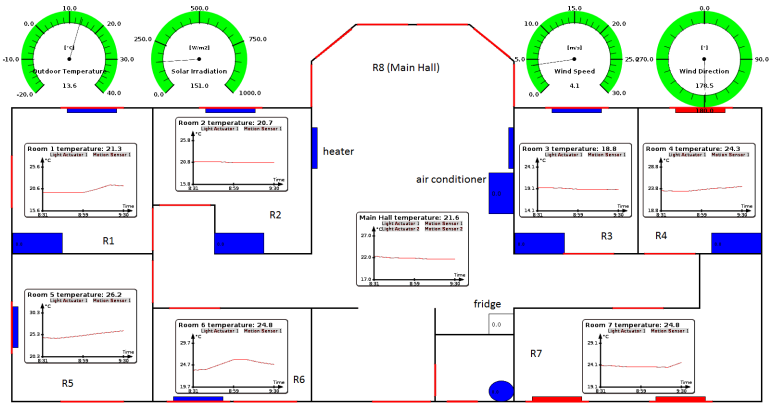


Figure 4.15: PowerFlexHouse layout.

Single-room model. Here the Power FlexHouse is modelled as one large room exchanging heat with the external environment [Tha08]. This crude approximation is convenient to catch the dominant heat dynamics of the building, which allows estimating and predicting its overall heat demand. With such approach, the space heating and cooling is modelled as a single appliance, so that it is not possible to control directly the temperature in rooms individually but control the building average temperature and consumption. Three states are used to describe the building heat dynamics: the first state is the indoor air temperature, T_i, the second state is the temperature of the building envelope, T_{om}, and the third state is the temperature

of the inner walls layers and floor, T_{im}. Assuming that the energy flow due to solar radiation is not absorbed by the outer walls layer, the energy is distributed between the inner walls layer and the indoor air. Therefore the differential equation for the inner wall layer temperature is:

$$C_{im}\frac{dT_{im}}{dt} = hA\left(T_i - T_{im}\right) + \dot{Q}_{im} = \frac{1}{R_i}\left(T_i - T_{im}\right) + A_w p\Phi_s + \hat{R}_{11,1}d\omega_1 \,, \tag{4.7}$$

where C_{im} is the total heat capacity of the inner walls, Φ_s is the solar radiation, R_i is the total thermal resistance between the inner walls and the indoor air, A_w is the total windows surface, and $\hat{R}_{11,1}$ is the diffusion term, which gives information for pinpointing model deficiencies [Baa+97]. The noise therm, ω_1 is a Markov process representing the uncertainty in the model.

The house envelope exchanges heat by convection to both the inside and the outside of the building:

$$C_{om}\frac{dT_{om}}{dt} = \frac{1}{R_{im}}\left(T_i - T_{om}\right) + \frac{1}{R_{am}}\left(T_a - T_{om}\right) + \hat{R}_{22,1}d\omega_2 \,, \tag{4.8}$$

where T_{om} is the outer walls layer temperature, C_{om} is the thermal capacity of the outer walls, R_{im} and R_{am} are the thermal resistances from outer wall layer towards the indoor and outdoor, and $\hat{R}_{22,1}$ is the diffusion term. Finally, the indoor air also receives thermal energy from solar radiation and the electrical space heaters, accounted in the terms Φ_s and Φ_h, so that the equation for the heat balance of the indoor air is:

$$C_i\frac{dT_i}{dt} = \frac{1}{R_a}\left(T_a - T_i\right) + \frac{1}{R_i}\left(T_{im} - T_i\right) + \frac{1}{R_{im}}\left(T_{om} - T_i\right) + A_w\left(1-p\right)\Phi_s + \Phi_h + \hat{R}_{33,1}d\omega_3 \,, \tag{4.9}$$

where C_i is the total heat capacity of the indoor air, R_a is the total resistance against heat flow to the outside, through windows and due to ventilation, and is the energy input from the electrical heaters. Also this equation presents the diffusion term $\hat{R}_{33,1}$. Finally, the measured temperature, T_r, is given by:

$$T_r = T_i + \hat{R}_2 d\omega_4 \,. \tag{4.10}$$

The model parameters are estimated using CTSM [Juh+13], and their values are found in [Tha08].

Multi-room model. Bonby and Parvizi in [BP12] describe the Power FlexHouse individual rooms' heat dynamics, including the influence of the heaters, the solar radiation and the outside temperature. The cooling effect due to wind is modelled as unknown noise, and each room is represented as an element exchanging heat not only with the outside environment but also with the surrounding rooms. This approach creates cross-correlation terms in the model, which describe the behaviour of the room heat heat gain from the adjacent rooms. The known disturbances are solar irradiance

and outside temperature, whereas the thermal power generated by the heaters is the controllable input. Following the grey-box physical modelling, the air temperature in room j is $T_{r,j}$:

$$\frac{dT_{r,j}}{dt} = \frac{1}{C_{r,j}} \left[\sum_{k \in S_j} \frac{1}{R_{r,kj}} \left(T_{r,k} - T_{r,j} \right) + A_{w,j} \Phi_s + \Phi_{h,j} + \sigma_{r,j} \frac{dW}{dt} \right] , \tag{4.11}$$

where $C_{r,j}$ is the air thermal capacity of room j, $R_{r,kj}$ is the thermal resistance between room j and room k, A_w is the window area, Φ_s is the solar radiance, $\Phi_{h,j}$ is the heating power in room j, $\sigma_{r,j}$ is the diffusion therm, and S_j is the set of rooms adjacent to room j.

The innovations brought by the multi-room model are: cross-correlation terms between rooms, which provide a more precise model of the building; quasi-diagonal system, which allows to use the same identification approach as in [BM11]. The parameters of the continuous time model in Eq. 4.11 are identified via CTSM and presented in [BP12]. An analogous approach is followed by Oldewurtel et al. in [Old+10], to which the interested reader can refer for a stochastic MPC formulation of building climate control.

Batteries and Electric Vehicles

Electric Vehicles are intended as energy buffers, which state of charge must be maintained within a given range. Thus, the correspondent model in the BBB taxonomy is the Battery. In the context of this research, the EVs are modelled as integrators with an efficiency parameter, and the energy consumption related to driving patterns is considered and simulated as disturbance to the MPC. A detailed model of electric storage units is presented in [Mar+12], from which it emerges that, in order to extend the lifetime, batteries should be operated between 20% and 90% of the state of charge (SOC). Moreover, when operating in this band, the batteries can be charged and discharged at controlled current, therefore can be reasonably modelled as integrators of rated efficiency.

Support batteries differ from the EVs by capacity and rated power, therefore a generic state space model for this type of units is presented below:

$$\begin{array}{l} \dot{x}\left(t\right) = Ax + Bu \\ SOC\left(t\right) = Cx \end{array} , \tag{4.12}$$

where u is the power flow at the inverter in kW, x is the battery state of charge in kWh and SOC is the battery state of charge normalised between 0 and 1. Note that the state and the input vector have both dimension 2. This is in order to take into account different efficiency for charging and discharging, therefore the system matrices are:

$$A = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} -\eta_c \\ -\eta_d^{-1} \end{bmatrix}, \quad C = \begin{bmatrix} rc^{-1} & rc^{-1} \end{bmatrix}, \quad u = \begin{bmatrix} u_{charge} \\ u_{discharge} \end{bmatrix},$$

where η_c and η_d are the inverter efficiencies for charging and discharging (usually they have the same value), and rc is the battery rated capacity. Note that u_{charge} is negative and $u_{discharge}$ is positive, therefore the negative signs in the B matrix, and that model 4.12 is in continuous time.

The approach used for EV modelling is white-box, since the necessary parameters (e.g. rated capacity, power and efficiency of inverters) are known from the devices datasheets.

Other appliances

Here we present the modelling approach for such appliances as washing machines or tumble dryers. This type of devices have a consumption profile that depends almost entirely on the operation mode, and therefore they correspond to the *Bakery* in the BBB taxonomy, which model is:

$$\begin{aligned} x_{k+1} &= a_d x_k + T_s u_k \\ x_N &= E_N \\ u_k &= P_k \quad , \quad P_k \in \mathcal{P} \end{aligned} \tag{4.13}$$

We can assume that some of them have flexible starting time, such as washing machines or dishwashers, while others don't (e.g. microwave ovens). In both cases the modelling approach used here can be either white-box or grey-box. White-box is used if the consumption profile is known a-priori, otherwise a grey-box identification is performed by measuring the consumption profile in the different operation modes. In any case the appliance operation is determined by the interaction with users.

Conclusion

This section presented some linear models of common household units, which flexibility can be used to provide ancillary services to the power system. The next section and Appendix B.2 introduce the reader to the subject of Model Predictive Control (MPC), with an overall overview of linear and quadratic MPC, centralised and distributed MPC, and serves as preamble to Sec. 4.3, which presents the original contribution of this research in DLC for DR applications.

4.2 Model Predictive Control for Demand Side Management

A model predictive controller (MPC) uses a model of the system under control to predict the process behaviour over a specified prediction horizon. In this way, it can optimise the sequence of control inputs with respect to a specified performance criterion. Consider a time-invariant linear model, or a linear approximation for a determined operating point, in discrete-time state space form:

$$S : \begin{cases} x_{k+1} = A_d x_k + B_d u_k + E_d d_k \\ y_k = C_d x_k + D_d u_k \end{cases} , \tag{4.14}$$

where k is the current time step, x is the system state space, y is the system output, A_d, B_d, E_d, C_d, D_d are the state space matrices, u is the control input, and d is the process disturbance. The cost function of an MPC controller can be designed to pursue different objectives, e.g. follow a reference signal, limit the control effort, and the cost of control action:

$$V(k) = \sum_{j=k}^{k+N} \left\| y_{k+j|k} - r_{k+j|k} \right\|^2_{Q(j)} + \sum_{j=0}^{N-1} \left\| \Delta u_{k+j|k} \right\|^2_{R(j)} + \sum_{j=0}^{N-1} \rho_{k+j} u_{k+j|k} \ . \quad (4.15)$$

Following the procedure presented in Appendix B.2, it is possible to formulate the MPC problem as least-squares optimisation:

$$\begin{aligned} \min_{x} \phi : & \ \tfrac{1}{2} x^T H x + f^T x \\ s.t. : & \quad Ax \leq b \end{aligned} \ , \quad (4.16)$$

which can be efficiently solved using commercially available and open-source optimisation libraries for Quadratic Programming (e.g. CPLEX, MOSEK, OjAlgo, Quadprog (Matlab), and many others).

Flexibility aggregation is crucial as single units often cannot provide valuable ancillary services to the System Operator due to their small size and market and regulation barriers. Therefore, aggregation is a way to allow residential units to participate in DR. An aggregator can control the overall power consumption of a portfolio of residential DSRs in order to follows a desired pattern, respects a set of constraints, or even react to price signals. On first instance, one can think of centralised MPC (CMPC) as a suitable control option for a cluster of DSRs, in this case the CMPC-based aggregator must embed the models of the physical resources in the control cluster.

Centralised MPC. Consider the case of geographical aggregation, the DSRs are branched to the same feeder and participate in the same control cluster. The aggregated power at the point of common coupling (PCC) is controlled to provide ancillary services to the SO (Fig. 4.16).

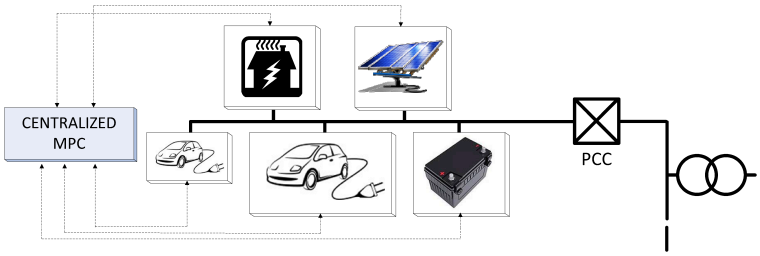


Figure 4.16: Centralised MPC architecture for Demand Response.

Recall the DSRs models in their discrete-time state space representation, e.g. Eq. 4.14, the CMPC is easily designed by assembling together the DSRs models as shown below:

$$x_c = \begin{bmatrix} x_1 \\ x_j \end{bmatrix}, u_c = \begin{bmatrix} u_1 \\ u_j \end{bmatrix}, d_c = \begin{bmatrix} d_1 \\ d_j \end{bmatrix}, y_c = \begin{bmatrix} y_1 \\ y_j \end{bmatrix}$$

$$A_c = \begin{bmatrix} A_1 & 0 \\ 0 & A_j \end{bmatrix}, \quad B_c = \begin{bmatrix} B_1 & 0 \\ 0 & B_j \end{bmatrix}, \quad C_c = \begin{bmatrix} C_1 & 0 \\ 0 & C_j \end{bmatrix} \tag{4.17}$$

$$D_c = \begin{bmatrix} D_1 & 0 \\ 0 & D_j \end{bmatrix}, \quad E_c = \begin{bmatrix} E_1 & 0 \\ 0 & E_j \end{bmatrix}$$

where the index $j \in \mathcal{S}$ indicates a specific subsystem. The CMPC problem is formalised in Eq. 4.18 for the PowerMax service and, since the units are connected to the same point of common coupling (PCC), their aggregated power consumption is explicitly taken into account in constraint 4.18f:

$$\min_{u_{c,t};\vartheta_{c,t};\gamma_{c,t}} V : \frac{1}{2}\left[\sum_{t=1}^{N} \|y_{c,t} - r_{c,t}\|_Q^2 + \sum_{t=0}^{N-1} \|\Delta u_{c,t}\|_R^2 + \sum_{t=0}^{N-1} \rho_t u_{c,t} + \sum_{t=1}^{N} \eta_t \vartheta_{c,t} + \sum_{t=1}^{N} \psi_t \gamma_{c,t}\right] \tag{4.18a}$$

$$s.t: \; x_{c,t+1} = A_c x_{c,t} + B_c u_{c,t} + E_c d_{c,t} \tag{4.18b}$$

$$y_{c,t} = C_c x_{c,t} + D_c u_{c,t} \tag{4.18c}$$

$$u_{c\,\min,t} \leq u_{c,t} \leq u_{c\,\max,t} \tag{4.18d}$$

$$y_{c\,\min,t} - \gamma_{c,t} \leq y_{c,t} \leq y_{c\,\max,t} + \gamma_{c,t} \tag{4.18e}$$

$$PCC_{\min,t} - \vartheta_{c,t} \leq u_{c,t} \leq PCC_{\max,t} + \vartheta_{c,t} \tag{4.18f}$$

$$\vartheta_t \geq 0 \tag{4.18g}$$

$$\gamma_t \geq 0 \tag{4.18h}$$

$$\tag{4.18i}$$

In Eq. 4.18 ρ_t is the energy price, while $\vartheta_{c,t}$ and $\gamma_{c,t}$ are the soft constraints for the power flow at the PCC and the systems outputs. Being this a centralised optimisation, the solution of Problem 4.18 is considered as benchmark for the global optimum. Note that Problem 4.18 can be adapted to the direct power control service by simply setting PCC_{min} and PCC_{max} to the same value, which is the cluster desired power.

Two problems arise from this approach: the necessity of a **central repository** of DSRs models, and the explosion of the **computation time** with the problem size. The necessity of resource models makes CMPC problematic to implement and deploy for cluster of DSRs of heterogeneous type, produced by different manufacturers, and placed in different parts of the grid. Furthermore, the computational effort for

solving MPC problems generally grows at a super-linear rate with the number of state variables involved. The simulation study presented below shows the growth of the CMPC computation time with respect to increasing number of units and different discretisation steps.

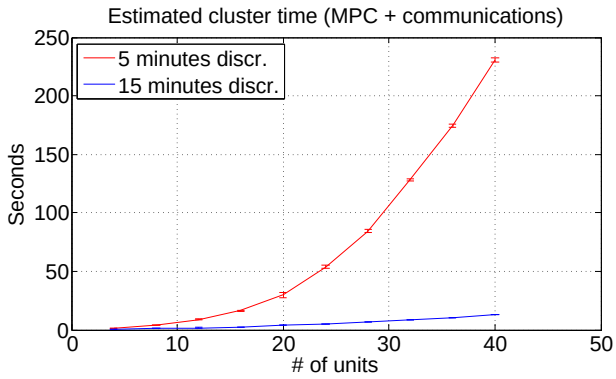


Figure 4.17: Estimated cluster time of centralised MPC for different cluster sizes.

This study does not account for the communication overheads, which depend on the specific network technology. Figure 4.17 reports the estimated cluster time with respect to different discretisation steps and cluster sizes. The prediction horizon is 24 hours. The *cluster time* is defined as the time needed to compute the control input (solve the optimisation problem) plus the time to retrieve units status and dispatch the control signal (communication delay, for which an educated guess is 100ms):

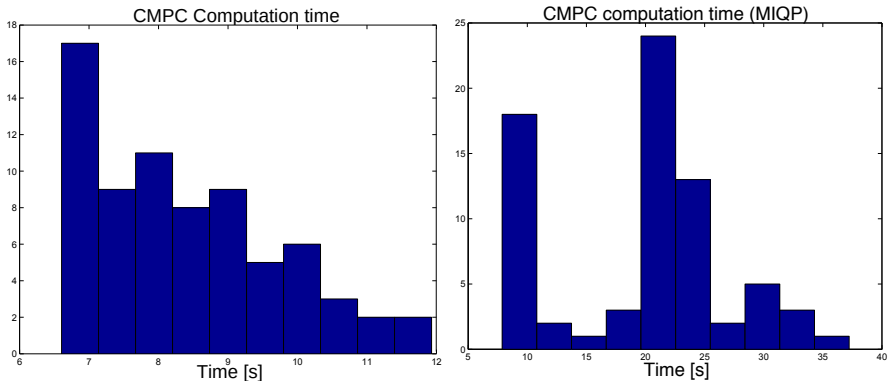
$$ECT = T_{MPC} + N_{units} * 100ms \ ,$$

where it is assumed that the communication between the units and the CMPC occurs in sequence. Table 4.18 compares the cluster time for different problem sizes and discretisation steps.

An increase of the sampling time to 15 minutes makes the problem much faster to solve. However, 15 minutes is a rather large interval to schedule EVs charging if the aggregator has to cope with unforeseen vehicle departures and arrivals. A resource becoming unexpectedly unavailable causes an imbalance in the consumption/production schedule, which cannot be resolved before the next MPC computation. Therefore the interest in having 5 minutes discretisation, or even lower. Note that in this study all the control variables are continuous, and in case some resources have binary control variables the computation time is dramatically affected, as shown in Fig. 4.19.

Table 4.18: CMPC: problem size and cluster time (CPU: Intel Core i5, 2.6 Ghz; RAM: 8 GB, 1.6 Ghz DDR3; OS X 10.9.5; Matlab R2011b with Mosek 7).

Pred. horizon: 24h	**CLUSTER SIZE (# OF UNITS)**			
	4	12	28	40
DISCRETISATION STEP	**PROBLEM SIZE (# VAR.)**			
5 minutes	2208	6072	13800	19596
15 minutes	736	2024	4600	6532
DISCRETISATION STEP	**ESTIMATED CLUSTER TIME [sec.]**			
5 minutes	1.36	8.74	84.44	230.75
15 minutes	0.56	1.80	6.46	12.71



((a)) ((b))

Figure 4.19: Computation time for CMPC: continuous quadratic programming (QP) (a) and mixed-integer quadratic programming (MIQP) (b).

Figure 4.19 compares the CMPC computation time, measured in several simulations, for a QP formulation and MIQP formulation in a cluster of 15 units. The discretisation step is 15 minutes, the prediction horizon is 24h and the units running on integer control variables are 6. The setup is a shared workstation equipped with Intel Xeon 3Ghz (8 cores), 16 GB ram, Windows Server 2008 64bit, Matlab R2012a with YALMIP toolbox and MOSEK optimisation libraries for QP and MIQP.

From this preliminary assessment, we conclude that managing multiple DERs with a single, feeder-wide MPC controller not only requires a large amount of DER state information to be shared with the central controller, but also the computational effort scales badly with the number of DERs.

Distributed MPC. An established approach for solving the curse of dimensionality relies on the decomposition of the MPC problem into smaller subproblems which can

be solved independently and locally. Convergence towards the overall goal is then achieved through a coordination mechanism, i.e. by communication between the individual solvers (Fig. 4.20).

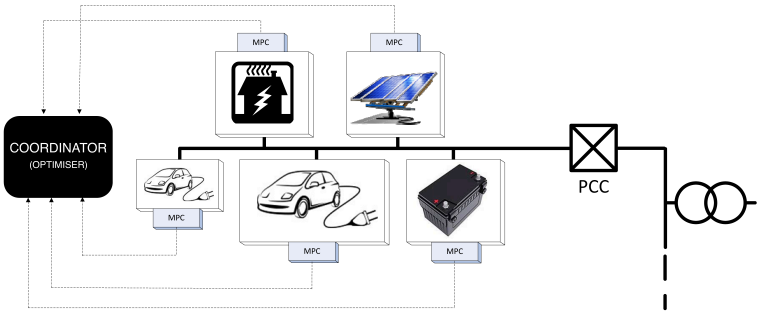


Figure 4.20: Distributed MPC architecture with coordination.

DMPC is based on information exchange between units' controllers, typically consisting of the future predicted control inputs or state variables computed locally at resource level. Scattolini presents in [Sca09] a survey on architectures for distributed and hierarchical MPCs, the state of the art technology, and a classification of DMPC schemes based on the information exchange between controllers, the type of connection between controllers, and the controllers objective function. According to this classification, DMPC schemes can be based on fully connected algorithms, in which every clonal controller communicates with all the other controllers in the cluster, or on partially connected algorithms, where each controller communicates only with neighbour controllers. Controllers can exchange information several times within each time step, in case of iterative algorithms, or only once per time step in case of non-iterative algorithms. Finally, controllers optimise a local performance index in independent algorithms, or they can optimise a global cost function in cooperative algorithms.

Different methods for distributed and parallel optimisation have been explored and proposed in order to overcome large scale tasks. We are particularly interested in those methods that use convex optimisation and, with this respect, our main reference is [Boy+11]. Here the reader can find the method of Dual Decomposition (DD) and Alternating Direction Method of Multipliers (ADMM), which are both based on the assumption that the objective function is separable with respect to a partition of the variable into sub vectors [Boy+11]:

$$f(x) = \sum_{i=1}^{N} f_i(x_i) ,$$

where $x = (x_1 \dots x_N)$, and x_i is a sub vector of x. In Appendix B.3 and B.4 we briefly introduce the DD and ADMM methods, with the aim of providing the reader

with a brief presentation, while for an exhaustive review of both methods and a rich list of references, the interested reader can refer to [Boy+11].

While DD and ADMM offer the means to solve the MPC problems in parallel at units level, they present the following specific characteristics:

> DD: needs continuous variables, convex objective functions (primal and dual problem), and accurate choice of the update parameter for the dual variable;
>
> ADMM: needs central coordination with optimisation, and presents slow convergence (often, at least few tens of iterations are needed).

Moreover, both methods assume, together with the decomposability of the master problem, a specific formulation of the optimisation problem at units level. These characteristics of DD and ADMM inspired the development of the SDMPC algorithm, presented in the next section.

4.3 Sequential Distributed Model Predictive Control

This section presents an innovative coordination scheme for distributed MPC, called Sequential Distributed Model Predictive Control (SDMPC), that is based on information sharing through a blackboard. According to the classification presented in [Sca09], SDMPC is an iterative, cooperative, and partially-connected algorithm. As it is further explained in Paper B, the controllers coordinate in sequential mode, from which the name *sequential*. Note that the therm SDMPC is introduced here and not in the paper, as an outcome of further discussions between the paper coauthors. A second paper integrates this chapter, Paper C, where the performance of the SDMPC is assessed with respect to a centralised MPC control algorithm. This latter study was conducted by D.E. Bondy and supported by G. T. Costanzo, in which a novel performance index for aggregators is introduced and the two MPC architectures are evaluated for the PowerMax service.

SDMPC is designed to solve a general consensus problem in a resource sharing setting, with centralised information sharing among agents [Cos+13a]. Consider again the case introduced in Sec. 4.2, where DSRs are aggregated basing on their geographical location, and the service provided is PowerMax. Instead of looking at the problem in its entire, one can identify two subproblems: 1) a global problem, called *master problem*, which consists in assigning a share of a determined resource (power flow at the PCC) to all clients (DSRs) avoiding overusing the resource, and 2) a *local problem* at client level, which consists in requesting a convenient amount of the shared resource (power consumption) to provide a service to the end user (Fig. 4.21).

This is the setting: each DSR has a local controller, hereafter called *agent*, that features MPC and shares information with a central repository, called *blackboard*. The blackboard does not perform any optimisation task but stores the information the agents put in it, so it can be thought of a marketplace. All the DSRs connected to the same PCC coordinate their operation through the same blackboard, thus they

belong to the same *control cluster*. Each agent in the cluster accesses the blackboard, reads the *global power schedule*, which is the cumulative schedule obtained from the other agents' power schedules, computes its local power schedule, and publishes it in the blackboard. We refer to *power schedule* as the time plan of power withdraws (positive) or injections (negative).

Together with the local power schedule the agents publish in the blackboard their *local consensus* flag, which is a boolean value that is *true* if the agent finds that the global power schedule plus its own one does not exceed the power limits at the PCC at any time, *false* otherwise. Each agent determines its local consensus flag, while the blackboard determines the *global consensus* flag, which is *true* if all the local consensuses are true or a maximum number of negotiations has been reached, *false* otherwise. This algorithm can also be used for the direct power control service.

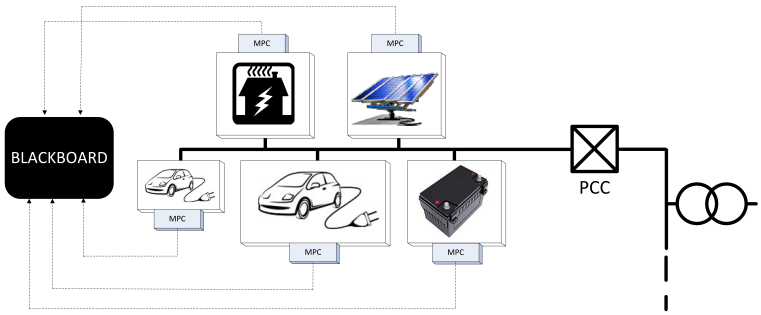


Figure 4.21: SMPC architecture with the central data repository, the blackboard.

The agents keep calling the blackboard until they find *true* for the global consensus, and publish their ideal (unconstrained) schedule to maximise the local comfort until the shared resource is not overexploited. If the shared resource is overused, between the agents starts a **collaborative game** with the following elements [Ras06]:

1. Players
    a) the DSR controllers (agents);
    b) the Blackboard (arbitrator);
2. Information
    a) cluster global power schedule and global consensus flag;
    b) DERs local power schedules and local consensus flags;
3. Actions
    a) Agents
        - retrieve cluster global power schedule, global consensus flag, and resource usage threshold;

- compute the local power schedule and local consensus flag;
- publish the local power schedule and consensus flag in the Blackboard;

b) Arbitrator

- compute the global consensus flag;

4. Payoffs (only DERs)

a) maximise the local comfort and guarantee a minimum level of service;

b) reduce the global resource overuse;

In order to avoid overusing the shared resource, each agent modifies its own request as long as its minimum local comfort level is met. In the worst case, i.e. the available resource is not enough to meet all agents' minimal requirements, each agent publishes its *minimal request*, or *minimal power schedule*, which is the power schedule that is enough to satisfy the DSR minimum comfort requirements. This updated plan contributes reducing the L_1 norm deviation of the global power schedule from the constraints. This last rule makes the SDMPC a *best-effort* algorithm, in the sense that if a feasible solution exists with respect to the constraints on the PCC, it is found. However, there is no guarantee that such solution corresponds to the best global welfare. Conversely, if no feasible solution exists, a sub optimal and best-effort solution is found to reduce the violations of the global power schedule on the PCC power constraints.

Figure 4.22: Agents behaviour with respect to total resource usage.

The global consensus flag, when set to true, determines the agents to close the negotiation round and proceed with dispatching the respective power schedules. A negotiation round is completed when all the agents have posted, or updated, their schedule to the blackboard. If the global consensus is not reached after a certain number of negotiation rounds, the blackboard forces the global consensus to true in order to close the negotiations and avoid deadlocks. Note that this is just one stopping criterion that is compatible with SDMPC, another one could be related to the resource overexploitation. After a certain number of negotiation rounds in which the resource is being overused, if the resource over exploitation does not decrease, the negotiation rounds are stopped.

Note that the structure, objective function and optimisation algorithm of the MPC controller in each agent can be independent from the other agents. Furthermore, each agent is able to: 1) read from the blackboard the global power schedule, the resource usage threshold, and the global consensus flag; 2) compute the forecasted

resource usage and publish the schedule in the blackboard, together with the local consensus; and 3) be *cooperative*, which implies to update the local schedule in order to compensate for other agent's needs.

Such structure makes the SDMPC algorithm more versatile for some applications than ADMM or DD, since each agent is not bounded to a specific optimisation formalism. Moreover, this scheme suits mixed-integer optimisation problems. Furthermore the size and complexity of MPC problems depends only on the resource models and not on the control cluster size.

Follows a study on the convergence and the computational complexity of the SDMPC algorithm based on simulations. Starting with presenting the benefits of the SDMPC algorithm in reducing the peak load, Fig .4.23 shows a comparison of the coordinated DSRs operation with respect to the non-coordinated operation (refer to Paper B for more details on the simulation setting), in a case in which the power constraints are not over stringent.

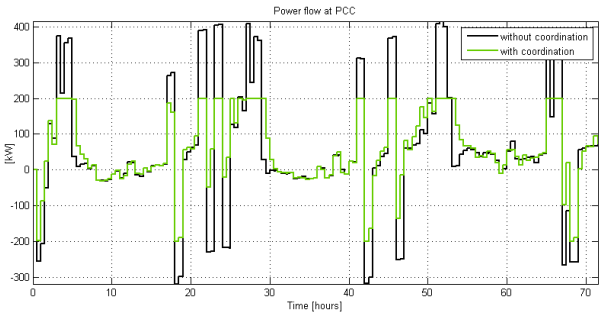


Figure 4.23: Power flow at the point of common coupling: units coordination versus units non-coordination.

Note that the units' power schedules in periods far from the load shedding is very similar to the uncoordinated behaviour, confirming that the SDMPC algorithm leaves units free to operate if the constraints on the total power are not attained. Immediately after a load shedding, there is a settling time for the *coordinated schedule* to match the non-coordinated schedule. Figure 4.24 shows how the agents reach a global consensus in rather few negotiation rounds, if the constraints on the PCC power are not over stringent. The lines of different colour show the maximum deviation of global schedule from the PCC limit. Different levels of the PCC limit correspond to different colours.

One can depict that, if the PCC limit is below a certain value, all the units publish the minimal schedule and the violation such constraint is not further reduced with the negotiation rounds (best effort solution). One can depict that 3 or 4 negotiation rounds are sufficient to settle all the power schedules. Note that the convergence time, thereafter called *cluster time* of the SDMPC algorithm is related to three factors: 1) the solution of the local MPC problems, which are fairly small size, 2) the

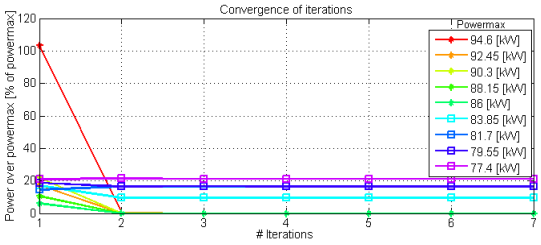


Figure 4.24: Study of algorithm convergence, exceeding power at PCC versus number of negotiation rounds.

amount of units in the cluster, and 3) the communication time and overheads, which are implementation-dependent. Therefore, first we discuss the main implementation details of SDMPC before proceeding with the estimation of the cluster time.

The blackboard is contacted by an agent, which reads the operational plans that the other agents have published and the PCC power limit. Basing on this information, the agent computes an optimal plan, that is published in the blackboard, together with the local consensus flag. Note that this is an iterative algorithm in which the access to the blackboard is sequential, motivated by the fact that every time an agent computes the power plan, it needs the most updated information regarding the other units plans. Coordination between agents, thereafter called clients, is performed in successive negotiation rounds, and the duration of the rounds is a lower bound for the achievable time granularity of the system. Figure 4.25 shows the sequence of information exchanges between a coordinator and a client during a regular bidding transaction; each round is divided into two time slots, a join slot and a negotiation slot.

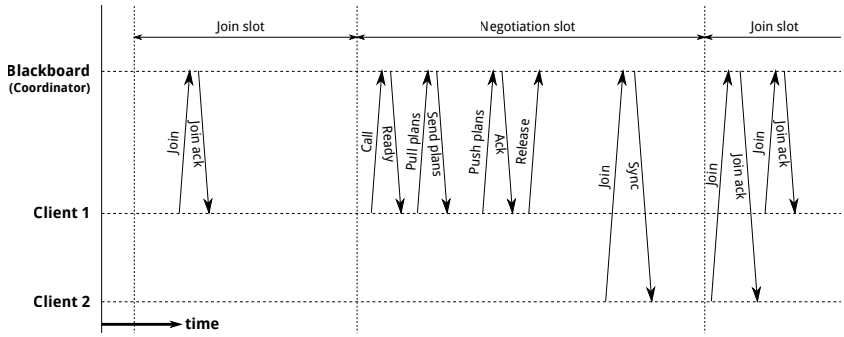


Figure 4.25: Joining and negotiation sequences for two clients. Client 1 is already synchronised to the blackboard slot timing while client 2 attempts to join during a negotiation slot and gets rejected.

The duration of the rounds, as well as the beginning and end of the two slots,

are defined by the blackboard with a Time Division Multiple Access (TDMA) policy. This allows several agents to take part to different negotiation rounds thanks to the time slots division [2]. In order for the coordination scheme to work, clients must be synchronised to the timing of the coordinator, i.e. the blackboard, and clients must have a similar notion of the beginnings and ends of negotiation rounds, and time slots. Some amount of synchronisation error Δt is tolerated by the algorithm as long as, for negotiation rounds of duration t_R, the condition $\Delta t << t_R$ is met. During the join slot, clients can announce their willingness to participate in the subsequent negotiation period, and those who fail to make this announcement are assumed inactive. This ensures that unresponsive controllers do not disturb the negotiations beyond that round.

In Fig 4.25 a client (agent) initially calls the coordinator (blackboard) to ask for a lock, after the confirm of which the client asks for the existing bids. The coordinator sends the bids, so the client can perform MPC-based optimisation of its power schedule based on the PCC constraints and the existing bids. Then, the client submits the result (bid) to the coordinator. If the coordinator acknowledges the validity of the bid, the client releases the lock on the blackboard. If the bid is not valid, the coordinator cancels the transaction, which implies removing the lock as well. A bid is deemed not valid if the data consistency check fails. If another client attempts to call while the blackboard is locked, it receives a sync signal (used to synchronise with the blackboard and interpreted as a rejection) and it waits for a random time before trying again. This feature enhances in the long run the fairness of the algorithm. In fact, although the units that access the blackboard earlier get a higher share of the power resource, if the first unit to get the bid is somehow random, then all the agents in the long run have equal probability to get the first bid.

Notice that, differently from the division of negotiation periods (join slot plus negotiation slot), the mere blackboard access is managed with a Carrier Sense Multiple Access With Collision Detection (CSMA/CD) policy [3]. Although the length of negotiation periods determines the upper bound of the cluster time (total time to settle the power schedules), its expected value is related to the number of participating agents. Furthermore, the random waiting time of the CSMA/CD-based access to the blackboard affects the maximum number of DSRs per cluster. In this context, consider the estimation of the cluster time:

$$ECT = 3 * \left(\sum_{j=1}^{N_{units}} T_{MPC_j} + N_{units} * 100ms \right) , \tag{4.19}$$

which, since the MPC time of the various controllers (T_{MPC}) does not depend on the cluster size, denotes a linear growth with the cluster size, N_{units} (the random waiting

[2] TDMA is used in communication networks as channel access method for shared medium networks. In our case, the agents access to the blackboard is performed with another technique, and only the negotiation periods are divided with a TDMA policy.

[3] In this implementation there are not jam signals being sent by the agents, as the agents do not communicate with each others, but only with the blackboard.

time is not taken into account here).

Table 4.26: SDMPC: problem size and cluster time (CPU: Intel Core i5, 2.6 Ghz; RAM: 8 GB, 1.6 Ghz DDR3; OS X 10.9.5; Matlab R2011b with Mosek 7).

Pred. horizon: 24h	**CLUSTER SIZE (# OF UNITS)**			
	4	12	28	40
DISCRETISATION STEP	**PROBLEM SIZE (# VAR.)-BUILD/BATT**			
5 minutes	1152/864	1152/864	1152/864	1152/864
15 minutes	384/288	384/288	384/288	384/288
DISCRETISATION STEP	**ESTIMATED CLUSTER TIME [sec.]**			
5 minutes	7.55	22.60	51.58	73.86
15 minutes	2.22	5.44	12.70	18.12

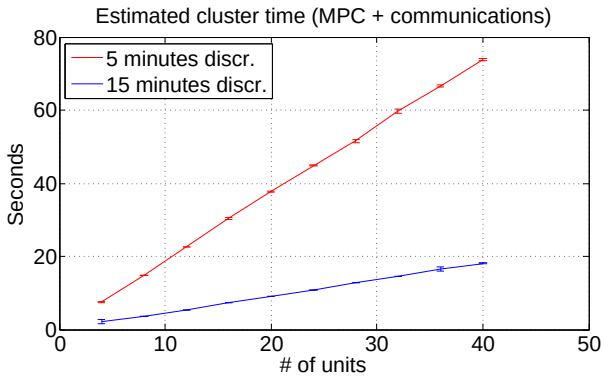


Figure 4.27: Estimated cluster time of centralised SDMPC for different cluster sizes. The waiting time related to CSMA/CD access to the blackboard is not accounted.

In order to determine the upper bound of the cluster size, consider the average transaction time per unit from Table 4.26 (cluster time divided by number of units): a conservative estimation is 1.84 seconds (5 min. discretisation), excluding the waiting time. In [ML83], Meditch proposes an analytical evaluation of the delay characteristics of CSMA/CD policies. It is found that the maximum channel throughput, which in our case represents the maximum cluster size achievable, is reached at 50% of the channel theoretical throughput, with an average waiting time of 10% of the transmission time. In our case, being the negotiation slot 240 seconds long and the unit transaction time 1.84 seconds, the estimated maximum cluster size in this setting is: 65 units with an optimal waiting time uniformly distributed on 10% $\pm$5% of the unit time. Although these represent quite conservative estimates, the acquired timing margin may help in coping with unexpected system delays, such as Java garbage collection procedures at agent and blackboard level.

We refer the reader to Paper B and Paper C for more details on the SDMPC algorithm and a performance assessment with respect to the service delivery to the final user and the System Operator. Regarding the CSMA protocol and networks delays analysis, [ML83] and [BC88] provide a valuable insight on this topic.

4.4 Conclusions

This chapter presents a standard approach DSRs grey-box modelling and two approaches to Direct Load Control based on MPC towards the application of a DSR portfolio management for the provision of ancillary services. This problem shows an intrinsic complexity in both modelling and control, which we aim at simplifying.

Although the CMPC architecture provides the optimal DSRs power scheduling, it has shown to scale badly with the cluster size and discretisation time. With this respect, distributed control represents an attractive approach. Two analysed algorithms in literature, Dual Decomposition and ADMM, guarantee convergence and optimality of the solution under more or less strict hypotheses; however they both are restricted by two facts: the local controller must be designed with a specific mathematical formalism, and the central coordinating agent must solve and optimisation problem.

The proposed solution, called Sequential DMPC, tackles both restricting aspects at the price of sub-optimal/best-effort solution and less scalability with respect to the cluster size with respect to DD or ADMM due to the sequential and exclusive communication between the agents and the blackboard. The advantages brought by SDMPC consist not only in an independent design of controller and a simple coordination system, but also in the capability of handling heterogeneous control variables. Moreover, SDMPC exhibits a collaborative behaviour under resource overuse to provide the highest quality of service to both end users and System Operators. SDMPC features higher robustness with respect to system failures if compared to the CMPC, which has a single point of failure that is the centralised controller, since the units can work in auto mode if the blackboard is down, and it allows an increase of approximately 63% of the maximum cluster size, from 40 units (CMPC) to 65 units (SDMPC) for the considered setting.

CHAPTER 5

Data-driven modelling and predictive control of DSRs

"In God we trust; all others must bring data."

William Edwards Deming [1]

This chapter presents two approaches based on artificial intelligence (AI) to demand side management. The first approach consists of two separate steps: system modelling via decision trees, and optimal control via Dynamic Programming, so it is denoted thereafter as model-assisted Dynamic Programming *(MADP). Although the presented MADP benefits from a simple design, the time required to system modelling is large and the generalisation of the learned model is limited. Therefore a second approach is introduced, which uses online learning and interaction with the system to directly determine the optimal control policy. This approach is based on* Model-assisted Batch Reinforcement Learning*, since it combines in-domain knowledge to boost the convergence and generalisation of the pure Reinforcement Learning-based controllers.*

[1] Dr. William Edwards Deming (14/10/1900 - 20/12/1993). He was an American engineer, statistician, professor, author, lecturer, and management consultant. Japanese economic renaissance after the second world war got large benefit from Deming ideas: better products design, high level of uniform product quality, improvement of product testing in the workplace and in research centres, and greater sales through markets.

5.1 Introduction

Grey-box physical modelling has been widely used for Model Predictive Control of TCLs, it is based on the system thermo-physical properties, and it relies on consistent prior knowledge of governing first principles. Bacher et al. applied Maximum Likelihood Estimation (MLE) to fit four different linear models of eight-room office building [BM11]. Yet, in that study the system was modelled in its whole, as a single-room office. Bondy and Parvizi presented instead a multi-room model for the same building following the same identification procedure [BP12]. In this latter study stability and accuracy issues were enlightened in relation to MIMO identification, as well as the fact that quality of data turned to be critical for convergence of estimation. Similar studies use thermal circuits as well [Old+12; Ma+12; Prí+11; Old+10; SBN14], all of them putting emphasis on the physical interpretation of the model, and the need of a pre-acquired understanding of the process to be modelled.

In this context, completely data-driven approaches are deemed interesting, sacrificing performance for practicality. Machine learning techniques have been gaining increasing attention in the past twenty years, since the availability of powerful computers and cheap memory allowed the pervasive use of artificial intelligence in a variety of applications that span from speech recognition to automated disease diagnosis, from image processing to data mining and intrusion detection in cyber-security applications [Bis+06; Has+05]. Zhao and al. provide in [ZM12] a review on modelling techniques, from engineering methods, to physical models, statistical models, and Machine Learning models.

Having all of these approaches specific pros and cons, it is difficult to state which one is the best in absolute sense. Hernandez Neto et al. in [NF08] compare the predicted building energy consumption between two approaches: detailed model-based, EnergyPlus, and Artificial Neural Networks (ANNs). It results that ANNs perform in average 3% better than Energy Plus with respect to the prediction error, due to the fact that ANNs can learn dynamics that stay unmodeled in EnergyPlus. However, EnergyPlus can provide building administrators with detailed information on how to improve building energy performance. An early application of ANNs to efficient building climate control is found in Neurobat [Mor+01], where ANNs are used to predict weather conditions and learn how that influences the internal temperature in a residential building, and Dynamic Programming [Ber+95] is used for optimal policy search.

This chapter presents two approaches based on artificial intelligence to building thermal modelling and control for DSM applications. The first method is the Model-assisted Dynamic Programming (MADP), which is detailed in Sec. 5.2 and consists of two phases: learning of system model via Extremely Randomized Trees (ERT), and policy computation via Dynamic Programming. In the optimisation enter also the forecasted disturbances on the system (solar radiation and external temperature) and dynamic energy pricing. A quick insight on ERTs and DP is presented in Appendix B.5 and B.6.

The second approach, presented in Sec. 5.3, is based on Model-assisted Batch

Reinforcement Learning (MABRL), and it is characterised by one main phase: concurrent model learning and computation of control policy. The optimal control policy is determined by jointly using data from the field and synthetic data from the support model, which is also obtained from the field data. As more data comes available from the field, the control policy produced by the MABRL algorithm converges to the optimum.

5.2 Model-assisted Dynamic Programming

The problem of building thermal modelling is here formalised as a supervised learning task of the function f:

$$T_i\left(t+1\right) = f\left(T\left(t\right), P\left(t\right), \phi\left(t\right), T_e\left(t\right)\right) \ , \tag{5.1}$$

where the function input is called *feature space*, and the function output *target value*. In this setting the target value is a scalar, i.e. the room temperature T_i at the next time step, and the input is a multidimensional space that includes measurements of internal temperature, solar radiation on the roof, ϕ, external air temperature, T_e, and power consumption of the air conditioners $P\left(t\right)$.

The experimental data is gathered in order to have a training set of twenty days and a test set of a couple of days. The one-second resolution data coming from VITO HomeLab is averaged over five minutes for the learning task in order to reduce the memory size (as detailed in Appendix B.5, the memory required by ERT is proportional to the size of the train data set). In Fig. 5.1 it is depicted the train dataset, where top graph reports the evolution of T_i, the second reports P, the third reports ϕ, and the fourth the T_e. The time step is 5 minutes.

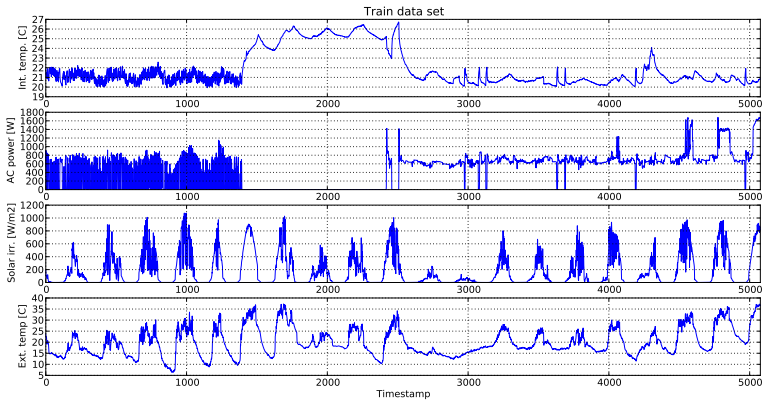


Figure 5.1: Train set for system identification via ERT.

The supervised learning task in Eq. 5.1 is carried using an ensemble of 100 ERTs trained on 5 minutes-discretised data with $n_{min} = 7$. [2]

The AC units are kept off for four days in order to be able to learn the uncontrolled system dynamics. In order to improve the model generalisation, in the learning phase the physical system should visits as many states as possible. Therefore, the AC units are controlled with a thermostatic control logic enhanced by *ε-greedy exploration* [SB98]:

$$u_k = \begin{cases} u \sim \text{Bernoulli}\,(0.5) & \text{if } \gamma \leq \varepsilon_j \\ u \text{ from thermostatic logic} & \text{if } \gamma > \varepsilon_j \end{cases} , \tag{5.2}$$

where $\gamma \sim U\,(0, 1)$. In this study the exploration factor ε_j reduces by half every four days:

$$\varepsilon_j = \frac{\varepsilon_0 \varsigma}{\varsigma + j - 1} , \tag{5.3}$$

where the decay factor ς is 4, the initial probability ε_0 is 0.4, and the the day index is $j \in \{1, 2, 3, \ldots\}$. In this way the temperature comfort bounds are always guaranteed and, thanks to the random actions, the controller has nonzero probability of selecting any control action in every encountered state. This technique improves the ERT regression performance, as the power consumed by AC units is in the feature space.

The regression performance of the ERT ensemble is depicted in Fig. 5.2 on a test data set (i.e. data that is not used in the learning phase), where the first graph shows the experimental data, in green, versus the model prediction, in blue.

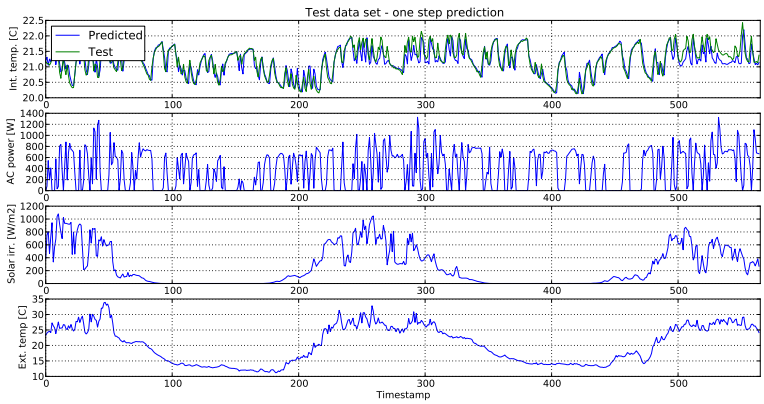
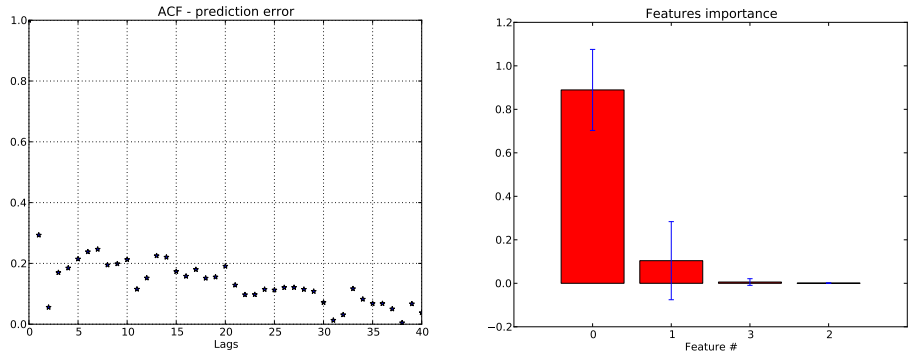


Figure 5.2: ERT forward prediction performance on a test set.

Figure 5.2 shows the model one-step prediction, while Fig. 5.3 presents the analysis of features importance and ACF of one-step prediction error.

[2] n_{min} is the minimum number of data points for which a node is not further split but it becomes a leaf. [Zho12] reports that a convenient choice of n_{min} for deterministic problems is 5, whereas for stochastic problems is 7.The choice of $n_{min} = 7$ is motivated by the noise affecting the measurements.



((a)) ((b))

Figure 5.3: ACF of model residuals (a), and features importance of Model 5.1 (b).

The residuals analysis via the ACF clearly shows that the ERT ensemble is not able to catch properly the slow dynamics. The analysis of features importance tells us that the temperature at a given time step is the most influencing feature for the prediction of the temperature at the following time step, while the second most important information is the power consumed by the air conditioners.

Denoting by x_t the feature space $(T_i(t), P(t), \phi(t), T_e(t))$, it is plausible to extend the proposed model with additional features from previous time lags (e.g. x_{t-1}, x_{t-2}, etc...). Fig. 5.4 shows the bias-variance analysis [Has+05] of the extended feature space, for time lags that go from 0 (reference case) up to 6. Increasing the feature space by one time lag allows reducing the RMS prediction error with respect to the reference case, yet further extensions of the model do not bring significant improvements.

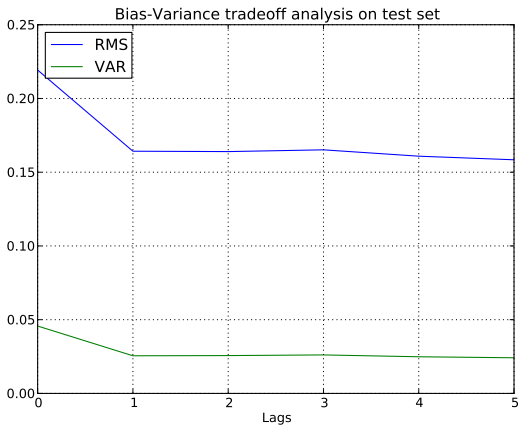


Figure 5.4: Bias-variance tradeoff for increasing lags of input features.

After performing the following model extension,

$$T_i(t+1) = f\left(T_i(t), P(t), \phi(t), T_e(t), T_i(t-1), P(t-1), \phi(t-1), T_e(t-1)\right) , \tag{5.4}$$

figures 5.5 and 5.6 show the one step prediction and residuals analysis.

The ACf in Fig. 5.6 (a) shows a slight improvement in the errors correlation, which has to be traded off with extending the DP optimisation problem by one additional dimension. As it will be clarified below, Dynamic Programming is a practical approach when the dimension of the state space is small (maximum three); therefore Model 5.4 is discarded in favour of Model 5.4.

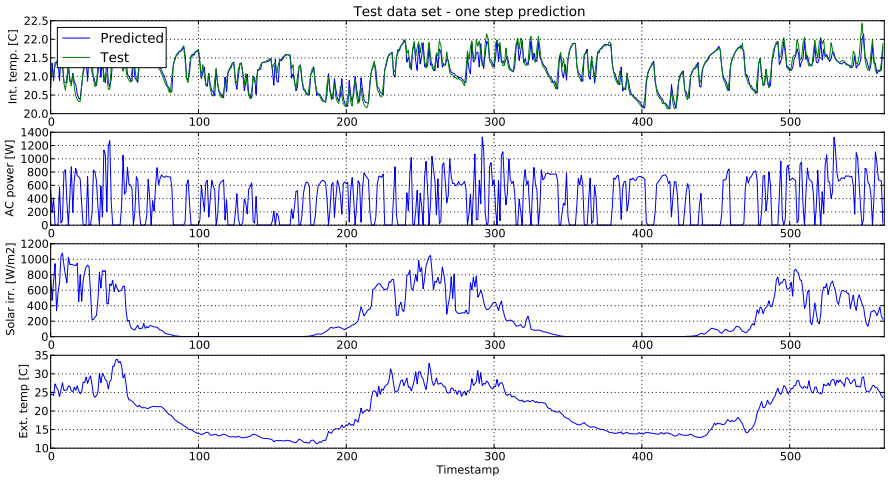
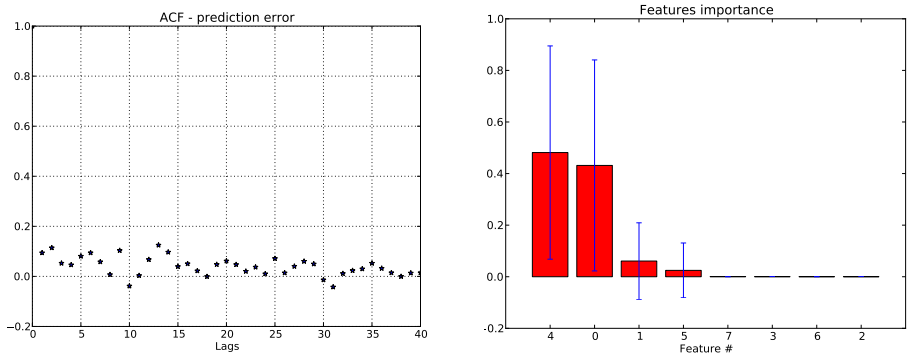


Figure 5.5: Train set for system identification via ERT. Extended feature space.



((a)) ((b))

Figure 5.6: ACF of model residuals (a), and features importance of Model 5.4 (b).

Once obtained the regressor, the optimal control policy is computed via Dynamic Programming, an optimisation technique that comes from Bellman's ingenious idea to cast an optimisation problem in terms of a multi-stage Markov Decision Process (MDP). In the DP setting the controller interacts with the system by means of three signals (Fig. 5.7): the state, which describes the state of the process under control; the control action, which leads the process to another state; the reward, which embeds information regarding the immediate performance of the control action and it can be seen as a design parameter that allows to specify the control objectives of the DP optimiser.

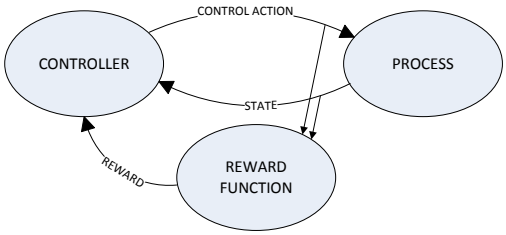


Figure 5.7: Flow of interaction in the DP setting.

DP is suitable for small-sized problems that exhibit the following properties [Bus+10]:

- overlapping subproblems - the problem can be broken down into subproblems;
- optimal substructure - the optimal solution to the original problem can be constructed efficiently from the optimal solutions to the subproblems.

The solution to a DP problem leads to the optimal control policy, called $h^*(\cdot)$ that, for any initial state x_0, maximises the total return:

$$u_k = h^*(x_k) \ , \tag{5.5}$$

where $(\cdot)^*$ denotes optimality. Notice the difference with the MPC, where a time-dependent open-loop control law was obtained. The total return associated to a generic policy h is given by the so-called Q-function [Bus+10], which corresponds to the sum of the rewards associated to the state transitions that occur starting from any initial state x_0 and following $h(\cdot)$ thereafter:

$$Q^h(x_0, u_0) = \rho(x_0, u_0) + \sum_{k=1}^{\infty} \gamma^k \rho(x_k, u_k) \ , \tag{5.6}$$

where $\gamma \in (0, 1]$ is a discount factor used to penalise rewards obtained after a large number of steps (it determines how far-sighted is the controller). Indeed γ allows trading-off the quality of the solution with the convergence rate of the DP algorithm.

The Markov property, which is at the basis of the theoretical guarantees of Dynamic Programming, allows DP to determine the next state and the reward by using only the current state and the current action together with:

1. the state transition function f, which determines the next state of the process given a specific control action:

$$x_{k+1} = f\left(x_k, u_k\right),$$

2. and the reward function ρ, which determines the reward associated to a specific control action taken at a specific state:

$$r_k = \rho\left(x_k, u_k\right).$$

It is worth to point out two facts about DP: 1) it does not need the analytical descriptions of f and ρ (they can be AI-learned regressors), and 2) it addresses the challenge of obtaining a long-term optimal control policy by using a reward function that gives only information about the immediate performance of a control action [Bus+10]. Buşoniu *et al.* in [Bus+10] present three classes of DP algorithms: Value Iteration, Policy Iteration, and Policy Search. Value Iteration uses the Bellman optimality equation to iteratively compute an optimal value function Q, from which an optimal policy is derived. Policy Iteration evaluates policies by constructing their value functions, and uses these value functions to compute new, improved, policies. Policy Search uses gradient-free optimisation to directly search for an optimal policy which maximises the return from any initial state.

The MADP controller proposed in this study is based on Value-iteration DP, which is further detailed in Appendix B.6, and validated in the HomeLab at the Flemish Institute for Technological Research (VITO). The setup consists of two air conditioners (Fig. 5.8), one temperature sensor to measure the room temperature in one point, one pyranometer to measure the solar irradiation on the roof, one temperature sensor to measure the external air temperature, and one power meter to measure the air conditioners power consumption. Note that, being the experiments carried during two different seasons, summer for MADP and winter for MABRL, cooling is considered for MADP and heating for MABRL.

As the experiments were carried in June, 2014, the targeted application is cooling. The time series of external disturbances is taken from the test set, depicted in Fig. 5.2, to produce the policy shown in the second chart of Fig. 5.9. The x axis represents the time, while the y axis the internal temperature discretised in 300 steps between 19.5 to 22.5 degree Celsius. This policy has been tested in VITO HomeLab, and the room air temperature is depicted in the top chart of Fig. 5.9.

The policy in Fig. 5.9 is a matrix where the columns represent the time, which is discretised by 5 minutes, and the rows represent the measured internal temperature. The air conditioners controller is operated by a script that dispatches the policy by reading it as a lookup table, where a 1 (red colour) corresponds to the control action *on*, and 0 (in blue) to *off*. The white dots represent a temperature reading corresponding to an activation of the AC units. Note that policy granularities are due to the fact that ERTs are discrete regressors, and the price profile is intentionally constructed in order to exhibit some variations in the envisaged time horizon.

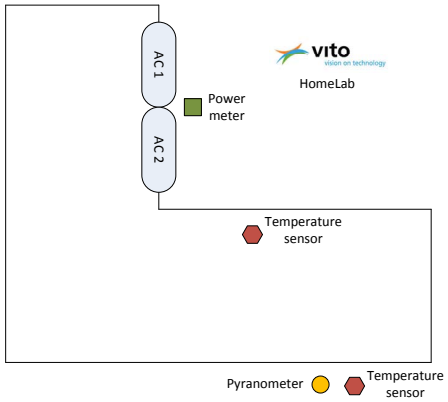


Figure 5.8: Experimental setup at Vlaamse Instelling voor Technologisch Onderzoek (VITO). Dimensions: 7.9m (L), 7.8m (W), 5m (H).

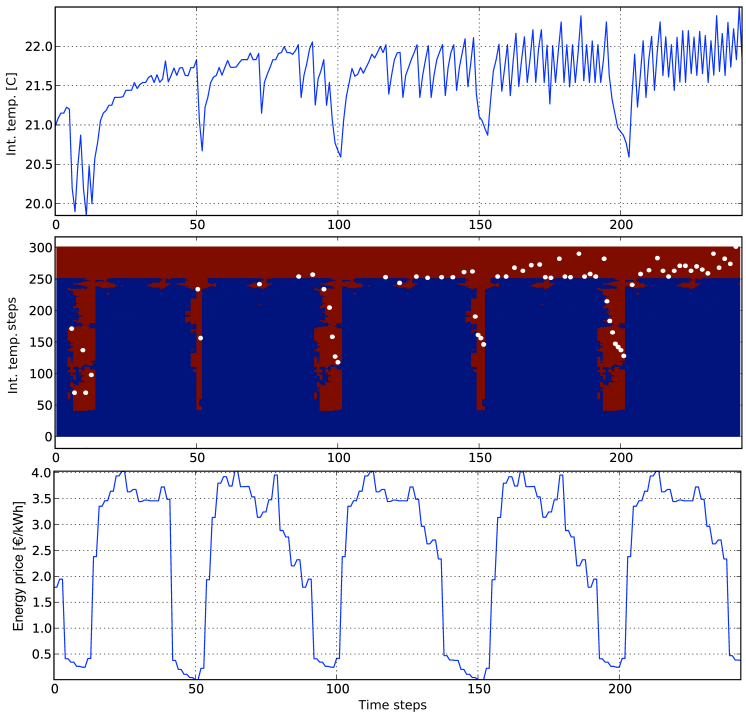


Figure 5.9: Test of optimal policy. The upper chart shows the room temperature, the one in the middle the policy map, while the lower one the energy price profile.

Summarising, the MADP approach here presented is characterised by an intrinsic simplicity in both modelling and optimisation. Modelling via ERT is a crucial procedure, since the regression performance is highly influenced by the choice of feature space, for which a feeling of the physical process is needed. Moreover, suitable discretisation time is chosen to trade off learning time, accuracy and use of memory. Concerning the optimisation, the used DP algorithm does not approximate the Q-function. Due to this, it is not practical to extend the state space, neither to further reduce the discretisation time. The policy produced exhibits granularities, as the ERTs provide a piece-wise linear model of the data, and the experience showed that twenty days of experimental data are needed for training.

In the next section we present an approach in which a policy-learning mechanism is designed instead of model-learning. Such algorithm is based on fitted Q-iteration Reinforcement Learning, and brings the novelty of using a support model to generate artificial trajectories of the physical system in support of the policy learning process. For that, this approach is called Model-assisted Batch Reinforcement Learning (MABRL).

5.3 Model-assisted Batch Reinforcement Learning

In general any learning task, supervised or non-supervised, benefits from large learning sets and extensive exploration of the system states. However, this is sometimes not possible due to learning time constraints and external contingencies that prevent the system under control to visit exhaustively its state space. In Sec. 5.2 we introduced how supervised learning can been used to obtain the process model, which is used afterwards to compute an optimal control policy.

An alternative approach is Reinforcement Learning (RL), a model-free control technique where an agent learns a control policy by interacting with the environment. Here the training information is obtained by evaluating the success or failure of different control actions on the system [Bar98] and, based on this information, the Q-function is constructed (Q-learning). Very similar to RL is Batch Reinforcement Learning (BRL), where the only difference with RL is that the policy is calculated offline using a batch of historical data. RL approach has been applied to HVAC systems in [HR11], and to heat pumps in [US13].

In order to reduce the convergence time of Q-learning, Model-assisted Batch Reinforcement Learning (BRL) has been proposed by Lampe and Riedmiller in [LR14] and by Ruelens *et al.* in [Rue+15]. MABRL uses expert knowledge, i.e. a support model, to generate trajectories outside the training set, improving the quality of control policies and, in general, speed up the convergence of RL. The basis of the presented approach is MABRL with fitted Q-iteration (FQI) [EGW05], which is used to obtain an approximation of the state-action function Q^* (Eq. B.24).

This section presents the original contribution of this thesis to AI-based control of DSRs for DSM. As previously stated, the target application is climate control, and Paper D presents an experimental analysis of data-driven control for building

heating system. Building upon [Rue+14; Bus+10; LR14], this work has the following contributions:

- BRL in the form of FQI [Ern+09] in combination with virtual trajectories [LR14] and policy shaping is applied to a HVAC system for a typical objective of dynamic pricing [FS10]. This effectively results in a data-driven solution for building climate control systems, combining state of the art BRL with domain knowledge;
- Quantitative and qualitative performance assessment of MABRL in a simulated and experimental environment, where the operation of an air conditioner is subject to dynamic energy pricing.

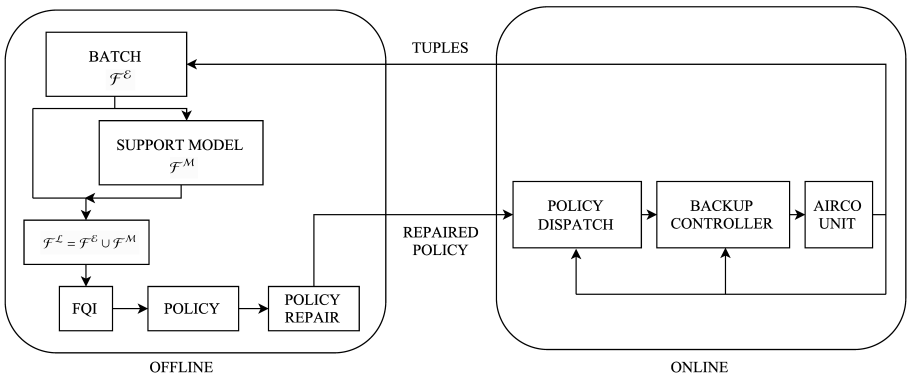


Figure 5.10: Overview of the information flow in this implementation of MABRL using FQI and policy repair.

Figure 5.10 shows the schematics of the proposed MABRL controller, where an offline phase and an online phase are depicted. The FQI algorithm is trained on the batch $\mathcal{F}^{\mathcal{L}}$, which is a combination of experimentally observed tuples $\mathcal{F}^{\mathcal{E}}$ and virtual tuples generated by a model $\mathcal{F}^{\mathcal{M}}$:

$$\mathcal{F}^{\mathcal{L}} = \mathcal{F}^{\mathcal{E}} \cup \mathcal{F}^{\mathcal{M}} \,. \tag{5.7}$$

The batch $\mathcal{F}^{\mathcal{E}}$ is used to update or re-train the support model, which is used to generate random data points. These are joined to the batch if the nearest experimental tuple in $\mathcal{F}^{\mathcal{E}}$ falls outside a predefined radius r, following a distance metric Δ. Then, the FQI algorithm generates a control policy using the batch $\mathcal{F}^{\mathcal{L}}$. The control policy is finally filtered to reduce the granularity before being dispatched to the air conditioners.

In this work, following the approach presented in [LR14], an ANN is used to represent a support model consisting of single-layer, single-output Extreme Learning Machines (ELMs) trained to predict the change of the internal temperature ΔT. ELMs

are used as they allow for fast training of the weights of the network at the expense of reduced regression performance. The latter is partially mitigated by combining multiple ELMs in an ensemble [Zho12].

The output of an ANN with p input neurons and n hidden neurons can be formulated as:

$$y(\mathbf{x}) = \sum_{i=1}^{n} \beta_i g_i(\mathbf{w_i} \cdot \mathbf{x},\, b_i) = \mathbf{G(x)}\boldsymbol{\beta}\,, \tag{5.8}$$

where $\mathbf{x} \in \mathbb{R}^p$ is the input vector, $\mathbf{w_i} \in \mathbb{R}^p \sim i.i.d.\, U(-1,1)$ is the weight vector connecting the input nodes with the i-th hidden node, $b_i \in \mathbb{R} \sim U(0,1)$ is the bias of the i-th hidden node, and $\beta_i \in \mathbb{R}$ is the output weight of i-th hidden node. $\boldsymbol{\beta} = [\beta_1 \dots \beta_n]^T$ is the output weight matrix, while $\mathbf{G(x)}$ is the output matrix of the hidden nodes, where the nodes activation function g is a sigmoid.

Since the parameters of the hidden nodes are randomly generated, training an ELM corresponds to determining the output weight matrix $\boldsymbol{\beta}$ based on the regularised least-squares solution of $\mathbf{G}\boldsymbol{\beta} = \mathbf{Y}$. As the training process is fast, finding the appropriate number of hidden nodes is done using cross-validation. Finally, the ensemble model output is given by a weighted average of the individual ELMs outputs. For a deeper insight on ELMs, we refer to [Cam+13], while the support model is fully detailed in Paper D.

In MABRL the problem of building climate control is formalised as a stochastic Markov decision process (MDP). This formulation is chosen with respect to the deterministic case in Sec.5.2 in order to account uncertainty on the prediction of system disturbances. Leaving the details on the MDP formulation to Paper D, it is worth remarking that an optimal control policy, here denoted by h^*, satisfies the Bellman equation:

$$J^{h^*}(\mathbf{x}) = \min_{\mathbf{u}} \mathop{\mathbb{E}}_{\mathbf{w}\sim P_w(.|\mathbf{x})} \{\rho(\mathbf{x}, h(\mathbf{x}), \mathbf{w}) + J^{h^*}(f(\mathbf{x}, h(\mathbf{x}), \mathbf{w}))\}\,, \tag{5.9}$$

which in this case is characterised by an expectation of the T-stage return. As for DP, typical techniques to find policies in an MDP framework are value iteration, policy iteration, and policy search [Bus+10]. In this work, value-iteration is considered.

Below a qualitative and quantitative assessment of the MABRL controller is presented. As the experimental setup is a living lab, exact external conditions cannot be reproduced from day to day. Thus, a model of the air conditioner and the lab room is used instead of the living lab for a quantitative assessment, and the thermal scheduling obtained using MPC is taken as benchmark. Figure 5.11 presents the economic performance of MABRL with respect to the standard thermostatic control (Default), the BRL, and the reference benchmark, MPC. One can depict that MABRL is able to find near optimal control policies in a learning time of approximately 20 days, after which the performance relative to a mathematical optimum is stable. However, adding virtual samples has limited contribution, with a slight economic advantage of the MABRL approach over the BRL.

In order to validate the MABRL convergence and performance with experimental data, the policy computation with respect to different training sets is performed and

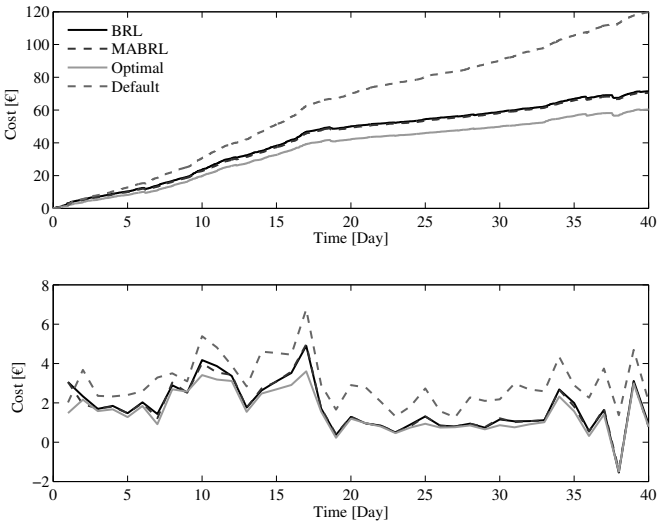


Figure 5.11: Cost performance of the different controllers: BRL (no virtual tuples), MABRL, Optimal (MPC), and Default (hysteresis). Top plot: cumulative electricity cost. Bottom plot: daily electricity cost.

shown in Figure 5.12. Here policies are organised in a matrix for different predicted profiles of the external temperature (from left to right) and increasing size of the learning batch (from top to bottom). No virtual tuples have been added. Starting from the top row, policies are computed using experimental batches of: two days, eight days, and sixteen days. The black areas in the policies correspond to the control action *switch on*. Consider Figure 5.12, from left to right only the profile of external temperature is changed. One can observe that the policies are highly influenced by the weather conditions: as soon as the external temperature is low, the policies correspond to basic thermostatic control (i.e. heat only when it is necessary). Thus MABRL effectively takes into account the forecast of exogenous information. Finally, note that the policies in the first row resemble each others, regardless of the different forecast of the external temperature, due to the fact that the training tuples are few.

Figure 5.13 shows the effect of shaping the policy, as discussed in Paper D. Here triangular membership functions [Bus+10] have been used with the constraint that the policy needs to be strictly decreasing with increasing indoor temperature. This is a direct consequence from the physical understanding that a room at a higher temperature is subjected to higher losses to the environment. The results depicted in Figure 5.13 show that the policies are smoothed following the monotonicity constraints.

Figure 5.14 shows the impact of virtual tuples on the convergence of the policy. The policy is computed using experimental data and virtual tuples in three different

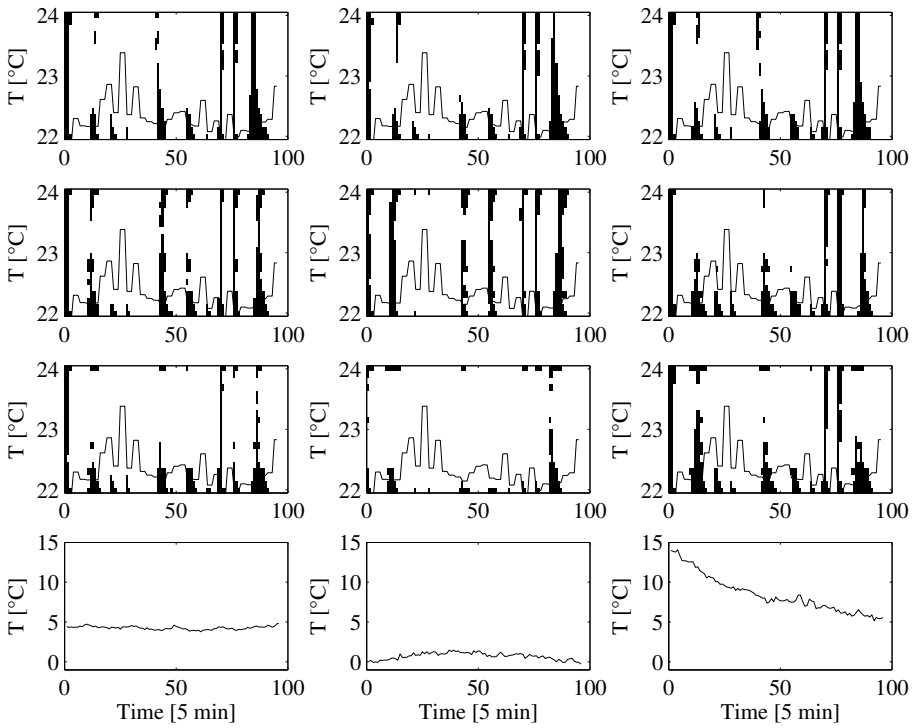


Figure 5.12: Policies projections obtained from experimental data for different forecast of the outside temperature, depicted in the lower row. From top to bottom data from 2, 8 and 16 days has been used. The lowest row indicates the forecasts of outside temperature for which the policies have been calculated.

proportions. The left chart in Figure 5.14 shows the computed policy with only few experimental data points (2 days) and no virtual tuples in support, called *early policy*. The right chart shows the policy computed using a larger set of experimental data (20 days), without virtual tuples, called *regime policy*. The middle chart shows the computed policy using few field data (2 days) and a large share of virtual tuples (400 samples) from the support model, which is trained on the basis of the experimental data available. This latter is called *model-assisted policy*. One can recognise that the model-assisted policy is much more similar to the regime policy than the early policy, concluding that the support model enhances the policy convergence. Furthermore, adding virtual tuples results in more smooth policies.

Figure 5.15 shows the implementation of a policy obtained after 12 days of training. The bottom chart depicts the internal temperature and the control policy, where the black areas correspond to the control action *switch on*. Note that, in order to avoid frequent switching of the air conditioners, the heating is kept on until the temperature

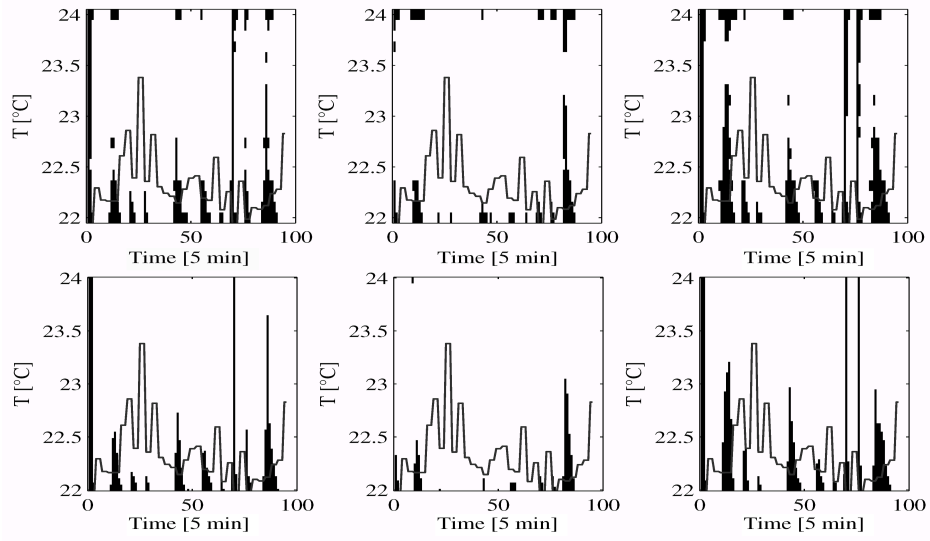


Figure 5.13: regime policies. Original form (row n. 3 from Figure 5.12), in the upper row, versus repaired policies, in the lower row.

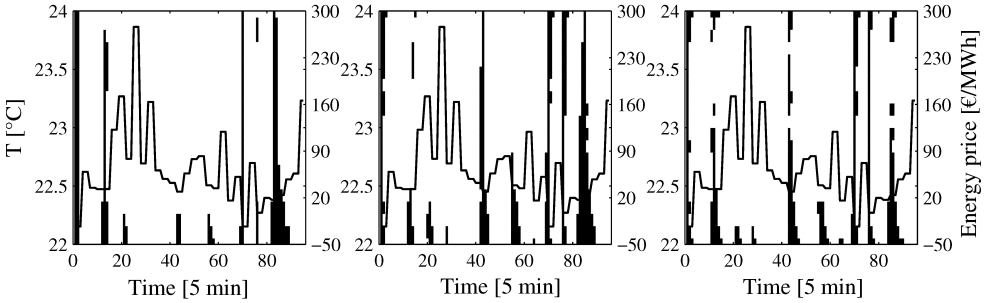


Figure 5.14: Computation of the closed-loop policies for different shares of virtual tuples over experimental data.

upper bound is reached. The middle chart reports the power consumption of the air conditioning units and the intraday balancing price. The second and first chart shows the external air temperature.

Figure 5.16 shows the implementation of a regime policy on a warmer day. In this case, since the thermal losses are moderate with respect to a cold day, the optimiser allows longer pre-heating in order to avoid high price zones. This doesn't occur in cold days, where the policy is much more similar to a simple thermostatic logic.

Summarising, the analysis based on simulations and experimental data showed that the policy convergence time for MABRL is approximately five times smaller

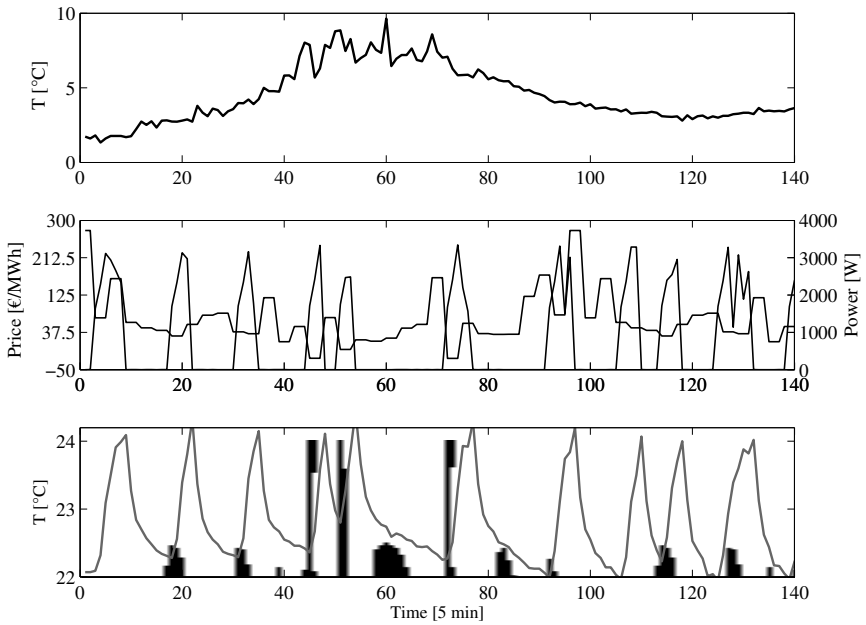


Figure 5.15: Experimental results 1. Top plot: outside temperature. Middle plot: consumed power by the HVAC and electricity price. Bottom plot: control policy obtained with MABRL and the resulting indoor air temperature.

than the one of BRL. Both techniques reach a performance within 90% of a mathematical optimum in approximately 20 days without grey-box system identification, and both controllers are immediately deployable. The MABRL technique, already introduced in literature, is here adapted and validated in a living lab for a climate control application, which constitutes the original contribution presented in Paper D.

5.4 Conclusions

This chapter presented two techniques for building climate control, Model-assisted Dynamic Programming and Model-assisted Batch Reinforcement Learning. MADP consists of two separate procedures that are both carried offline, model learning, and policy computation. Differently to BRL, MABRL learns directly the control policy using experimental data and virtual data coming from a support model. This allows faster policy convergence and controller deployment time.

Some time was spent for gathering the experimental data and troubleshoot all the practical problems that arose when coupling the air conditioners control software (system excitation and policy dispatch) with the physical setup. MADP shows reasonable overall performance, at the price of a learning period comparable to grey-box physical

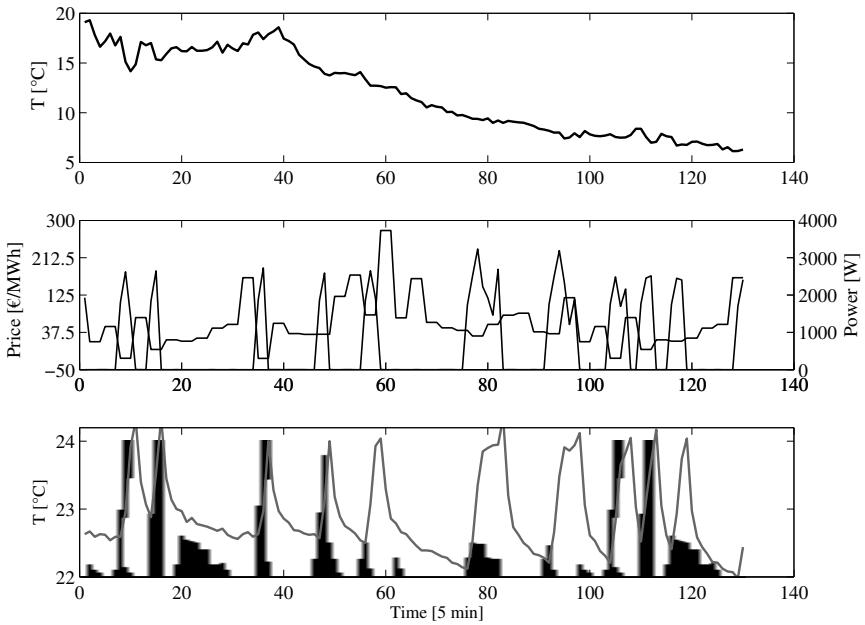


Figure 5.16: Experimental results 2. Top plot: outside temperature. Middle plot: consumed power by the HVAC and electricity price. Bottom plot: control policy obtained with MABRL and the resulting indoor air temperature.

modelling. However, the model baed on ERT requires a sufficient understanding of the system in order to properly define the learning task.

Therefore Model-assisted Batch Reinforcement Learning is investigated, which features faster learning as no accurate model has to be trained in advance. Simulation studies and experimental tests show that the control policies obtained with MABRL converge to a configuration that is very similar to the policies obtained with BRL, in a fifth of the time. However, even if both BRL and MABRL offer significant savings with respect to the default thermostatic control, no significant economic gain has been observed in this setting between BRL and MABRL.

Finally, the policies obtained from both BRL and MABRL showed that the thermal inertia of a relatively-small lab may be sufficient to respond to imbalance prices, but not to day-ahead prices. A similar study on larger buildings would definitely be an interesting continuation of this research.

CHAPTER 6

A model-free approach to Direct Load Control

"Design is not just what it looks like and feels like.
Design is how it works."

Steve Jobs

This chapter presents an approach for abstracting the flexibility of thermostatically controlled loads and using it in a model-free aggregation framework for DLC, called COMMELEC, that was developed at Ecole Polytechnique Federale de Lausanne, outside the context of this research. After introducing the the basics of COMMELEC, the original contribution of this research is presented, which consists in an innovative algorithm for Load Admission Control that allows abstracting the flexibility of thermostatically controlled loads and using it within COMMELEC. A simulation study concludes this chapter, providing a proof of concept of the presented approach and algorithm.

6.1 Introduction

Two approaches based on predictive control for DSM via DLC have been presented so far; the first uses standard grey-box approach for modelling and quadratic programming for predictive control, while the second uses AI techniques for modelling and gradient-free optimisation for predictive control.

This chapter is dedicated to the design of a model-free control strategy for interfacing thermostatically-controlled loads (TCLs) with a framework for real-time control of active distribution networks by means of explicit power set points. This framework, called COMMELEC, offers a composable approach to frequency and voltage control in medium and low voltage grids with a large penetration of controllable loads and distributed and stochastic production [Ber+14b].

The algorithm introduced here, called Load Admission Control (LAC), allows electric space heating scheduling for Direct Power Control (DPC) applications, and it can be directly extended to any TCL. Being it model-free and based on online scheduling, LAC is of easier scalability and deployment compared to MPC-based schemes. The comfort of the end customer is prioritised with respect to the DR service provision, which allows exploiting the TCLs only within their local comfort settings.

After a brief introduction to COMMELEC in Sec. 6.2, the main concepts for abstracting TCLs flexibility and the LAC algorithm are presented in Sec. 6.3 and further detailed in Paper E (Appendix A). Section 6.4 outlines the final conclusions.

6.2 An introduction to COMMELEC

This framework, called COMMELEC [Ber+14b], is based on request/response communication protocol between the devices controllers (Resource Agents, RAs) and the grid controller (Grid Agent, GA). The RAs export generic and standardised information about devices status to the GA that computes, in real-time, the optimal power injections at the grid nodes that the various RAs are requested to realise. Thus, the design of a Resource Agent is specific to device type, whereas the GA is totally generic and can be adapted to any network. Thanks to this standardised interface between the agents, it is possible to completely disregard devices details in the overall network control and easily upscale (or downscale) the entire architecture. Each agent speaks for and controls the subsystem under its responsibility, which can consist of other electrical grids or resources (loads, generators and storage devices), by using explicit power set points. An agent can be assigned the role of leader of one or more other agents, which we term the followers of that leader.

The agents communicate with each other by using a simple Advertisement/Request protocol, and using some simplified quantitative information about their devices capabilities and internal states. In particular, the follower agents periodically advertise to their leader the following three elements: (i) PQt profile, that is a region in PQt space of setpoints that the subsystem is willing to implement; (ii) Virtual

cost function, that acts as proxy for the internal state of the system, the preferred setpoints, and how close the subsystem is close to its operational constraints; and (iii) Belief function, representing all the possible values of (P;Q) that the subsystem may implement in practice when it receives a certain setpoint.

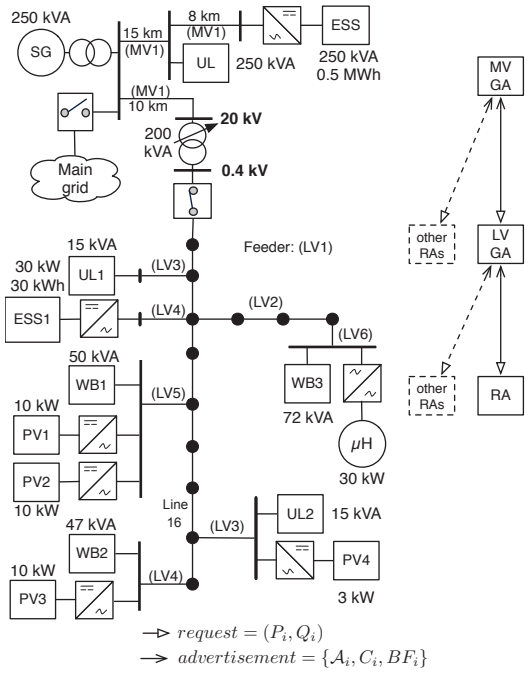


Figure 6.1: Reference distribution network. The LV part of the grid is composed by the CIGRE LV microgrid benchmark proposed by TF C6.04.02 [P+05]. The MVGA is the leader of one or more LVGA, which themselves are leaders of one or more RAs (dotted arrows).

Observe the difference between PQt profile and belief function: the former indicates the setpoints that this subsystem is willing to receive, whereas the latter indicates all the possible operating conditions that may result from applying a received setpoint. The PQt profile is used by the grid agent to compute an "optimal" setpoint, whereas the belief function is used to determine the region of the safe operation of the grid. All agents monitor in real time the state of the grid device(s) and/or the agents (the GAs) under their responsibility and compute the general messages to send to their leader agent. When receiving a new request, each agent projects the power setpoint to the current updated PQt boundaries in order to ensure its individual safety. We refer the reader to [Ber+14b] for further details on the protocol and the computation of the setpoints of the GA.

6.3 Space heating scheduling for real-time explicit power control

The algorithm here developed for DLC is based on a modified version of the Load Admission Control (LAC) algorithm presented in [Cos+12b]. The LAC algorithm has been previously tested in experimental trials for the PowerMax service with Power FlexHouse [Cos+12a], and it has been here modified and extended in order to: 1) abstract from the resource state the information needed in the COMMELEC protocol, and 2) perform Direct Power Control service, since the RAs in COMMELEC have to implement specific power set points rather than limit the power to a maximum level.

In the following, we refer to the agent for TCLs as Load Agent (LA), which generality and model-independent design this section is intended to enlighten. The TCL considered in this study is a building with its space heating system. However, the presented LA can be adapted to TCLs of any size and kind, provided that the thermal control system allows controlling the temperature independently in each control volume, where at least one measurement of internal temperature is available. Note that this approach is model-free in the sense that no model of the TCL thermal inertia and insulation is needed; however a model of the actuator is necessary, even a trivial one. In the case considered here, being the heating modules resistive electric heaters, the actuators are accounted in the control algorithm as controllable constant-power loads.

The LA advertises three information to the GA: PQt profile, Belief function, and Virtual Cost function; then it receives back the power set point from the GA to be implemented. The computation of the load **PQt profile** is directly related to the notion of flexibility, which a property of the load that allows it to be interrupted and re-started again without causing user discomfort or device failure. Specifically, the flexibility of a single room is determined by its internal air temperature: if it is within the comfort bounds, the heater can be switched on or off without compromising the user comfort. Instead, if the room temperature is below the lower comfort bound (T_{LB}) the heater *must* run, and if the temperature is above the upper comfort bound (T_{UB}) the heater *must not* run (Fig. 6.2)

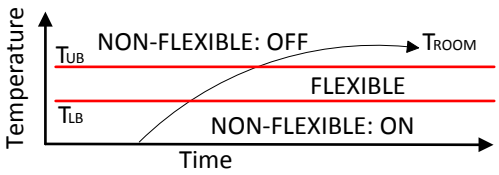


Figure 6.2: The flexibility of operation of a heating system is related to the system state itself.

The flexibility of the whole system (building) is determined by its subsystems (rooms) flexibilities. The rooms whose temperature is below T_{LB} determine the building minimum power consumption, while the maximum consumption is limited by those rooms whose temperature is above T_{UB}. The (P, Q) flexible region is determined by those rooms which temperature is within the comfort requirements $[T_{LB}, T_{UB}]$. One can immediately see the potential scalability of this computation.

The **belief function**, henceforth denoted by B, describes the ability of the Load Agent to map a requested power setpoint, u, coming from the Grid Agent, to actual load consumption. In this study, the building is composed by eight rooms with one electric space heater of each. The heaters can be either switched on or off, therefore the load can implement only discrete active power setpoints from $0kW$ to $8kW$ in steps of $1kW$. Hence the *belief function* simplifies to the *nearest integer* function:

$$B(u) = \{nint(u)\} , \tag{6.1}$$

where u is a power setpoint received from the GA. Note that, thanks to the LAC algorithm, the LA achieves tracking a power setpoint from the GA with an accuracy of $\pm 0.5kW$ (Fig.6.3).

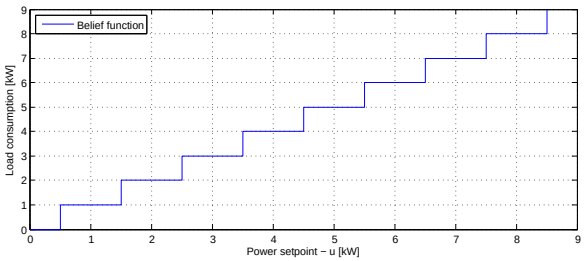


Figure 6.3: Example of belief function for discrete load steps of $1kW$

The **Virtual Cost function** serves as a proxy for load internal conditions and it associates a cost to every operating point inside the PQt profile. The cost is computed per every single subsystem (room) with respect the possible control actions: *heat*, and *not heat*. Since the LA responds to a received power setpoint by switching on the heaters in some rooms and switching off the heaters in the remaining rooms, we compute the total cost of a received power setpoint as the cost associated to heating the selected rooms, plus the cost associated to not heating the remaining rooms.

In each room the cost of heating is zero when the room temperature is in the middle of the comfort band, T_M, positive when it is above, and negative when it is below. This design choice is based on the consideration that the building, or any TCL, exhibits flexibility for longer time if the temperature in all control volumes, in this case the rooms, is the middle of the comfort zone(s). Equation 6.2 shows the cost of heating a specific room j given the internal temperature T_j:

$$c_h\left(T_j\right): \begin{cases} 1+\left(T_j-T_{UB_j}\right)^2 & : T_j \geq T_{UB_j} \\ \dfrac{T_j-T_{M_j}}{T_{UB_j}-T_{M_j}} & : T_{LB_j} < T_j < T_{UB_j} \\ -1-\left(T_{LB}-T_j\right)^2 & : T_j \leq T_{LB_j} \end{cases}, \tag{6.2}$$

while the cost of not heating the same room is given by $-c_h\left(T_j\right)$.

Note that in this case study only active power operating points are considered, and the LA does not need any information about the building model to compute the cost function. The Load Agent uses the temperature readings to compute the terms c_h, then it calls the LAC algorithm (presented in Paper E) to compute the control vectors U for one admissible power setpoints $x \in PQt$. To each x corresponds the cost $C\left(x\right)$:

$$C\left(x\right) = \sum_{j \in \mathcal{R}} \left[U\left(j\right) c_h\left(T_j\right) - \left(1-U\left(j\right)\right) c_h\left(T_j\right)\right], \tag{6.3}$$

where j is a generic control volume belonging to the set of control volumes $\mathcal{R}$, and $U\left(j\right)$ is the component of the control vector corresponding to room j. The entire Virtual Cost function for the load is obtained by computing the control vector U for all setpoints in the PQt profile and iterating Eq. 6.3 over the obtained control vectors.

The proposed algorithm is tested in the COMMELEC simulator, developed at EPFL, which grid configuration is depicted in Fig 6.1. Two case studies are presented below, and the DSRs on which we shall focus are: the building (UL2), the water boiler (WB2) and the PV (PV3), depicted in Fig. 6.1. From the simulation it emerges that the LV Grid Agent dispatches the power setpoints in order to avoid congestions in Line 16, which connects the considered DSRs with the rest of the network.

In the first simulation, the PQt profile of the LA is initially between $-6kW$ and $-8kW$, as six rooms out of eight are below the comfort zone (Fig. 6.4). The LA is able to track the explicit power setpoint coming from the LVGA and the building flexibility increases as soon as all the rooms enter the comfort zone (upper chart). At this time, the LVGA fully exploits the load flexibility when dealing with local resource contingencies and line constraints.

In the second simulation, due to the initialisation of the internal temperature in all the rooms within the comfort zone, the PQt profile is constrained between $0kW$ and $-8kW$ (Fig. 6.5), in a way that the building exhibits the maximal flexibility to the LVGA from the beginning. Figure 6.5 shows also that only the rooms in the lower part of the comfort zone are heated, while the LVGA operates the water boiler WB2 in advance with respect to the previous simulation. When WB2 is switched on, the LA is requested to reduce the building consumption in order to not overload the feeder. At about 1200s, the decreased PV production brings the LVGA to reduce the WB2 power and restore the building heating so that all the rooms steer toward the middle of the comfort zone (i.e. to the minimum virtual cost).

Since the PQt profile reports the admissible operating power of the resource, assuming that only power set points within this region are implemented, the user comfort is always guaranteed, and the flexibility information upon the GA is updated

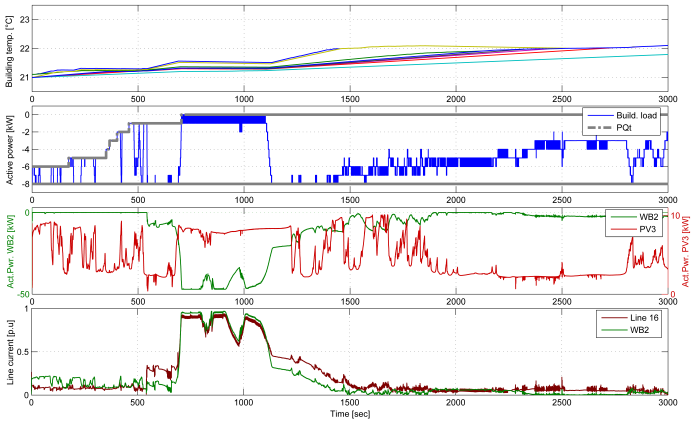


Figure 6.4: Simulation with LA - rooms initialized below the comfort zone. From the top chart: temperature of building rooms, building power consumption, WB2 consumption and PV3 production, building and connecting line (Line 16) currents.

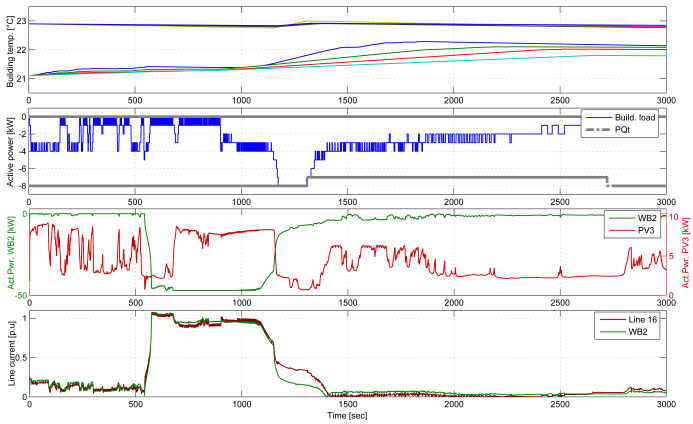


Figure 6.5: Simulation with LA - rooms initialised within the comfort zone. From the top chart: temperature of building rooms, building power consumption, WB2 consumption and PV3 production, building and connecting line (Line 16).

at every advertisement/request iteration cycle which, in the case of COMMELEC, is 100ms. Please refer to Paper E for further details on the control algorithm and the entire Load Agent, and to [Ber+14b] for further details on COMMELEC.

6.4 Conclusions

This chapter presented a systematic approach to abstract the flexibility of a building heating system and interface it with the composable framework for real-time grid control, COMMELEC, i.e. to be able to use a building's thermal inertia for real-time control. The presented simulations show the efficacy of the LA to operate the load within the user comfort requirements and communicate generic information to the GA to perform explicit power control.

The main benefit brought by this approach consists in the simplicity on the Load Agent side, and the accomplishment of a variety of tasks, such as voltage and frequency control, thanks to the design of the control framework and the Grid Agent. However, it remains open the question on how to assess the cluster available flexibility in terms of shiftable energy. This latter topic is certainly of relevant importance and interest, and it is under investigation of the COMMELEC developing team at EPFL.

CHAPTER 7

Conclusions

Through this thesis the reader is introduced to the concepts of Smart Grids, Demand Side Management, and Direct Load Control. As discussed in Ch. 2, political, economical and strategical reasons brought energy independency and sustainability high on the agendas of many governments. Together with an increase of non dispatchable production, not only grid stability and power quality issues arise, but also energy markets and regulations are affected. Smart Grids enhance the functionality of the power networks by using information technology, data analysis and control theory, in any possible fashion to countless applications. In this thesis Demand Side Management is further investigated towards residential applications, with a mainly focus on Direct Load Control. Three mainstreams are investigated: physical grey-box modelling with model predictive control, machine learning-based modelling and predictive control, and model-free real-time control.

Chapter 4 presents physical modelling via grey-box system identification. Such approach is very intuitive if the system governing principles are known, and it can be used with efficient optimisation techniques for predictive control. However, carrying out proper identification is time and resource demanding, and when it comes to aggregation, the size of MPC problems grows rapidly, especially in centralised control solutions. For this reasons distributed control architectures, which scale better, are deemed interesting. A novel algorithm for distributed predictive control, called DSMPC, is introduced. It offers the possibility of independent controller design, provided that few requirements are satisfied, it can handle binary, integer and continuous control variables, it does not rely on a central coordinating optimiser, and it provides a best-effort solution in few iterations. Yet, the main drawback of this algorithm stays in the scalability which, although being better than the centralised solution, is connected to the sequential and iterative communication between units.

Chapter 5 introduces an alternative modelling approach that aims at tackling the issues of long identification time and necessity of grey-box physical modelling by means of artificial intelligence. Two techniques are investigated and presented: Model-assisted Dynamic Programming (MADP), and Model-assisted Batch Reinforcement Learning (MABRL). MADP consists of a preliminary modelling phase and a successive optimisation phase. System modelling is carried with Extremely Randomized Trees, which features simplicity and intuitively, but generalise less than finely-tined tray-box models and are highly memory demanding. Based on ERT model, Dynamic Programming is used to compute the optimal closed-loop control policy with respect to dynamic energy pricing. A DLC scheme in this context, which is not truly in-

vestigated here, may be implementable by means of simulation-based optimisation. The second solution investigated is MABRL, which allows computing directly an optimal policy from a learning data set in a fifth of the time with respect to Batch Reinforcement Learning (BRL).

Chapter 6 presents the third and last approach analysed in this research: model-free real-time control. This approach offers a systematic way to abstract the flexibility of thermal loads and use it in a framework for control of active distribution networks. The unit controller uses information on the internal temperature and the comfort settings in order to compute the load flexibility. Yet, the model-free approach cannot be readily used for predictive control, and the estimation of cluster flexibility constitutes an interesting further development.

From the results presented in this research it is evident that, depending on the envisaged application and in-domain knowledge, some techniques result more applicable and viable than others. As seen, MPC can be preferable in case the DSR model or a deep in-domain knowledge is available. Alternatively, techniques based on artificial intelligence offer an attractive alternative for large-scale deployments. The question whether or not a generic and universally-applicable approach to DSM can be found remains open after this research. However, it is conviction of the author that good knowledge of different techniques and their applicability can lead to innovative and brilliant results on system integration.

APPENDIX A

Paper collection

This chapter includes the publications submitted as part of this thesis, to which the different chapters refer to. A list of related publications is also included. These last ones, even though having high affinity to this thesis, do not constitute part of it. For each paper we report the title, the journal or conference, and the publication date.

Publications part of the thesis

- Paper A - Grey-box Modeling for System Identification of Household Refrigerators: a Step Toward Smart Appliances. 4th IEEE Youth International Conference on Energy, 2013. DOI: 10.1109/IYCE.2013.6604197. Authors: Giuseppe Tommaso Costanzo, Peder Bacher, Fabrizio Sossan, Mattia Marinelli, and Henrik Madsen.
- Paper B - A Coordination Scheme for Distributed Model Predictive Control: Integration of Flexible DERs. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2013. DOI: 10.1109/ISGTEurope.2013.6695474. Authors: Giuseppe Tommaso Costanzo, Oliver Gehrke, Esteban Morales Bondy, Fabrizio Sossan, Jacopo Parvizi, Henrik Madsen, and Henrik W. Bindner.
- Paper C - Performance Assessment of Aggregation Control Services for Demand Response. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2014. DOI: 10.1109/ISGTEurope.2014.7028779. Authors: Daniel Esteban Morales Bondy, Giuseppe Tommaso Costanzo, Kai Heussen, and Henrik W. Bindner.
- Paper D - Experimental analysis of data-driven control for a building heating system. Submitted to Elsevier Sustainable Energy, Grids and Networks (SEGAN) on July 14, 2015. Available at: http://arxiv.org/abs/1507.03638.. Authors: Giuseppe Tommaso Costanzo, Sandro Iacovella, Frederik Ruelens, Tim Leurs, and Bert J. Claessens.
- Paper E - Electric Space Heating Scheduling for Real-time Explicit Power Control in Active Distribution Networks. IEEE Innovative Smart Grid Technolo-

gies (ISGT) Europe, 2014. DOI:10.1109/ISGTEurope.2014.7028798. Authors: Giuseppe Tommaso Costanzo, Andrey Bernstein, Lorenzo Reyes Chamorro, Henrik W. Bindner, Jean-Yves Leboudec, and Mario Paolone.

Other publications related to this thesis

- Paper F - An Experimental Study on Load-Peak Shaving in Smart Homes by Means of Online Admission Control. IEEE Innovative Smart Grid Technologies (ISGT) Europe, 2012. DOI: 10.1109/ISGTEurope.2012.6465658. Authors: Giuseppe Tommaso Costanzo, Anna Magdalena Kosek, Guchuan Zhu, Luca Ferrarini, Miguel F. Anjos, and Gilles Savard.

- Paper G - A System Architecture for Autonomous Demand Side Load Management in Smart Buildings. IEEE Transactions on Smart Grid, Vol. 3, no. 4, December 2012. DOI: 10.1109/TSG.2012.2217358. Authors: Giuseppe Tommaso Costanzo, Guchuan Zhu, Miguel F. Anjos, and Gilles Savard.

- Paper H - Power Admission Control With Predictive Thermal Management in Smart Buildings. IEEE Transactions on Industrial Electronics, Vol. 62, no. 4, April 2015. DOI: 10.1109/TIE.2014.2387091. Authors: Jianguo Yao, Giuseppe Tommaso Costanzo, Guchuan Zhu, and Bin Wen.

Grey-box Modeling for System Identification of Household Refrigerators: a Step Toward Smart Appliances

Giuseppe Tommaso Costanzo,
Fabrizio Sossan and Mattia Marinelli
Center for Electric Power and Energy
Technical University of Denmark, Risø Campus
4000 Roskilde (DK)
{guco, faso, matm}@elektro.dtu.dk

Peder Bacher and Henrik Madsen
DTU Compute
Technical University of Denmark
2800 Lyngby (DK)
{pb,hm}@imm.dtu.dk

***Abstract*—This paper presents the grey-box modeling of a vapor-compression refrigeration system for residential applications based on maximum likelihood estimation of parameters in stochastic differential equations. Models obtained are useful in the view of controlling refrigerators as flexible consumption units, which operation can be shifted within temperature and operational constraints. Even if the refrigerators are not intended to be used as smart loads, validated models are useful in predicting units consumption. This information can increase the optimality of the management of other flexible units, such as heat pumps for space heating, in order to smooth the load factor during peak hours, enhance reliability and efficiency in power networks and reduce operational costs.**

Keywords*—*Refrigerators, Stochastic processes, System identification, Load shifting.

I. Introduction

The World Business Council for Sustainable Development estimates that in most countries buildings account approximately for the 30-40% of total energy consumption [1]. Energy consumption in a building can be related to such applications as space heating and building automation (including security systems and ICT infrastructures) or to human activities. It emerges that controlling loads with building automation systems can enhance the overall demand flexibility and enable a win-win situation, where customers adjust their consumption upon economic inducements and utilities avoid grid overloads by spreading the demand during off-peak periods [2]. In this context, validated models of appliances are necessary in the design of systems for residential demand side management and in testing and benchmarking controllers for energy consumption in Smart Buildings.

Devices or processes associated to thermal storage present intrinsic flexibility in consumption as long as their operation is managed within certain comfort bounds. One example is space heating, which can be used for peak shaving [3], but also other types of thermal storages (such as refrigerators or water chillers) offer flexibility in consumption.

This paper presents the grey-box modeling of a vapor-compression refrigeration system for residential applications using stochastic differential equations (SDEs). The grey-box approach offers the possibility of providing a combined physical and statistical description of the system. The identified models are useful in the view of controlling refrigerators as flexible consumption units, which operation can be shifted within temperature and operational constraints. Even if the refrigerators are not intended to be used as smart loads, validated models are useful in predicting units consumption. This information can increase the optimality of the management of other flexible units, such as heat pumps for space heating, in order to smooth the load factor during peak hours, enhance reliability and efficiency in power networks and reduce operational costs. Household refrigerator modeling and performance assessment has been previously addressed with such approaches as dynamic simulation [4], steady state simulation [5], or CFD models [6].

The motivation to this study is to provide simple, ready-to-use and validated lumped parameter (stochastic state space) models for household refrigerators. The approach used is formed by forward model selection and validation based on experimental data and statistical testing. The software used is CTSM [7], which is based on maximum likelihood estimation. Parameters as thermal masses, evaporator thermal resistance, U-value of insulation and refrigeration cycle Coefficient of Performance (COP) are identified for each model in terms of expected value and variance. Convergence of estimation is also troubleshot.

II. Experimental setup

The experimental setup consists of: household refrigerator of capacity 60 liters with freezer bay and single compressor, power meter DEIF-MIC2, ADAM-6024 ADC card, four calibrated temperature sensors TI-LM35, one remotely controlled power outlet. Every second the refrigerator internal temperatures, ambient temperatures and refrigerator active power consumption are synchronously measured. Given the stratification of temperatures in the refrigeration chamber, two sensors are used in order to provide the average internal temperature. The same approach is used for determining the ambient temperature.

The refrigerator thermostat is set to supply the minimum temperature such that it is possible, within a temperature range, to enable or disable the compressor operation directly via the controlled power outlet.

III. Model of refrigeration cycle

This section presents a simple model for vapor-compression refrigeration system based on steady state one-

dimensional heat transfer equations. It is deemed valid to develop a static model of the vapor-compression cycle since it has faster dynamics if compared to the cold storage. Figure 1 shows a schematic representation of common refrigeration system for household applications.

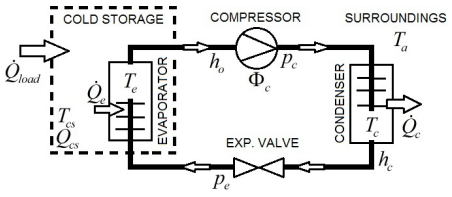


Fig. 1. Single stage vapor-compression refrigeration system.

A simple model of the system is:

$$\frac{dQ_{cs}(t)}{dt} = \dot{Q}_{load}(t) - \dot{Q}_e(t)\,, \tag{1}$$

where:

$$\begin{aligned} &dQ_{cs} = m_{cs} c_{cs} dT_{cs} \\ &\dot{Q}_{load} = UA_{cs}\left(T_a - T_{cs}\right) \\ &\dot{Q}_e = \dot{m}_r\left[h_o\left(p_e\right) - h_c\left(p_c\right)\right] \approx COP \cdot \Phi_c \\ &\dot{m}_r = N_c \alpha \rho_r\left(p_e\right) \end{aligned} \tag{2}$$

In Eq.2, m_{cs} is the cold storage mass and c_{cs} is its specific heat capacity. h_o and h_c are the evaporation and condensation enthalpies at the evaporation and condensation pressures, respectively p_e and p_c. UA_{cs} is the overall transmittance coefficient from the refrigeration chamber to the ambient and $\dot{m}_r$ is the refrigerant mass flow rate. COP is the overall coefficient of performance, here defined as the ratio between $\dot{Q}_e$, the thermal power extracted at evaporator side, and Φ_c, the refrigerator electrical consumption.

IV. Grey box modeling

Grey-box modeling is a framework for identifying a system description that combines prior physical knowledge of the system with information obtained from experimental data. For parameters estimation and system control it is convenient to use stochastic state space models [8], where the dynamical part of the model, the state, is described by Stochastic Differential Equations (SDEs) and the output is given by a discrete time algebraic equation describing how the observations are linked to the state. The parameters estimation and uncertainty assessment is obtained with statistical methods [9]. A stochastic differential equation (SDE) is a differential equation where one or more terms are stochastic processes, so that the solution is a stochastic process itself.

This section presents three different models of increasing complexity, all of which are developed under the hypotheses of: homogeneous materials, linear cooling cycle with constant COP and neglect of freezer compartment.

It is convenient to use electric thermal equivalent models in order to easily depict the models' structure and relate the identified parameters to physical quantities such as thermal transmittances and efficiency coefficients.

A. Model T_i

Here the refrigeration chamber is represented with a thermal mass, C_i, while the envelope (insulation) is modeled with a pure thermal resistance, R_{ia} (Fig. 2):

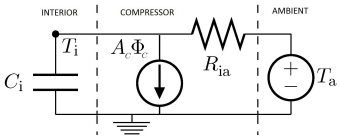


Fig. 2. Refrigerator preliminary model (electrical equivalent): T_i.

The compressor has a direct refrigeration effect, so that it is modeled as a current generator. This model is a single state stochastic state space model:

$$\begin{aligned} dT_i &= \left[\tfrac{1}{C_i R_{ia}}\left(T_a - T_i\right) - \tfrac{1}{C_i} A_c \Phi_c\right] dt + \sigma_1 dw \\ y_{t_k} &= T_{i,t_k} + e_{t_k}, \quad e_{t_k} \sim N(0, \sigma_e^2) \end{aligned}\,, \tag{3}$$

where A_c is the cycle COP and w is a standard Wiener process independent from the residual e_{t_k}. T_a is the ambient temperature, T_i is the refrigeration chamber temperature and Φ_c is the compressor active power consumption. Parameters' units are:

$$T_i = [^oC]\,,\; R_i = \left[\tfrac{^oC}{W}\right],\; C = \left[\tfrac{J}{K}\right],\; A_c = [scalar]\,,\; \Phi_c = \left[\tfrac{kJ}{s}\right].$$

B. Model $T_i T_{evap}$

This model extends the previous one by accounting for the heat transfer between the refrigeration chamber and the evaporator. This leads to an additional state for the evaporator temperature, T_e:

$$\begin{aligned} dT_i &= \left[\tfrac{1}{C_i R_{ia}}\left(T_a - T_i\right) + \tfrac{1}{C_i R_{ei}}\left(T_e - T_i\right)\right] dt + \sigma_1 dw_1 \\ dT_e &= \left[\tfrac{1}{C_{evap} R_{ei}}\left(T_i - T_e\right) - \tfrac{1}{C_{evap}} A_c \Phi_c\right] dt + \sigma_2 dw_2 \\ y_{t_k} &= T_{i,t_k} + e_{t_k}, \quad e_{t_k} \sim N(0, \sigma_e^2) \end{aligned}\,, \tag{4}$$

where w_1, w_2 and e_{t_k} are independent.

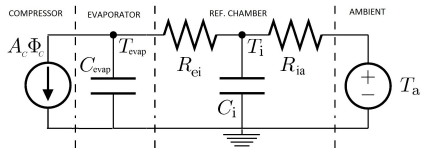


Fig. 3. Refrigerator model (electrical equivalent): $T_i T_{evap}$.

C. Model $T_i T_{evap} T_e$

Here the $T_i T_e$ model is extended by adding a state to the envelope and separating the envelope thermal resistance in inner resistance, R_{ie}, and outer resistance, R_{ea}:

$$\begin{aligned} dT_{evap} &= \left[\tfrac{1}{C_{evap} R_{evi}}\left(T_i - T_{evap}\right) - \tfrac{1}{C_{evap}} A_c \Phi_c\right] dt + \sigma_1 dw_1 \\ dT_i &= \left[\tfrac{1}{C_i R_{evi}}\left(T_{evap} - T_i\right) + \tfrac{1}{C_i R_{ie}}\left(T_e - T_i\right)\right] dt + \sigma_2 dw_2 \\ dT_e &= \left[\tfrac{1}{C_e R_{ie}}\left(T_i - T_e\right) + \tfrac{1}{C_e R_{ea}}\left(T_a - T_e\right)\right] dt + \sigma_3 dw_3 \\ y_{t_k} &= T_{i,t_k} + e_{t_k}, \quad e_{t_k} \sim N(0, \sigma_e^2) \end{aligned}\,, \tag{5}$$

where w_1, w_2, w_3 and e_{t_k} are independent. Follows the electric equivalent model:

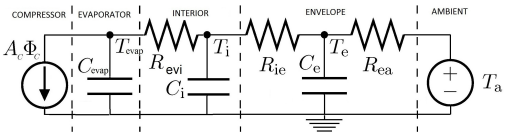


Fig. 4. Refrigerator model (electrical equivalent): $T_i T_{evap} T_e$.

V. A-PRIORI PARAMETERS

Grey-box modeling can benefit from calculated or judged value of parameters to be used as initial value for the estimation process. This section presents an initial estimation of physical parameters for the refrigeration chamber, including the glass shelves and the plastic drawer. The refrigerator insulation is assumed to be made by extruded expanded polystyrene (XPS).

A. Refrigeration chamber (thermal mass)

a) Air (0°C, sea level, dry air):

$$\begin{aligned} &c_{v-air} = 1297 \tfrac{J}{m^3 K}, \quad V_{air} = 0.111456\, m^3 \\ &C_{air} = c_{v-air} V_{air} \simeq 145 \tfrac{J}{K}. \end{aligned} \tag{6}$$

b) Glass (tempered glass):

$$\begin{aligned} &V_{shelf(1,2)} = 8.25 \cdot 10^{-4} m^3,\ V_{shelf(3)} = 4.41 \cdot 10^{-4} m^3 \\ &\rho_{glass} = 2500 \tfrac{kg}{m^3},\ c_{m-glass} = 0.84 \tfrac{J}{gK} \\ &m_{glass} = \rho_{glass} \left(2 \cdot V_{shelf(1,2)} + V_{shelf(3)}\right) = 5.232\, kg \\ &C_{glass} = c_{m-glass} m_{glass} \simeq 4395 \tfrac{J}{K}. \end{aligned} \tag{7}$$

c) Plastic (a rough estimation for the drawer):

$$\begin{aligned} &\rho_{polyetylene} = 910 \tfrac{kg}{m^3},\ V_{drawer} \simeq 7.096 \cdot 10^{-4} m^3 \\ &m_{drawer} = \rho_{polyetylene} V_{drawer} \simeq 0.65\, kg \\ &c_{m-polyetylene} = 1.67 \tfrac{J}{g\,K} \\ &C_{plastic} = m_{drawer} c_{polyetylene} \simeq 1086 \tfrac{J}{K}. \end{aligned} \tag{8}$$

d) Total thermal mass of refrigeration chamber:

$$C_i = C_{air} + C_{glass} + C_{plastic} = 5626 \frac{J}{K} \tag{9}$$

B. Envelope: thermal mass and resistance

It is reasonable to assume that the insulation layer has size: 44cm depth (D), 55cm height (H), 48cm width (L), and 3.5cm thickness (d):

$$\begin{aligned} &\rho_{poly} = 50 \tfrac{kg}{m^3},\ c_{m-poly} = 1.3 \tfrac{J}{g\,K},\ \lambda_{poly} = 0.033 \tfrac{W}{mK} \\ &S_{envelope} = 1.4344 m^2,\ V_{envelope} = d \cdot S_{envelope} \simeq 0.043\, m^3 \\ &m_{envelope} = \rho_{poly} V_{envelope} = 2.15\, kg \end{aligned}$$

e) Total thermal mass and resistance of the envelope:

$$\begin{aligned} &C_e = m_{envelope} c_{m-poly} = 2797 \tfrac{J}{K} \\ &R_e = (\tfrac{1}{\lambda_{poly}} \cdot \tfrac{d}{S}) \simeq 0.74 \tfrac{K}{W} \end{aligned} \tag{10}$$

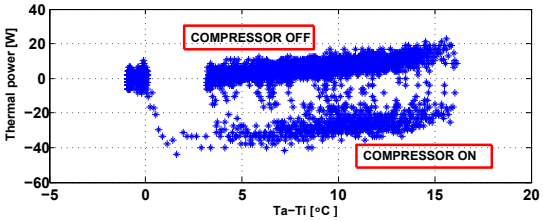


Fig. 5. Refrigerator operation: thermal power at refrigeration chamber v.s. temperature drop.

C. Refrigeration cycle (COP)

Figure 5 shows the total thermal power acting on the refrigeration chamber versus the temperature drop. When the compressor is not operating, the thermal power coming from the ambient accounts for approximately 8W, whereas during the refrigeration cycle the total thermal power at the refrigeration chamber is approximately -30W. Therefore the compressor generates approximately -38 thermal watts with an average electrical consumption of 50 watts, so that an initial value of the COP is:

$$COP \simeq 0.76. \tag{11}$$

The COP could seem low, but notice that here it is approximated by the ratio between thermal power extracted from the refrigeration chamber and the electrical power consumed by the compressor and hence it includes also the mechanical and electrical efficiency.

VI. SYSTEM IDENTIFICATION

Parameter estimation is carried out using CTSM, which provides a tool for semi-physical modeling and identification of dynamic systems based on stochastic differential equations [10]. CTSM provides methods for estimating unknown parameters of the model from data, including parameters in the diffusion term, using either the maximum likelihood (ML) method [11] or the maximum a posteriori (MAP) method. Both methods allow for several independent data sets to be used and are both sound statistically based estimation methods, which means that once the parameters have been estimated, statistical methods can be applied to investigate the quality of the model [12].

Figure 6 shows the process of model identification and validation. A first set of data, called *identification data* (see Fig. 7), is used for estimating model parameters and initial values of the states. Then the model, using the power input and room temperature from the same identification data, is used to calculate the one-step ahead predictions of the output. These predictions are subtracted from the measured output to form the residuals, which are analyzed for their white noise properties. If the model prediction residual is statistically *close* to white noise, the model is good [9]; therefore the auto correlation function is used to analyze the residuals (see, eg., Fig. 9). This procedure is called *model validation*.

A model can also be validated with another data set (see, eg., Fig. 14). If the results are good, this procedure gives a good indication of model robustness and correct identification.

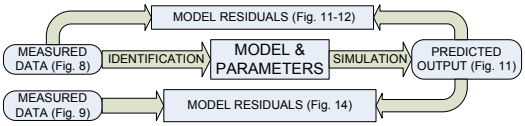


Fig. 6. Process of parameters identification and model validation.

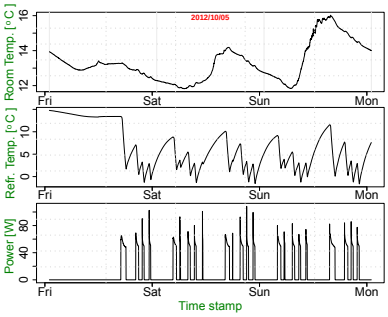


Fig. 7. Refrigerator operation: ambient temperature, internal temperature and electrical power consumption - identification data set.

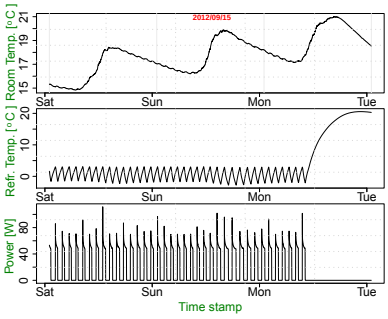


Fig. 8. Refrigerator operation: ambient temperature, internal temperature and electrical power consumption - validation data set.

A. *Parameters of the T_i model*

TABLE I. IDENTIFIED PARAMETERS: T_i MODEL

PARAMETER	VALUE	STD. DEV.
R_{ia}	1.4749	2.5617
C_i	$8.9374 \cdot 10^3$	$1.5481 \cdot 10^4$
$T_i(0)$	14.774	$2.9795 \cdot 10^{-2}$
A_c	0.58092	1.0075
$exp(\sigma_1)$	-5.4552	$1.2511 \cdot 10^{-2}$
$exp(e)$	-24.332	75.437
$Loglikelihood$	7995.168	

Figure 9 shows the residuals analysis of the model described by (3) with respect the identification data set. The first graph on the left presents the auto correlation function (ACF) of residuals, the graph in the middle the periodogram and the graph on the right the cumulated periodogram. High correlation of residuals at low values of lags indicates that the dynamics are not well modeled, and hence it is concluded that the model is too simple to describe the dynamics.

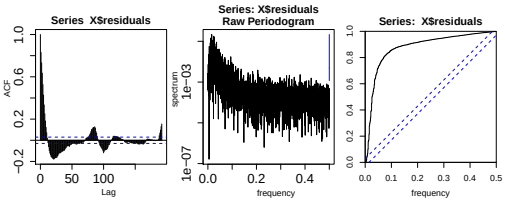


Fig. 9. Model residuals analysis: T_i.

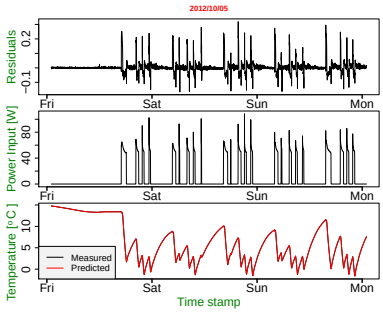


Fig. 10. Residuals, input and output of Model: T_i.

Figure 10 presents the residuals (top chart), the power input (mid chart) and the predicted and measured temperature in the refrigerator. From these plots it is possible to depict that model residuals are higher at the beginning of refrigeration cycle. Such situation was expected, since the non-linearities and complexity of the refrigeration cycle are not considered in the model. When the compressor is off, the prediction error is low and residuals are similar to white noise. Due to its inaccuracies and the identified missing dynamics from Fig. 9, this model is not further validated. In the next subsections, the second group of charts (eg., Fig. 10) is omitted for brevity.

B. *Parameters of the T_iT_{evap} model*

TABLE II. IDENTIFIED PARAMETERS: T_iT_{evap} MODEL.

PARAMETER	VALUE	STD. DEV.
R_{ia}	$9.0188 \cdot 10^{-1}$	$3.5460 \cdot 10^{-2}$
R_{ei}	$9.0348 \cdot 10^{-1}$	$2.5121 \cdot 10^{-1}$
C_i	$1.1600 \cdot 10^4$	$1.6529 \cdot 10^2$
C_e	$3.4342 \cdot 10^2$	$9.9157 \cdot 10^1$
$T_i(0)$	14.774	$1.0263 \cdot 10^{-2}$
$T_{evap}(0)$	16.181	3.6991
A_c	0.8383	$2.6217 \cdot 10^{-2}$
$exp(\sigma_1)$	$-1.7406 \cdot 10^1$	$5.6451 \cdot 10^{-2}$
$exp(\sigma_2)$	$-8.9551 \cdot 10^{-1}$	$2.6646 \cdot 10^{-1}$
$exp(e)$	$-1.2246 \cdot 10^1$	$1.1364 \cdot 10^{-1}$
$Loglikelihood$	12096.4351	

The residual analysis in Fig. 11 shows a clear improvement of model T_iT_{evap} compared to T_i and the cumulative periodogram is almost inside the confidence bands.

C. *Parameters of $T_iT_{evap}T_e$ model*

Model $T_iT_{evap}T_e$ outperforms in data fitting and the cumulative periodogram stays in the confidence bands. Figures 13

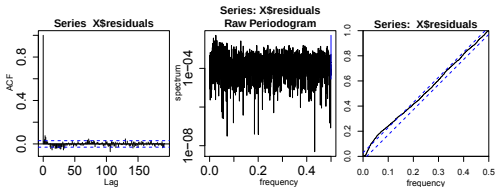


Fig. 11. Model residuals analysis: T_iT_{evap}.

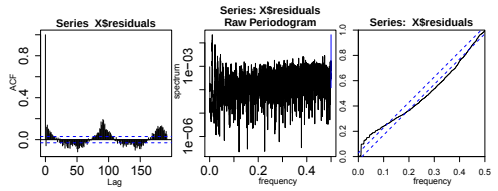


Fig. 12. Model residuals analysis: T_iT_{evap} - validation data set.

TABLE III. IDENTIFIED PARAMETERS: $T_iT_{evap}T_e$ MODEL

PARAMETER	VALUE	STD. DEV.
R_{ea}	$7.2869 \cdot 10^{-2}$	$1.8571 \cdot 10^{-2}$
R_{evi}	2.2431	$5.1971 \cdot 10^{-1}$
R_{ie}	3.7394	1.9380
C_i	$4.4245 \cdot 10^{3}$	$2.2810 \cdot 10^{3}$
C_e	$1.0755 \cdot 10^{4}$	$2.4514 \cdot 10^{3}$
C_{evap}	$1.9177 \cdot 10^{1}$	4.8643
$T_i(0)$	14.774	$8.6339 \cdot 10^{-3}$
$T_e(0)$	14.38	6.1042
$T_{evap}(0)$	18.568	5.5536
A_c	$2.1808 \cdot 10^{-1}$	$1.1258 \cdot 10^{-1}$
$exp(\sigma_1)$	$-1.7661 \cdot 10^{1}$	$1.2498 \cdot 10^{1}$
$exp(\sigma_2)$	$-2.0051 \cdot 10^{1}$	2.4919
$exp(\sigma_3)$	$-6.2477 \cdot 10^{-1}$	$1.0131 \cdot 10^{-1}$
$exp(e)$	$-1.1766 \cdot 10^{1}$	$7.9326 \cdot 10^{-2}$
$Loglikelihood$	12306.517	

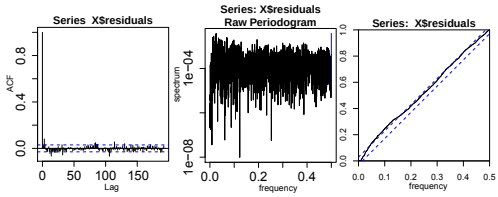


Fig. 13. Model residuals analysis: $T_iT_{evap}T_e$.

and 14 present the residuals analysis using respectively the identification data set and the validation data set. Hence it is concluded that this model seems capable of describing the observed dynamics of the refrigerator.

D. Model selection

Previous estimation trials have shown that model $T_iT_{evap}T_e$ leads to the highest likelihood value (12306) and best residuals properties. However, model T_iT_{evap} has good residuals properties and a high likelihood value (12096). Moreover, identified parameters of model T_iT_{evap} are closer to the prior estimates, compared to the parameters of $T_iT_{evap}T_e$ model, and using the validation data set it is found that T_iT_{evap} has the best performance. Therefore the choice of model T_iT_{evap} as reference model for the given setup.

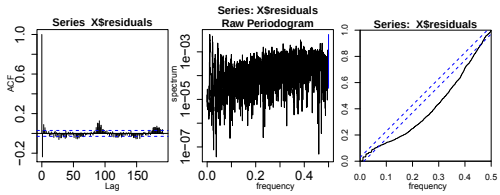


Fig. 14. Model residuals analysis: $T_iT_{evap}T_e$ - validation data set.

VII. CONCLUSION

This study showed an application of grey-box stochastic modeling for household refrigeration systems. Identified models are simple, reliable and, since they are SDE-based, they can be used for forecasting, control and simulation. Thanks to the diffusion terms, model uncertainties are also provided. This study represents for the authors a starting point for the development of intelligent control of such systems as thermal storages for providing power balancing services to the utility in a Smart Grid context.

ACKNOWLEDGMENT

The authors acknowledge the financial support of iPower, a project within the Danish Strategic Platform for Innovation and Research within Intelligent Electricity (www.iPower.dk).

REFERENCES

[1] World Business Council for Sustainable Development (WBCSD), "Transforming the market: Energy efficiency in the buildings," 2009.

[2] B. Flynn, "Key smart grid applications," *The Protection & Control Journal (G.E)*, Dec. 2008.

[3] G. Costanzo, A. Kosek, L. Ferrarini, G. Zhu, M. Anjos, and G. Savard, "An experimental study on load-peak shaving in smart homes by means of online admission control," in *IEEE Innovative Smart Grid Technologies Europe (ISGT)*, Oct. 2012.

[4] C. J. Hermes and C. Melo, "Assessment of the energy performance of household refrigerators via dynamic simulation," *Applied Thermal Engineering*, vol. 29, no. 56, pp. 1153 – 1165, 2009.

[5] C. J. Hermes, C. Melo, F. T. Knabben, and J. M. Gonalves, "Prediction of the energy consumption of household refrigerators and freezers via steady-state simulation," *Applied Energy*, vol. 86, no. 78, pp. 1311 – 1319, 2009.

[6] J. Gupta, M. R. Gopal, and S. Chakraborty, "Modeling of a domestic frost-free refrigerator," *International Journal of Refrigeration*, vol. 30, no. 2, pp. 311 – 322, 2007.

[7] Danish Technical University - Department of Informatics and Mathematical Modeling, "Continuous time stochastic modeling," 2005.

[8] J. Møller, N. Carstensen, and H. Madsen, *Stochastic State Space Modelling of Nonlinear systems - With application to Marine Ecosystems*, ser. IMM-PHD-2010. Technical University of Denmark (DTU), 2011.

[9] H. Madsen, *Time Series Analysis*, ser. Texts in statistical science. Chapman & Hall/CRC, 2008.

[10] N. R. Kristensen and H. Madsen, "Continuous time stochastic modelling, CTSM 2.3 - mathematics guide," DTU, Tech. Rep., 2003.

[11] H. Madsen and P. Thyregod, *Introduction to General and Generalized Linear Models*. CRC Press, 2010.

[12] N. R. Kristensen, H. Madsen, and S. B. Jørgensen, "Parameter estimation in stochastic grey-box models," *Automatica*, vol. 40, no. 2, pp. 225 – 237, 2004.

A Coordination Scheme for Distributed Model Predictive Control: Integration of Flexible DERs

Giuseppe Tommaso Costanzo,
Oliver Gehrke, Daniel Esteban Morales Bondy,
Fabrizio Sossan and Henrik Bindner
Center for Electric Power and Energy
Technical University of Denmark, Risø Campus
4000 Roskilde (DK)
{guco, olge, bondy, faso, hwbi}@elektro.dtu.dk

Jacopo Parvizi and Henrik Madsen
DTU Compute
Technical University of Denmark
2800 Lyngby (DK)
{jacop, hmad}@dtu.dk

***Abstract*—This paper presents a control scheme based on distributed model predictive control (DMPC) for coordinating flexible distributed energy resources (DER) of heterogeneous type in the Smart Grid with minimum system integration effort. This approach can be used for reducing the peak power exchange between the grid and a cluster of units in the same feeder in price-driven demand response applications. Preliminary simulations prove that the proposed coordination scheme for DMPC succeeds in coordinating flexible DER unit, achieving significant peak shaving when required. The rationale of this approach consists in coordinating independent units equipped with local MPC controller via simple information passing and hiding in the local controllers the units' dynamics.**

***Keywords*—Model Predictive Control, Smart Buildings, Smart Grid, demand response, DER integration, distributed control.**

I. Introduction

In the last two decades, factors such as increased global energy demand, speculation on fossil fuels and global warming have generated a high interest in renewable energy sources. The increase in electricity production from distributed renewable energy resources (RES), creates the need for new ways of providing ancillary services. This is important in order to maintain safe and reliable operation of the power system, both at transmission and distribution levels. Electricity production stemming from wind and solar energy can help to significantly reduce the carbon footprint of human activities and, due to favorable legislation, the installation of environmentally friendly production units in low voltage (LV) feeders has increased in the past decade. However, if corrective actions are not taken, an increased share of distributed production from renewable resources can threaten the grid stability. Therefore the integration of DERs in the distribution grid has to be operated judiciously.

It is foreseen in the near future an increase of distributed storage, provided by electric vehicles (EVs), and flexible consumption enabled by demand-side management (DSM) in buildings. This technology leads to higher flexibility and predictability of energy demand, allowing such services as peak shaving, valley filling, load shifting and real-time grid congestion management in the Smart Grid. For an insight on DSM practice, we refer to [1], [2], [3] and the references therein.

This paper presents a coordination scheme for a control architecture based on DMPC for the aggregation and coordination of multiple DER units that are connected behind a single Point of Common Coupling (PCC) in the LV grid. These units could be EVs, HVAC systems, electric water heaters and refrigerators, whose consumption is intrinsically flexible. The coordinated operation of these units allows the provision of aggregated services to the Distribution System Operator (DSO) such as peak shaving or voltage support [4], [5]. The presented approach may allow for the connection of several units with a potential high peak consumption on the same feeder while ensuring that existing flow constraints for lines, circuit breakers or transformers are respected. The main motivations for a distributed solution are computational scalability, modularity, and easy system integration.

This paper is organized as follows: Section II presents the case study and Section III the DERs models, whereas the proposed architecture for DMPC is presented in Section IV. Preliminary simulation results are presented in Section V.

II. Case Study

This case study presents the coordination of multiple flexible DERs that may be installed in the same distribution feeder. The goal of coordinating units is to constrain the aggregated consumption/production of a cluster (group of units) to a fixed value or to a specific schedule provided by a higher-level entity in the Smart Grid, such as an aggregator or the DSO. Such service is called *PowerMax* and it is presented in [4]. Although the case study presented in this work considers three types of unit: building space heating and air conditioning (climate control), photovoltaic array (PV) and electric energy storage (EVs and local batteries), this coordination scheme is also suitable for other types of unit and allows units to join and leave the cluster dynamically. Figure 1 shows the configuration for this case study.

Each unit is equipped with a local model predictive controller (MPC), which optimizes the unit operation with respect to a local objective and it complies with local and global constraints. Controllers' local objective can range from minimizing operating costs to maximizing comfort, while local constraints ensure minimal comfort requirements and safe unit operation. The units in the cluster have also to respect a global constraint which relates to the power flow at the point of

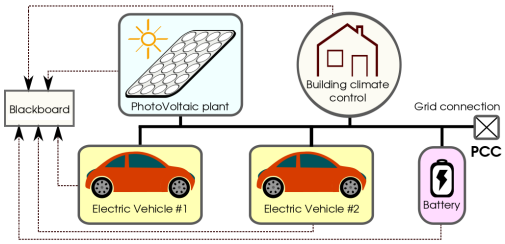


Fig. 1. Case study setup.

connection between the cluster and the main grid. Each unit, by retrieving information about other units consumption and the PCC power limit from the cluster blackboard, computes the local power plan by solving a constrained optimization problem. The next section presents, for each unit type, the model and the operational requirements (local constraints).

III. Models of DERs

1) Building and space heating: the model for building heating demand is taken from [6] (model $T_iT_eT_h$), where a single room-equivalent model is used for an office building of eight rooms equipped with electric space heating and air conditioning units. Model 1, which has three states, represents the dynamics of the (equivalent) room internal temperature, the (equivalent) heater temperature and the envelope (building insulation) temperature:

$$dT_i = \frac{1}{C_i}\left[\frac{1}{R_{ie}}\left(T_e - T_i\right)dt + \frac{1}{R_{ih}}\left(T_{hc} - T_i\right)dt + A_w\Phi_s dt\right] + \sigma_i dw_i \tag{1a}$$

$$dT_e = \frac{1}{C_e}\left[\frac{1}{R_{ie}}\left(T_i - T_e\right)dt + \frac{1}{R_{ea}}\left(T_a - T_e\right)dt\right] + \sigma_e dw_e \tag{1b}$$

$$dT_{hc} = \frac{1}{C_h}\left[\frac{1}{R_{ih}}\left(T_i - T_{hc}\right)dt + \Phi_{in}dt\right] + \sigma_h dw_h \tag{1c}$$

where Φ_{in} and Φ_s are respectively the electric power consumed for climate control and the solar irradiation. Although Model 1 does not include the air conditioning units, in this study, as a simplifying assumption, the air conditioner dynamics is assumed equivalent to the heater dynamics. The fact that the heater dynamics is about two orders of magnitude faster than the room dynamics justifies the choice of assuming the AC unit *"fast enough"*, as the heater is, with respect to the room heat dynamics. Therefore the cooling effect of air conditioning is accounted with a negative sign on Φ_{in} and a multiplicative factor of 3 (average COP from AC units datasheet). Finally T_i, T_e and T_{hc} are, respectively, the room internal temperature, the envelope temperature and the heating/cooling module temperature.

2) Batteries and EVs: in this work the EVs are modeled as batteries and driving patterns are considered and simulated as disturbances to the MPC. A detailed model of electric storage units is presented in [7], from which it emerges that, for lifetime purposes, the batteries should be operated between 20% and 90% of the state of charge (SOC). Moreover, when operating in this band, the batteries can be charged and discharged at constant current and can be reasonably modeled as integrators of rated efficiency. The support batteries differ from the EVs by capacity and rated power, therefore a generic state space model for this type of units is presented below:

$$dx = A_s x\ dt + B_s u\ dt \tag{2a}$$

$$SOC = C_s x \tag{2b}$$

where u is the power flow at the inverter in kW, x is the battery state of charge in kWh and SOC is the battery state of charge normalized between 0 and 1 and it is related to the battery rated capacity.

3) PV array: this unit is supposed to operate at maximum power point (MPP) and the unit local controller, using the forecast of solar radiance, wind speed and external temperature, publishes in the blackboard the expected power production using the model presented in [8], [9]. Every time the weather forecast is updated the PV controller accordingly updates the PV production plan in the blackboard. In this study, as simplification, the PV unit publishes a production plan which consists in the actual produced power from a real PV plant with added a normally distributed noise of 10% of the power measured at each instant. Then, at simulation time, the effective power produced by the unit is given by the real measurements decurted of the noise. This approach is considered in order to simulate inaccuracies on PV modeling and weather forecasts.

IV. Control Architecture

The computational effort for solving MPC problems generally grows at a superlinear rate with the number of state variables involved. The exact order is problem specific and depends on the coupling between the state variables as well as the chosen solving method. Managing multiple DERs with a single, feeder-wide MPC controller would not only require a large amount of DER state information to be shared with the central controller. The computational effort of such a setup would also scale badly for larger numbers of units. An established approach for solving the curse of dimensionality relies on the decomposition of the MPC problem into smaller subproblems which can be solved independently and locally. Convergence towards the overall goal is then achieved through a coordination mechanism, i.e. by communication between the individual solvers. An overview and categorization of distributed MPC variants is given in [10], [11]. In the presented approach, the MPC computations are carried out at unit level by the local unit controller. In order to meet the system constraint of limited power flow at the PCC, controllers have to coordinate because the fulfillment of the system constraint may not be derived only by measuring physical variables observable by the individual units.

Coordination between autonomous entities in a distributed system has been an active area of research for many years [12]. Many different ways of achieving such coordination have been proposed; however, as with many engineering problems, no one-size-fits-all solution exists and the specific trade-offs of a particular method have to be weighed against the needs of the distributed system in question. These needs may include scalability of the solution, synchronization speed, access concurrency, fault tolerance, level of trust between entities, convergence speed and others. The communication effort of the distributed MPC system presented in this paper needs to be scalable to a certain degree, but the speed of synchronization is not critical. A common approach for this

type of system is the use of a tuple space for information exchange, typically implemented as a blackboard [13] , i.e. a central data repository whose content all processes can retrieve and to which all processes can write. For this type of solution, the communication effort – as defined by the message count – for each iteration scales linearly with the number of units. A notable drawback is that a single point of failure is introduced at the blackboard itself which could be partially mitigated by adding a replication and failover mechanism for the blackboard.

A. Algorithm for MPCs coordination

In the presented scheme all units access the blackboard asynchronously, retrieve the power plans of the other units, compute their own plan accordingly and publish it in the blackboard. Units have exclusive access to the blackboard, which is a resource that is blocked by one unit at the time during the processes of data retrieving, power plan computing and data publishing. As soon as the available power at the PCC decreases, units tend to reduce their consumption (or production). However, if the operating or comfort requirements of a unit cause a violation on the constraint on the PCC power flow, a *minimal* power plan that minimizes the exceeding power at the PCC, called *power over max*, is published in the blackboard and the consensus flag related to that plan is set to false. Units keep calling the blackboard if they need to update their operation plan until the consensus flag is set to true or the iterations of such process reach a predefined limit, L, which is set equal to all units and serves to avoid deadlocks. When those conditions hold, the units are entitled to operate according to the local power plan for the next time frame. Thanks to such mechanism the minimal operation and comfort requirements of units are always satisfied by the local controllers, at the price of occasional violations of the limits on the PCC power flow in the case of over stringent constraint. Algorithm 1 presents the coordination scheme for DMPC, while the following subsections present the unit local MPC formulation for the building climate control and control of electric storage.

B. Building controller

The building controller uses a discretized version of Model 1 ($\Delta t = 30\,min$) and keeps the internal temperature at time t, $T_{i,t}$, within the comfort bounds $[T_{\min,t}, T_{\max,t}]$ and trades off energy expenses with user comfort by solving the following optimization problem:

$$\min_{\Phi_{h,t},\Phi_{c,t},\vartheta_t} J = \sum_{t=1}^{N} \left[\|T_{i,t} - R_t\|_Q^2 + p_t \left(\Phi_{h,t} + \Phi_{c,t}\right) + \rho\vartheta_t \right] \quad (3a)$$

$$subject\ to:$$

$$x_{t+1} = A_b x_t + B_b \Phi_{h,t} + F_b \Phi_{c,t} + E_b T_{a,t} + S_b \Phi_{s,t} \quad (3b)$$

$$T_{i,t} = C x_t \quad (3c)$$

$$0 \le \Phi_{h,t} \le U_{\max,t} \quad (3d)$$

$$0 \le \Phi_{c,t} \le C_{\max,t} \quad (3e)$$

$$T_{\min,t} \le T_{i,t} \le T_{\max,t} \quad (3f)$$

$$\vartheta_t \ge 0 \quad (3g)$$

$$PCC_{\min,t} - \vartheta_t \le \Psi_t + \Phi_{h,t} + \Phi_{c,t} \le PCC_{\max,t} + \vartheta_t \quad (3h)$$

The objective function in Eq. 3a is composed by three terms. The first term penalizes the deviation of $T_{i,t}$ from the setpoint, R_t; the energy expense is minimized in the second term, where p_t is the energy price over the prediction horizon N taken from the Nordpool Elspot market; finally *power over max*, ϑ_t, is minimized in the third term with the parameter ρ.

Algorithm 1 Coordination of MPCs

Variables:

- N: prediction horizon
- Ψ: aggregated units operation plan
- $P_{\max}, P_{\min}$: bounds on contracted power at PCC
- U_j: operation plan of unit j, $U_j \triangleq \{u_{j,t}\}$
- L: iteration limit for consensus

Require: Initialize $U_j = 0,\ \forall j$
  consensus = FALSE
  i = 0
  while (not(consensus) OR i $\le$ L) **do**
    for all j **do**
      $\Psi = \sum_k U_k\ , k \ne j$
      U_j = compute MPC$_j$(state$_j$, Ψ)
      if $P_{\min}(t) < \Psi(t) + U_j(t) < P_{\max}(t)\ ,\ \forall t$ **then**
        consensus = TRUE
      else
        consensus = FALSE
      end if
      push U_j to the blackboard
    end for
    i = i + 1
  end while
  dispatch $u_{j,1}, \forall j$

The MPC operates the heater and the air conditioner independently, so that the control input Φ_{in} in Eq. 1c is divided in: $\Phi_{h,t}$ for heating and $\Phi_{c,t}$ for cooling. In this way the discretized model in Eq. 3b presents two control inputs ($\Phi_{h,t}$ and $\Phi_{c,t}$) and two disturbances ($T_{a,t}$ and $\Phi_{s,t}$). Equations 3b and 3c represent the building thermal dynamics for the MPC, Eq. 3d and 3e set the bounds on control inputs, whereas the comfort bounds for the building internal temperature are stated in Eq. 3f. Finally, violations of limits on the PCC power flow are allowed by the soft constraints in Eq.3g and 3h. The optimization problem defined in Eq. 3 is reformulated as a least squares problem, which is solved via quadratic programming algorithm as presented in [14].

C. Battery and EV controller

The controller for the electric storage strives at minimizing the energy cost while keeping the SOC within given bounds. The MPC problem is formulated as follows:

$$\min\ J = \sum_{t=1}^{N} [p_t u_t + \rho\vartheta_t] \quad (4a)$$

$$subject\ to:$$

$$x_{t+1} = A_s x_t + B_s u_t \quad (4b)$$

$$SOC_t = C_s x_t \quad (4c)$$

$$U_{\min,t} \le u_t \le U_{\max,t} \quad (4d)$$

$$SOC_{\min,t} \le SOC_t \le SOC_{\max,t} \quad (4e)$$

$$PCC_{\min,t} - \vartheta_t \le \Psi_t + u_t \le PCC_{\max,t} + \vartheta_t \quad (4f)$$

The cost function is expressed in Eq. 4a, which consists in a term for minimizing the energy cost (p_t is the Nordpool Elspot price) and a term, ρ, for minimizing the *power over max*. Constraints in Eq. 4b, 4c implement the energy

storage dynamics, while Eq. 4d accounts for inverter power capabilities. Constraints in Eq. 4e allow the MPC controller to keep the battery state of charge between given operational bounds. This constraint can be used to bring the SOC to a convenient level, accounting for the users driving patterns and the unit's availability. If the controlled unit is a battery, the SOC is not time dependent. Finally, Eq. 4f penalizes the excess of power flow at the PCC. A similar MPC formulation for vehicle charging is presented in [15].

In the control scheme proposed in this work no computation is carried out at the data repository level, which can be as simple as an html page. In this way the devices optimization is entirely distributed to a local level. Using a blackboard with asynchronous units access and negotiation allows to keep both system and information exchange simple; moreover there is no need for hierarchical coordination, since this is achieved by means of coupling constraints in the local MPCs. All of these factors enhance the system interoperability, allowing to integrate and coordinate devices of heterogeneous type and with totally different operation requirements and constraints. Ultimately, if a data repository is not available, each controller can broadcast its plan to the other controllers. In this case the number of messages exchanged considerably grows with the number of units per group.

V. Simulation Study

The simulation study shows the effectiveness of the proposed algorithm in limiting the aggregated power flow of a cluster of units composed by 100 buildings, 10 local batteries, 20 large EVs, 20 small EVs and 20 PV modules. This population of units is created by using the models from literature with parameters that are normally distributed with a 10% variance around the nominal value. Figure 2 presents the aggregated power flow at the PCC without units coordination (black line, with peaks up to +420kW and -320kW) and with units coordination (green line). It is shown that the proposed algorithm successfully limits the aggregated power flow between the given bounds of $\pm 200\ kW$.

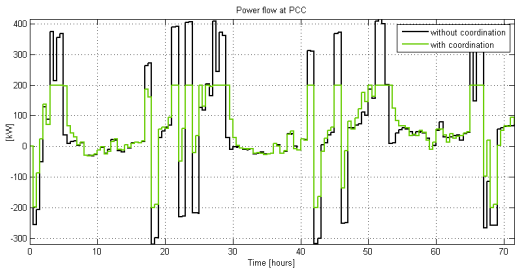


Fig. 2. Power flow at the PCC: units coordination v.s. units non-coordination.

Note that the coordination scheme has no effect when the units' aggregated consumption is *spontaneously* below the given limit of $\pm 200\ kW$, eg. the green and black line are superposed in the time frame between 30h and 40h (Fig. 2). Figure 3 shows the power exceeding the PCC limit versus the number of coordination iterations, for different maximum power requirements. It emerges that the algorithm converges in two iterations if the available power at the PCC is sufficient

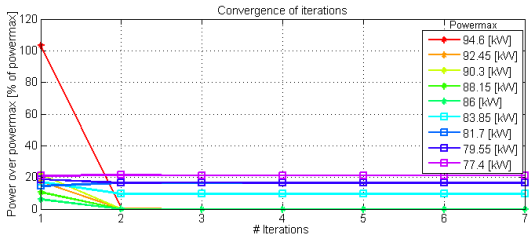


Fig. 3. study of algorithm convergence: exceeding power flow at the PCC v.s. number of iterations.

to satisfy the units minimum requirements. Conversely, if the available power is not sufficient, the units controllers keep trying to coordinate without success and the consensus is not reached. Such limit for the given setup is between 86 and 84 kilowatts. Focusing now on single units' operation, Fig. 4 and Fig. 5 show the range of all buildings' internal temperature and power consumption, respectively for the cases of units non-coordination and units coordination. By inspecting Fig. 4 and

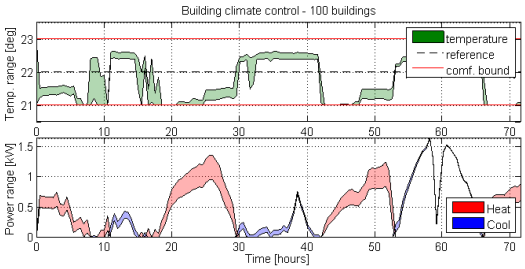


Fig. 4. Building thermal management without units coordination.

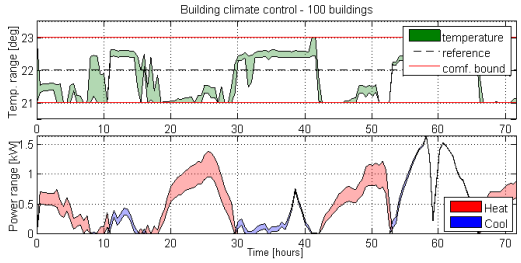


Fig. 5. Building thermal management with units coordination.

Fig. 5 it emerges that the buildings contribute in few occasions to shape the cluster consumption.

Figures 6 and 7 show the operation of the energy storages: batteries, small and big EVs. The range of units' SOC and power flow are shown for the cases of units non-coordination and coordination. It is observed in Fig. 6 that, when units are not coordinated, the control inputs lay in a narrow range. This behavior is due to the economic formulation of the MPCs: all units tend to charge when the price of energy is low and discharge when the price is high, causing peaks in the power flow at the PCC. When units are coordinated via DMPC, the trajectories of control inputs and SOC stay within a wider range, showing the units flexible operation (Fig. 7).

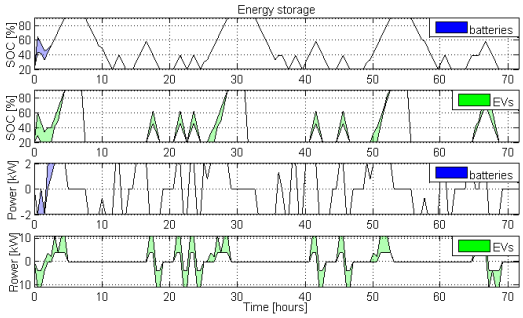


Fig. 6. Battery operation without units coordination: range of units' SOC and power flow.

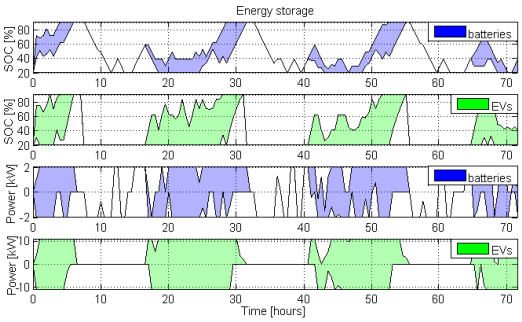


Fig. 7. Battery operation with units coordination: range of units' SOC and power flow.

Fig. 8 shows the predicted and realized PV production, and the NordPool Elspot price.

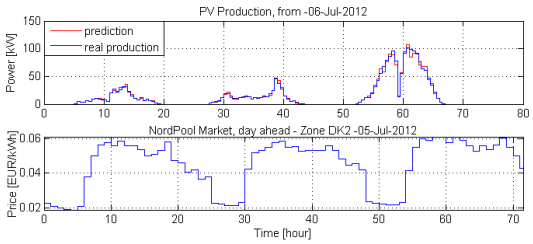


Fig. 8. Energy price and PV production.

The simulation study shows that the proposed architecture is capable of coordinating units in a cluster so as to shave the peak demand. However, although the algorithm convergence and optimality of the solution have not been formally proven, simulation trials show that given sufficient available power at the PCC to serve the units minimum requirements, the coordination scheme is successful and units flexibility is fully exploited. Moreover, it has been observed in simulation with smaller clusters that the speed of convergence does not depend by the cluster size.

VI. Conclusions

This paper presents an approach for the integration of flexible DERs in the Smart Grid based on DMPC coordination via coupling constraints. In the presented study, this control scheme allows minimizing the impact of PV plants and EVs in the distribution grid and the advantages brought from the blackboard-based coordination include scalability and easy system integration. It is planned as future study to investigate fairness of resource allocation among units, assess reliability problems in case of downtime of the blackboard and comparison with other DMPC algorithms.

Acknowledgments

The authors acknowledge the financial support of iPower, a project within the Danish Strategic Platform for Innovation and Research within Intelligent Electricity (www.ipower-net.dk).

References

[1] C. W. Gellings, "The concept of demand-side management for electric utilities," *Proceedings of the IEEE*, vol. 73, no. 10, pp. 1468–1470, 1985.

[2] L. Chan, N. Li, S. H. Low, and J. Doyle, "Two market models for demand response in power networks," in *IEEE SmartGridComm 2010*, Gaithersburg, MD, USA, October 2010, pp. 4569–4574.

[3] Z. Yua, F. Haghighata, B. C. Fungb, and H. Yoshinoc, "A decision tree method for building energy demand modeling," *Energy & Buildings*, vol. 42, no. 10, pp. 1637–1646, 2010.

[4] N. C. Nordentoft, Y. Ding, L. H. Hansen, P. D. Cajar, P. Brath, H. W. Bindner, and C. Zhang, "Development of a dso-market on flexibility services," iPower, Tech. Rep., 2013.

[5] O. Gehrke and F. Isleifsson, "An aggregation friendly information model for demand side resources," in *Local Computer Networks (LCN), 2010 IEEE 35th Conference on*, 2010, pp. 1019–1023.

[6] P. Bacher and H. Madsen, "Identifying suitable models for the heat dynamics of buildings," *Energy and Buildings*, vol. 43, no. 7, pp. 1511 – 1522, 2011.

[7] F. Marra, G. Y. Yang, C. Traholt, E. Larsen, C. Rasmussen, and S. You, "Demand profile study of battery electric vehicle under different charging options," in *Power and Energy Society General Meeting, 2012 IEEE*, 2012, pp. 1–7.

[8] W. De Soto, S. Klein, and W. Beckman, "Improvement and validation of a model for photovoltaic array performance," *SOLAR ENERGY*, vol. 80, no. 1, pp. 78–88, 2006.

[9] C. Koch-Ciobotaru, L. Mihet-Popa, F. Isleifsson, and H. Bindner, "Simulation model developed for a small-scale pv-system in a distribution network," in *Proceedings of the 8th International Symposium on Applied Computational Intelligence and Informatics-SACI 2012*, 2012, pp. 257–261.

[10] R. Negenborn, B. D. Schutter, and J. Hellendoorn, "Multi-agent model predictive control: A survey," Delft University of Technology, Tech. Rep. 04-010, August 2009.

[11] "Architectures for distributed and hierarchical model predictive control a review," *Journal of Process Control*, vol. 19, no. 5, pp. 723 – 731, 2009.

[12] T. W. Malone and K. Crowston, "The interdisciplinary study of coordination," *ACM Comput. Surv.*, vol. 26, no. 1, pp. 87–119, Mar. 1994. [Online]. Available: http://doi.acm.org/10.1145/174666.174668

[13] V. Jagannathan, R. Dodhiawala, and L. Baum, *Blackboard Architectures and Applications*, ser. Perspectives in Artificial Intelligence. Academic Press, 1989.

[14] R. Halvgaard, N. Poulsen, H. Madsen, and J. Jorgensen, "Economic model predictive control for building climate control in a smart grid," in *Innovative Smart Grid Technologies (ISGT), 2012 IEEE PES*, 2012, pp. 1–6.

[15] R. Halvgaard, N. Poulsen, H. Madsen, J. Jorgensen, F. Marra, and D. Bondy, "Electric vehicle charge planning using economic model predictive control," in *Electric Vehicle Conference (IEVC), 2012 IEEE International*, 2012, pp. 1–6.

Performance Assessment of Aggregation Control Services for Demand Response

Daniel Esteban Morales Bondy, Giuseppe Tommaso Costanzo, Kai Heussen, Henrik W. Bindner
Department of Electrical Engineering
Technical University of Denmark
Email: {bondy,guco,kh,hwbi}@elektro.dtu.dk

***Abstract*—Much effort has been put into creating aggregation algorithms that will provide services to the grid. In comparison, little attention has been given to performance assessment under the service delivery perspective. This paper formulates a performance index for aggregation control algorithms designed to provide ancillary services. A case study is presented where the index is applied to two controllers with same service objective. Through the index it is shown that centralized solution is able to provide a better Quality of Service to its costumers.**

Keywords*—Control Service, Distributed Ancillary Services, Performance Assessment, Aggregator, Demand Side Management.

I. Introduction

The future increase in energy production from Renewable Energy Sources (RES) may lead to a power system where production is distributed, and where the Transmission System Operators (TSOs) require a larger amount of balancing services. At the same time, the increase in Distributed Energy Resources (DERs) brings new challenges to the Distribution System Operators (DSOs), which may need new kind of ancillary services [1]. It is anticipated that DER owners will be able to provide services to the system operators via Demand Side Management (DSM).

An Aggregator is a market player whose business case is to manage DER units in its portfolio and use their inherent consumption flexibility to participate in one or more ancillary service markets, i.e. it provides the necessary coordination for DSM. A general classification of different aggregation methods is presented in [2]. An example of direct control can be found in [3]. An analysis and evaluation of indirect control architectures can be found in [4].

Since the Aggregator has contractual obligations with customers and system operators, it is important that the control algorithm the Aggregator uses proves suitable to task. From a service perspective, an aggregation algorithm is considered suitable if it performs according to the specified quality of service (QoS). In this case, a service is the control of a DER portfolio in such way that it fulfills the needs of both the DER owners and the System Operator. The QoS is the service performance that makes an acceptable service provision. A good QoS is essential to the functioning of the power grid.

Little attention has been given to the problem of performance assessment of aggregator controllers seen from a service delivery perspective. This paper approaches the problem by presenting two main points:

- both ancillary services and DSM have minimum QoS requirements that need to be respected. In this work we propose a way of evaluating the aggregation control algorithms according to specific requirements;
- a performance index suitable for evaluating, from the Aggregator point of view, the quality and reliability of aggregation control algorithms.

The paper is organized as follows: Section II gives a general description of concepts relevant to the definition of the index, while the index itself is defined in Section III. A case study is presented in Section III-C and further research is discussed in Section IV.

II. Method Background

A. Ancillary Services

Ancillary services are acquired by TSOs in order to ensure the stability of the system and can generally be divided into primary, secondary and tertiary ancillary services [5]. Each class of ancillary services has different purposes and work on different time scales.

In the Danish system, producers are allowed to bid into the ancillary services market once they have been approved by the TSO. In order to be approved, the producers must prove that they are able to deliver the relevant services within the specified requirements defined in [6]. The TSO defines in the quality of service (QoS) which are the acceptable errors in service delivery with respect to a reference power schedule.

Furthermore, it is expected that new ancillary services will appear in the near future, which will help the DSOs resolve congestion and voltage problems [1]. Throughout this paper, the recurring example of an ancillary service is the *PowerMax*, one of the new DSO services. This service is discussed further in Sec.III-C.

B. Asset Management Service

Since the flexibility of individual DERs is too small to provide services to the system operators, an Aggregator pools the flexibility of the units, and present their flexibility in the market as a single entity, see Fig.1. In this way the Aggregator provides ancillary services to the DSO or, through a Balance Responsible Party (BRP), the TSO. The Aggregator and the BRP could be the same entity, but if they are not, the Aggregator should not work against the balancing responsibilities of the BRP. The concept of an Aggregator operating a portfolio

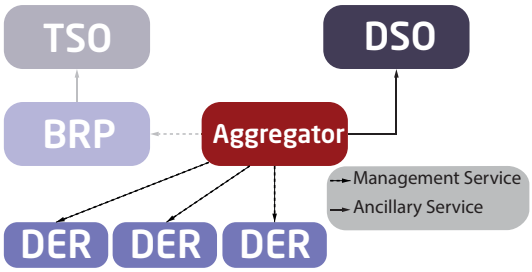


Fig. 1. The setup of the power system with DSM.

of DER units in order to deliver ancillary services is referred also as Demand Side Management (DSM) [7].

The objective of a DER is to satisfy the needs of its owner. The Aggregator provides the DER owners with an asset management service that ensures the QoS to customers is adequate. Providing services to the grid is only a secondary (and optional) function, which the Aggregator must take advantage of within the constraints of the primary function.

The power consumption of DERs varies greatly depending on the daily routines of their owners and the meteorological conditions. Due to the varying and distributed nature of the DERs, the Aggregator must be able to evaluate if its control algorithms are capable of providing DSM to the system operators and, at the same time, manage the primary functions of the DERs.

C. Control Performance Assessment

When talking about evaluation of controllers, there already exists a wide field of theory: Control Performance Assessment (CPA). The theory in this field is applied mostly in the process industry, and for a thorough overview of its application we refer to [8]. It seems natural to transfer the well studied concepts of CPA to the field of DSM.

From [9] we have that the measures used to evaluate a controller usually fall within three categories: Quality related criteria, Reliability related criteria, Energy efficiency related criteria.

The two first types of criteria can be directly related to service provision, for which the quality criterion is the performance of the asset management service towards the DER owner; and the reliability criterion, which is the performance of the service delivered to the DSO or Balance Responsible Party (BRP). Energy efficiency criteria are not considered in this work.

One approach to CPA is to establish the deterministic properties the closed-loop system must have, e.g. settling time and steady-state error [10] and summing up the criteria over time such that:

$$\eta = \frac{J_{opt}}{J_{act}} \tag{1}$$

where J_{opt} is the theoretical optimal value of the performance criterion(which is usually impossible to achieve in reality), J_{act} is the actual measured value of the criterion. Since $J_{opt} > J_{act}$, then $\eta \in [0, 1]$.

Given that the requirements of service delivery can easily be translated into time-domain deterministic measures, this work presents a deterministic approach. Using the concepts presented in this section, the performance index is defined in the next section.

III. The Performance Index

A. Defining the index

This section presents the main contribution of the paper: the formulation of a performance index that can be used to assess aggregation control algorithms designed to provide ancillary services. The index presented here is defined for post-simulation analysis, and represents the performance of the control algorithm over the whole time horizon.

Using the square root of the Integral Square Error index (i.e. the 2-norm, as defined in e.g. [11]), the following performance criterion is defined:

$$J = \sqrt{\int_0^N \left(\sum_{k=1}^{M} |e(t)_{D,k}|^2 Q_D + |e(t)_U|^2 Q_U \right) dt} \tag{2}$$

where $e(t)_D$ and $e(t)_U$ are, respectively, the time-dependent errors for the asset management service and the ancillary service. The units controlled by the Aggregator are denoted by the index k, the unit portfolio is of size M and N is the time horizon over which the services are provided. Finally, Q_D and Q_U are scaling factors that convert the errors into percentages so that $e(t)_D$ and $e(t)_U$ are comparable.

The actual measured performance criterion of the aggregation control algorithm is scaled with the maximum permissible error, J_{max}, such that the performance index is given by:

$$\eta = \frac{J_{act}}{J_{max}} \tag{3}$$

where $\eta \in [0, 1]$ and for which values close to zero represent good performance of service delivery. Contrary to index (1), which gives an intuition of how close performance is to the optimum, index (3) gives an intuition of how far performance is from the worst case scenario. The index was designed this way because the theoretical optimum of service delivery is $J_{opt} = 0$.

It must be noted that the performance criterion only measures the permissible error defined in the contract of the service, and service non-delivery is measured separately. Therefore, whenever the algorithm performs outside the established limits, then $J(t)_{act} = J(t)_{max}$ and the non-delivery counter is increased.

It is important to stress that the QoS has already been defined by the entity requesting the service, and this index is specifically designed to evaluate how close the service provision conforms to the defined QoS. Having defined what the performance index measures, we will proceed with establishing how to obtain the required values to estimate the index.

B. Calculating the index

Calculating the performance index requires the following steps:

1) Identify and model the service requirements and error in service provision.
2) Estimate J_{act}.
3) Calculate J for operation on the requirement boundaries ($J(t)_{max}$).
4) Calculate η by normalizing J_{act} with J_{max}.

For the first step, the service requirements must be defined and translated into measurable errors. For some services, the error can be stated as a tracking error, e.g. $e = y_{ref} - y_{meas}$. In other cases, service requirements are defined by operation within bands, which may lead to error defined as:

$$e(x) = \begin{cases} x_{min} - x & \text{if } x \leq x_{min} \\ 0 & \text{if } x_{min} \leq x \leq x_{max} \\ x - x_{max} & \text{if } x \geq x_{max} \end{cases} \quad (4)$$

This step is a service-specific problem and it is non trivial.

The second step requires to compute J_{act} using measurement data from the unit portfolio. This can be a challenge for evaluation in field deployment. In this paper it is assumed that the measurement data will be available, either through a DSO or a third party metering company.

The actual performance of the aggregation algorithm can be found through two different methods:

On-line monitoring – This method brings the added benefit of being able to use the index for performance monitoring and diagnosis at runtime, but the downside of being communication intensive.

Post-delivery analysis – This method is less communication intensive, but does not permit to take remedial actions at run time if a aggregation controller is not working as expected.

Usually services have some acceptable error (see Sec. II-A) which can be interpreted as the hard boundaries for the service delivery. The third step requires the calculation of J along these boundaries, in this way, the maximum allowed error is found for the service. The boundaries are based on the service models presented in the first step. By adding the maximum permissible error for all services, J_{max} is obtained. Normalizing the performance measure with the J_{max} gives an intuitive value of the performance of the control algorithm.

In the following subsection an example of how to determine J_{max} is presented.

C. An example: DSO Service PowerMax

For demonstration purpose, in this paragraph J_{max} for the PowerMax service is calculated.

Typically, the service will be contracted several months ahead the actual delivery. The activation schedule (On and Off triggers), the maximum power cap (P_M), the maximum duration of the service per activation (T_M), and the quality of service (*QoS*) are defined when contracting the service. The contract is valid for a period of several months, where the Aggregator is obliged to follow the established schedule.

The limits specified in the QoS [1] for the PowerMax service are presented here:

- Deviation from On trigger: $\pm$ 15 min. per day

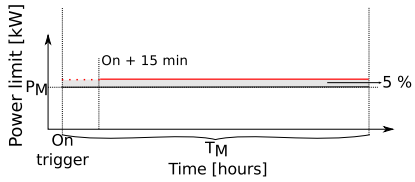


Fig. 2. The PowerMax service requirements, where the red line represent the boundaries for the permissible error, and the shaded area represents the error in service delivery which is within the limits established in the QoS.

- Deviation in size of service (dependent on P_M): Max. $\pm 5\% P_M$
- Acceptable no. of unsatisfactory activations(non-delivery): Non-del $= 4$

A graphical representation of these service requirements is depicted in Fig. 2. It is clear that the maximum acceptable error in service delivery is the shaded area. Note that the limit for non-delivery of service during the first 15 minutes of activation is dotted due to the fact that non-delivery is not counted during this period. The specifications for counting unsatisfactory activations are not clarified in [1], so it is assumed that breaking the QoS limits on one sampling period counts as one non-delivery. In the case where the service is not respected in 3 consecutive(or non-consecutive) sampling periods, Non-del $= 3$.

For example, in the case where $P_M = 5\,kW$, $T_M = 4\,h$ and the power is measured once an hour, $J_{max} = 200$, as it represents the square root of the square of the maximum (100%) permissible error over 4 hours.

This case study presents the aggregation of multiple flexible DERs via coordinated operation: 75 DERs installed in a suburban residential area, which are all connected to the same feeder leading to a 10/0.4 kV transformer. The transformer is rated to a maximum power flow of 200 kW, which is sufficient under the current load circumstances.

This case study addresses a scenario with high electric vehicle (EV) penetration, low photo-voltaic (PV) penetration and electric space heating in all households. Furthermore all DERs connected to the same LV feeder offer their flexibility to the same Aggregator. Then, the proposed performance index for service provision is evaluated for two different aggregation control algorithms: Centralized soft Model Predictive Control (C-MPC) and Distributed soft Model Predictive Control (D-MPC).

D. The reference case: without units coordination

In this section we make a scenario hypothesis for year 2050 regarding PV and EV penetration in a distribution feeder in a rural area and present simulation results. The following units are connected to the LV transformer:

- 40 buildings with electric climate control: resistive space heating with maximum load of 10 kW and air conditioning with a maximum load of 5 kW.
- 20 large EVs, with a battery size of 25 kWh, 11 kW.

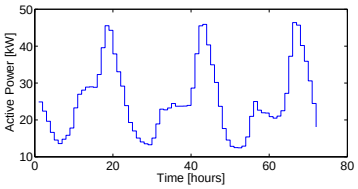


Fig. 3. The nonflexible load of the households under the transformer. The sample is statistically represantative of danish households.

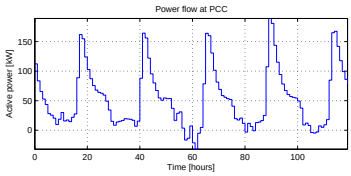


Fig. 4. PCC aggregated power flow reference case of units non coordination: Demand Response based on day ahead energy price.

- 10 small EVs, with a battery size of 14 kWh, 3.3 kW.
- 5 PV (polycrystalline) installations of 6 kW rated power each.

The PV installations provide forecasts of the production for one day ahead. To simulate uncertainty in the forecasts, Gaussian noise has been added to real data of PV production according to:

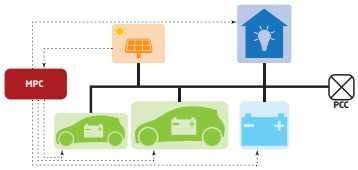
$$P_{PV-F,t} = P_{PV-T,t} + v_t, \quad v_t \sim N\left(0, \alpha\sqrt{P_{PV-T,t}}\right) \quad (5)$$

where $P_{PV-F,t}$ is the forecasted PV power production at time t and $P_{PV-T,t}$ is the actual power production at time t (from historical data). Uncertainty in solar radiation and ambient temperature are modelled in the same way. The actual power production time-series used in this case covers the same days as [12].

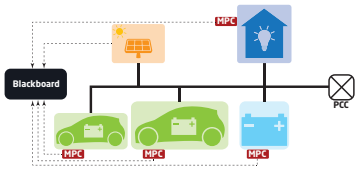
The load related to households is divided into climate control (flexible load) and everything else (non-flexible load). The building climate control is operated on MPC basis for minimum deviation from temperature set point. Regarding the non flexible household loads, a 5 day (one hour-sampled) profile of the non flexible load of 40 households is depicted in Fig. 3.

The EVs leave the charging station at a uniform randomly distributed time between 6am and 8am, and are plugged again at a uniform randomly distributed time between 4pm and 6pm. The EVs operate on dumb charging, i.e. they try to fully charge as soon as they are connected to the grid. By running a simulation of the described scenario without units coordination, the results shown in Fig. 4 are obtained.

EVs operating on dumb charging can cause peak consumption up to 190 kW. Given that the transformer capacity is 200 kW and it is customary to reserve 30% of the transformer capacity for emergency operations [13], the DSO aims at keeping the load below 140 kW and limit the inverse power flow at the substation. Thus, the DSO can sign a contract



(a) The setup of the Centralized MPC scheme.



(b) The setup of the DMPC scheme as seen in [12].

Fig. 5. The setup of the two Aggregation algorithms to be compared.

for PowerMax service (see Sec. III-C) with an Aggregator which, at any time, operates Demand Response via Direct Load Control (DLC) [2] in order to limit the power flow at the transformer. The maximum capacity available at the transformer is therefore 140 kW for direct power flow and -10 kW for inverse power flow.

The rest of this section presents the C-MPC and D-MPC formulations. For the formulation of the mathematical models we refer to [14] for the battery model and to [15] for the building space heating model (modified, as proposed in [12]). For the modelling of the services, we apply the method described in Sec.III-B. A discussion on the simulation results concludes this section.

E. The Centralized Model Predictive Control scheme

In this scheme the Aggregator contains the control algorithm to centrally manage all the units in its portfolio (Fig. III-E). Since the Aggregator optimizes its portfolios consumption through MPC, it has detailed knowledge of the state and dynamics of its portfolio. The units portfolio is the same as of the reference case. The C-MPC control problem is formulated as quadratic optimization with soft constraints (as seen in e.g. [16]):

$$\min_{u_t,\vartheta_t} \; J = \sum_{t=1}^{N} \left[\|y_t - r_t\|_Q^2 + \rho\vartheta_t + \psi\gamma_t \right] \quad (6a)$$

$$subject\ to: \quad (6b)$$

$$x_{t+1} = Ax_t + Bu_t + Ed_t \quad (6c)$$

$$y_t = Cx_t + Du_t \quad (6d)$$

$$u_{\min,t} \le u_t \le u_{\max,t} \quad (6e)$$

$$y_{\min,t} - \gamma_t \le y_t \le y_{\max,t} + \gamma_t \quad (6f)$$

$$PCC_{\min,t} - \vartheta_t \le u_t \le PCC_{\max,t} + \vartheta_t \quad (6g)$$

$$\vartheta_t \ge 0 \quad (6h)$$

$$\gamma_t \ge 0 \quad (6i)$$

where r_t is the output reference, γ is the weight for output soft constraints and ρ penalizes the *power over max*. Since this

MPC controller is centralized, the system matrices in Eq. (6b) and Eq. (6c) are formed by block diagonal-adding each of the systems' respective matrices. Being the set of units $\mathcal{S} = \{1..N\}$, it follows:

$$x = \begin{bmatrix} x_1 \\ x_j \end{bmatrix}, u = \begin{bmatrix} u_1 \\ u_j \end{bmatrix}, d = \begin{bmatrix} d_1 \\ d_j \end{bmatrix}, y = \begin{bmatrix} y_1 \\ y_j \end{bmatrix}$$

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_j \end{bmatrix}, \quad B = \begin{bmatrix} B_1 & 0 \\ 0 & B_j \end{bmatrix}$$

$$C = \begin{bmatrix} C_1 & 0 \\ 0 & C_j \end{bmatrix}, \quad D = \begin{bmatrix} D_1 & 0 \\ 0 & D_j \end{bmatrix} \tag{7}$$

$$E = \begin{bmatrix} E_1 & 0 \\ 0 & E_j \end{bmatrix}, \quad \vartheta = \begin{bmatrix} \vartheta_1 \\ \vartheta_j \end{bmatrix}, \gamma = \begin{bmatrix} \gamma_1 \\ \gamma_j \end{bmatrix}$$

where the index $j \in \mathcal{S}$ and the system in Eq. (6b) and (6c) is extended with all the units belonging to the set $\mathcal{S}$.

F. The Distributed Model Predictive Control scheme

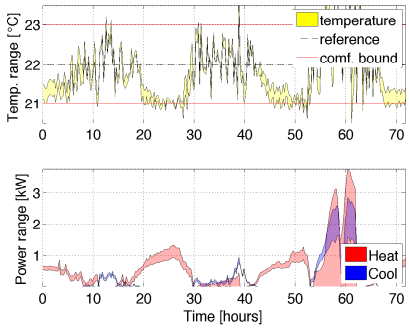
In the D-MPC formulation units within the same cluster retrieve the power plans of the other units, compute their own plan accordingly and publish it in a blackboard. Note that in this case study, differently from what it has been proposed in [12], the unit controllers have soft constraints on the outputs (temperature for buildings and SOC for batteries and EVs). In this algorithm, as soon as the units publish their consumption plan, the available power at the PCC decreases in a way such that the subsequent units communicating with the blackboard tend adjust their plan accordingly. After a negotiation period the units are entitled to operate according to the power plan that has been published in the blackboard for the next time frame. Figure III-E shows the configuration for the D-MPC. This is an example of transactional control [2], where the unit power consumption negotiated.

G. Comparison and discussion of results

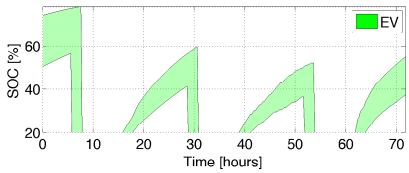
Certain assumptions have been made with regards to controllers:

- The EVs are preferably kept operating in the range $SOC = [0.2, 0.9]$ due to battery life concerns [14], although it is possible to operate in $SOC = [0.0, 1.0]$.
- The comfort band for the households lie in the band $T_{ref} = 22°C \pm 1°C$. The concept of non-delivery is not used in the asset management services, but the absolute boundaries for user comfort bands lie on $T_{ref} = 22°C \pm 1.5°C$.
- The required PowerMax service is of $P_M = 90kW$ each day in the periods of 16:30 to 20:30.
- The time sampling is of 15 minutes and the prediction horizon is of 23 hours.
- The EVs are not capable of providing Vehicle to Grid (V2G) services, i.e. EVs only charge.

These assumptions lead to the results presented Fig. 6-8 as well as in Table I. The following conclusions can be made: 1)



(a) Households temperatures and power consumption



(b) EV State of Charge

Fig. 6. Simulation results for the D-MPC with $\alpha = 0.1$

TABLE I. RESULTS OF 3-DAY SIMULATION

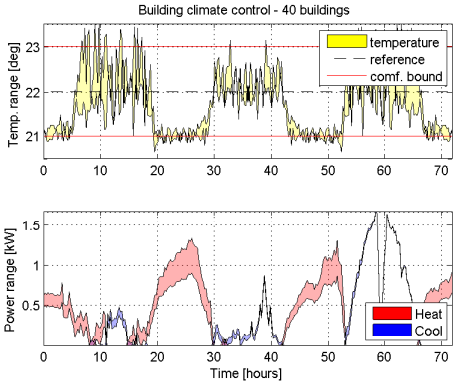
	D-MPC		C-MPC	
α	0.1	0.2	0.1	0.2
Non-delivery	0	0	0	0
η	0.0075	0.0160	xxxx	0.0153

From the graphs it can be seen that both controllers are quite good at staying within the QoS limits, as can be seen in the fact that none of the controllers have non-delivery and η is small.2) Controller performance is sensible to prediction uncertainties, as can be seen in the varying values of η depending on the uncertainty α. 3) In terms of service provision, the C-MPC outperforms the D-MPC. This arises from the fact that the C-MPC has absolute control of all units and determines a global optimum. 4) Due to behaviour difference between the local EV controllers in the D-MPC scheme, and the behaviour of the C-MPC, the power consumption of the EV is very different (compare Fig. III-G and Fig. III-G). This also leads to a vast difference in the power flow at PCC (see Fig. 8). 5)By comparing the values of η in both cases, it is clear that the difference comes from the behaviour of the household heating, where the C-MPC clearly manages to deliver a better quality service than the D-MPC.

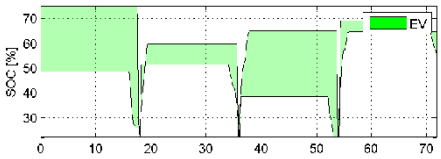
IV. OUTLOOK

The index is useful for the systematic evaluation of the adequacy of different control architectures providing ancillary services. As modelled in this paper, the index only gives a general idea of the performance of the control algorithm, but future work could include a method for differentiating the sources of high error. It is also expected that the performance index can be part of validating Aggregators for ancillary service provision.

In order to do a successfull evaluation of an aggregation

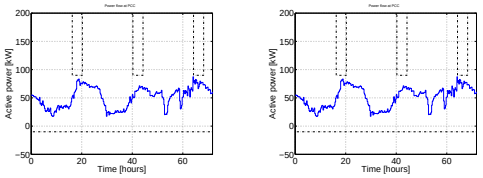


(a) Households temperatures and power consumption



(b) EV State of Charge

Fig. 7. Simulation results for the C-MPC with $\alpha = 0.1$



(a) Total power load for C-MPC (b) Total power load for D-MPC

Fig. 8. Power load at Point of Common Coupling for the controllable loads

algorithm, it is important that the QoS specifications of the future ancillary services are well defined. This is a challenge in itself since many of the ancillary services assume a production baseline, which poses many problems (see e.g. [17]).

The evaluation of aggregation control algorithms is an important part of a general verification framework for service providing control algorithms. Future work will also include performance assessment in laboratory implementation, as well as modelling of ancillary services provided through DSM.

V. Conclusion

We wanted to formulate a method for assessing performance of control service provision of aggregation algorithms. In order to do this, a performance index was defined by drawing inspiration from the field of Control Performance Assessment.

A case study was presented, where two different control algorithms were evaluated and the results were presented.

Evaluation of aggregation algorithms is expected to be an important part of validation of Aggregators providing DSM.

Acknowledgment

The authors acknowledge the financial support of iPower (www.ipower-net.dk). The authors thank Shi You from DTU Elektro for providing data for the non-controllable loads.

References

[1] N. C. Nordentoft, C. Zhang, Y. Ding, L. H. Hansen, P. Cjar, H. Bindner, and P. Brath, "Development of a DSO-market on flexibility services," iPower, Tech. Rep., mar 2013.

[2] A. Kosek, G. T. Costanzo, H. Bindner, and O. Gehrke, "An overview of demand side management control schemes for buildings in smart grids," in *Proc. IEEE Int. Conf. Smart Energy Grid Eng. (SEGE'13)*, Oshawa, Canada, aug 2013, p. to be published.

[3] B. Biegel, J. Stoustrup, J. Bendtsen, and P. Andersen, "Model predictive control for power flows in networks with limited capacity," in *2012 American Control Conference*, 2012, pp. 2959–2964.

[4] K. Heussen, S. You, B. Biegel, L. H. Hansen, and K. B. Andersen, "Indirect control for demand side management - A conceptual introduction," in *Proc. 3rd IEEE PES Int. Conf. Innovative Smart Grid Technologies Europe*, Berlin, Germany, oct 2012.

[5] Y. Rebours, "A comprehensive assessement of markets for frequency and voltage control ancillary services," Ph.D. dissertation, University of Manchester, School of Electrical and Electronic Engineering, 2008.

[6] Energinet.dk, "Ancillary services to be delivered in denmark tender conditions," On-line, oct 2012.

[7] G. Strbac, "Demand side management: Benefits and challenges," *Energy Policy*, vol. 36, no. 12, pp. 4419 – 4426, 2008, foresight Sustainable Energy Management and the Built Environment Project. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0301421508004606

[8] M. Jelali, "An overview of control performance assessment technology and industrial applications," *Control Engineering Practice*, vol. 14, no. 5, pp. 441 – 466, 2006, intelligent Control Systems and Signal Processing ICONS 2003, IFAC International Conference on Intelligent Control Systems and Signal Processing. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0967066105002479

[9] T. Green, R. Izadi-Zamanabadi, and H. H. Niemann, "On the choice of performance assessment criteria and their impact on the overall system performance - The refrigeration case study," in *Proc. IEEE Conf. Control and Fault Tolerant Systems (SysTol'10)*, Niece, France, 2010, pp. pp. 624–629.

[10] K. Åstrom, "Assessment of achievable performance of simple feedback loops," *International journal of adaptive control and signal processing*, vol. 5, no. 1, pp. 3–19, 1991.

[11] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control-Analysis and Design*, 2nd ed. Chichester, England: Wiley, 2005.

[12] G. T. Costanzo, O. Gehrke, D. E. M. Bondy, F. Sossan, H. W. Bindner, J. Parvizi, and H. Madsen, "A coordination scheme for distributed model predictive control: Integration of flexible DERs," in *Proc. 4th IEEE PES Int. Conf. Innovative Smart Grid Technologies Europe*. Copenhagen, Denmark: IEEE, oct 2013, p. to be published.

[13] T. M. Engel, "Optimal utilization of the electricity distribution network," 2011, in danish.

[14] F. Marra, G. Y. Yang, C. Traholt, E. Larsen, C. Rasmussen, and S. You, "Demand profile study of battery electric vehicle under different charging options," in *Power and Energy Society General Meeting, 2012 IEEE*, 2012, pp. 1–7.

[15] P. Bacher and H. Madsen, "Identifying suitable models for the heat dynamics of buildings," *Energy and Buildings*, vol. 43, no. 7, pp. 1511 – 1522, 2011.

[16] G. Prasath and J. B. Jørgensen, "Soft constraints for robust mpc of uncertain systems," *International Symposium on Advanced Control of Chemical Processes*, p. 54, 2009.

[17] S. Borenstein, "Money for nothing," www.energypolicyblog.com/2014/05/14/money-for-nothing/, 2014.

Experimental analysis of data-driven control for a building heating system

G.T. Costanzo[a], S. Iacovella[b,d], F. Ruelens[b,d], T. Leurs[b], B.J. Claessens[c,d,*]

[a] *Danish Technical University, Department of Electrical Engineering, Frederiksborgvej 399, 4000 Roskilde, Denmark.*
[b] *Electrical Engineering of KU Leuven, Kasteelpark Arenberg 10, bus 2445, 3001 Leuven, Belgium.*
[c] *Flemish Institute for Technological Research (VITO), Boeretang 200, B-2400 Mol, Belgium.*
[d] *EnergyVille, Thor Park Poort Genk 8130, 3600 Genk, Belgium.*

Abstract

Driven by the opportunity to harvest the flexibility related to building climate control for demand response applications, this work presents a data-driven control approach building upon recent advancements in reinforcement learning. More specifically, model-assisted batch reinforcement learning is applied to the setting of building climate control subjected to a dynamic pricing. The underlying sequential decision making problem is cast on a markov decision problem, after which the control algorithm is detailed. In this work, fitted Q-iteration is used to construct a policy from a batch of experimental tuples. In those regions of the state space where the experimental sample density is low, virtual support samples are added using an artificial neural network. Finally, the resulting policy is shaped using domain knowledge. The control approach has been evaluated quantitatively using a simulation and qualitatively in a living lab. From the quantitative analysis it has been found that the control approach converges in approximately 20 days to obtain a control policy with a performance within 90% of the mathematical optimum. The experimental analysis confirms that within 10 to 20 days sensible policies are obtained that can be used for different outside temperature regimes.

Keywords:
Thermostatically controlled load, batch reinforcement learning, demand response, data-driven modelling, fitted Q-iteration

1. Introduction

Perez *et al.* estimate that 20 to 40% of the global energy is consumed in buildings [1]. About half of this energy is used for Heating, Ventilation and Air Conditioning (HVAC) [2]. As a consequence, control strategies for HVAC have received considerable academic attention in recent years. A popular class of control strategies is that of model-based strategies, such as Model Predictive Control (MPC) [3]. MPC for HVAC systems has been largely investigated in the recent literature [4, 5, 6, 7, 8], in both of its main aspects, modelling [9, 10], and control [11]. In MPC, at regular time intervals, a control action is selected by solving an optimization problem over a finite time horizon, which is typically a day for HVAC control. In MPC the impact of future disturbances, such as internal heating and meteorological conditions, is taken into account using forecasts. Predictive control allows using the load flexibility related to thermal storage, e.g. through the thermal inertia of the building or through direct heat storage [12].

This flexibility can be harvested to enable demand response and provide load control services, which value has been increasing together with the share of renewable energy in the production mix. Examples of services are peak shaving and valley filling for a distribution system operator [13], ancillary services towards a transmission system operator [14, 15] or energy arbitrage [16]. However, deploying MPC can be a challenging task. The most significant challenge is to derive an accurate model which, in the case of thermal control, has to include the thermal dynamics and the actuation model. In [17], Širokỳ *et al.* give a detailed report on implementation issues of MPC controllers for building heating systems.

In this context, completely data-driven approaches are deemed interesting, sacrificing performance for practicality. One possible embodiment uses data-driven model in combination with an optimization algorithm to obtain a control policy [18]. Alternatively, it is possible to learn directly the control policy by estimating a state-action value function through interaction with the system. For example in [19], Reinforcement Learning (RL), a model-free control approach is applied to building thermal storage. In RL, the policy is updated online, i.e. at each time step. In Batch Reinforcement Learning (BRL), on the other hand, the policy is calculated offline using a batch of historical data. Even though (B)RL is getting more mature [20], as discussed in [21], combining techniques of RL with prior (domain) knowledge is a logical control paradigm. It is towards this direction that this paper is positioned, i.e. in applying BRL in combination with prior knowledge to the operation of a building climate control system for demand response applications.

The basis of our approach is BRL with Fitted Q-Iteration (FQI) [22, 23], where the learning of an optimal control policy is enhanced by *virtual* data coming from a model. For this reason, such approach is called Model-Assisted Batch Reinforcement Learning (MABRL) as discussed in [24].

*Corresponding author: tel.: +3214335910;
e-mail: { bert.claessens@vito.be}

In Section 2 an overview of the related literature is provided and the contribution of this work is explained. Following the approach presented in [25], in Section 3 the building thermal scheduling is formalised as a sequential decision making problem under uncertainty. In Section 4 MABRL is detailed, while Section 5 presents a quantitative and qualitative assessment of the performance of the controller. Finally, Section 6 outlines the conclusions and discusses future research directions.

2. Related Work

This section gives a non-exhaustive overview of related work regarding MPC and RL for building climate control, after which the main contributions of this work are explained.

2.1. Model Predictive Control

When considering building climate control, MPC has received considerable attention in the recent literature [6, 7, 8, 26]. The overview of practical issues related to the implementation of an MPC controller can be found in [27]. The key elements of an MPC comprise mathematical models describing the building dynamics, comfort requirements and exogenous information such as user behavior and outdoor temperature. This information is used to cast an optimization problem that is solved to define optimal control actions with respect to a defined objective function, subject to constraints provided by the model.

In typical embodiments of MPC one tries to formalize the problem as a mixed integer problem to allow using fast solvers with performance guarantees. Therefore, a Linear Time Invariant model (LTI) of the system under control is to be identified. If no domain knowledge is available, black-box identification techniques are used, such as subspace identification methods [28, 29]. Alternatively, gray-box models can be used, where the model structure is defined and the parameters are estimated using experimental data [9]. In the context of thermal modelling a number of studies use thermal circuits [30, 31, 32, 33, 34].

Advanced climate control allows, besides efficient use of energy and comfort management, integration within aggregation schemes to provide ancillary services and portfolio management in demand side management [35]. For example, in [36] the aggregated flexibility of a cluster of buildings is used to provide balancing services using an aggregate-and-dispatch approach.

An alternative for LTI modelling is to use non-linear data-driven models, such as artificial neural networks (ANNs) [18, 37], in combination with Dynamic Programming (DP) [38] to compute a control policy. This form of control can be seen as a form of RL [39].

2.2. Reinforcement Learning

As discussed in Section 1 RL is a model-free control technique whereby a control policy is learned from interactions with the environment. A well established reinforcement learning method is Q-learning [40] where the state-action value function, or Q-function, is learned. Compared to techniques provided in the previous section, RL mitigates the risk of model-bias [24] as a policy is build around the data. When considering Q-learning and its applications to demand response, mainly traditional Q-learning has been used [41, 42, 19]. More recently BRL [43, 44] in the form FQI [21] has been investigated. The main advantage of BRL is the practical learning time required for convergence (20-40 days in [43, 44]) which comes at the cost of an increased computational complexity. Although BRL can rival the performance of MPC techniques, as indicated in [21], the context of demand response allows to add *prior knowledge* to the optimal control problem that can result in faster convergence. A first approach uses prior knowledge by shaping the policy, obtained with FQI, by means of constrained regression [22]. A second approach is described by Lampe *et al.* in [24]. Here virtual data from a model is used together with experimental data to obtain an approximation of the Q-function (state-action value function).

Building upon [43, 22, 24], this work has the following contributions:

- BRL in the form of FQI [21] in combination with virtual trajectories [24] and policy shaping is applied to a HVAC system for a typical objective of dynamic pricing [45]. This effectively results in a data-driven solution for building climate control systems, combining state of the art BRL with domain knowledge;
- Quantitative and qualitative performance assessment of MABRL in a simulated and experimental environment, where the operation of an air conditioner is subject to dynamic energy pricing.

3. Problem Formulation

Before presenting the control approach in Section 4, this Section formulates the decision-making process as a Markov Decision Process (MDP) [38, 46]. An MDP is defined by its state space X, its action space U, and a transition function f:

$$\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k), \tag{1}$$

which describes the dynamics from $\mathbf{x}_k \in X$ to $\mathbf{x}_{k+1}$, under the control action $\mathbf{u}_k \in U$, and subject to a random process $\mathbf{w}_k \in W$, with probability distribution $p_w(\cdot, \mathbf{x}_k)$. The reward accompanying each state transition is r_k:

$$r_k(\mathbf{x_k}, \mathbf{u_k}, \mathbf{x_{k+1}}) = \rho(\mathbf{x}_k, \mathbf{u}_k, \mathbf{w}_k) \tag{2}$$

which is here considered to a cost, as it accounts for the energy price. Therefore, the objective is to find a control policy $h : X \rightarrow U$ that minimises the T-stage cost starting from state $\mathbf{x_1}$, denoted by $J^h(\mathbf{x_1})$:

$$J^h(\mathbf{x_1}) = \mathbb{E}\left(R^h(\mathbf{x_1}, \mathbf{w_1}, ..., \mathbf{w_T})\right), \tag{3}$$

with:

$$R^h(\mathbf{x_1}, \mathbf{w_1}, ..., \mathbf{w_T}) = \sum_{k=1}^{T} \rho(\mathbf{x_k}, h(\mathbf{x_k}), \mathbf{w_k}). \tag{4}$$

It is worth remarking that an optimal control policy, here denoted by h^*, satisfies the Bellman equation:

$$J^{h^*}(\mathbf{x}) = \min_{\mathbf{u}} \mathop{\mathbb{E}}_{\mathbf{w}\sim P_w(.|\mathbf{x})} \{\rho(\mathbf{x}, h(\mathbf{x}), \mathbf{w}) + J^{h^*}(f(\mathbf{x}, h(\mathbf{x}), \mathbf{w}))\} \quad (5)$$

Typical techniques to find policies in an MDP framework are value iteration, policy iteration, and policy search [22]. As mentioned earlier, in this work MABRL (related to value iteration) is considered.

3.1. State description

Following the approach presented by Ruelens *et al.* in [25], it is assumed that the state space X consists of: time-dependent state information X_t, controllable state information X_{phys}, and exogenous (uncontrollable) state information X_{ex}:

$$X = X_t \times X_{phys} \times X_{ex}. \quad (6)$$

In the following, each component of the state space is detailed.

3.1.1. Timing

The time-dependent information component X_t contains information related to timing. In this implementation the quarter hour during the day has been used:

$$X_t = \{1, \ldots, 96\}, \quad (7)$$

in order to identify behavioral daily patterns. Extending this with, e.g. the day of the week, can be done at little extra cost. However, this extension is outside the scope of this work.

3.1.2. Physical representation

The controllable state information $x_{phys,k}$ consists of the indoor air temperature, T_k:

$$x_{phys,k} = T_k \mid \underline{T}_k < T_k < \overline{T}_k \quad (8)$$

where $\underline{T}_k$ and $\overline{T}_k$ denote the lower and upper bound set by the end consumer.

3.1.3. Exogenous Information

The exogenous (uncontrollable) information $\mathbf{x_{ex,k}}$ is considered to have an impact on $x_{phys,k}$, but it is invariant for control actions $\mathbf{u_k}$. In this study the exogenous state information consists of the outside temperature, $\mathbf{T_o}$, and the solar radiance, $\mathbf{S}$:

$$\mathbf{x_{ex,k}} = (T_{o,k}, S_k). \quad (9)$$

In this work it assumed that a forecast of the outside temperature and the solar radiance is available when constructing the policy h, as will be detailed in Section 4.2 ($\hat{.}$ is used to denotes a forecast).

3.1.4. Control action

In this work the control action is a binary value indicating if the HVAC system should switch ON or OFF:

$$u_k \in \{0, 1\}. \quad (10)$$

The control action of the previous control event u_{k-1} is also added to the state information, as it is relevant for the dynamics of the HVAC system. In fact, its value will be used to avoid too frequent switching as discussed in Section 3.2. As a result, the final state vector is defined as:

$$\mathbf{x_k} = (x_{t,k}, T_k, T_{o,k}, S_k, u_{k-1}). \quad (11)$$

As this state vector only contains part of the actual state of the system, a common approach to enrich the state vector is to add previous state samples [47]. This, however, results in an increased state dimension that could be reduced by means of feature extraction, e.g. non-linear principal component analysis [48]. In this work however, the state vector is defined according to (11).

3.2. Backup controller and physical realisation

The HVAC system is assumed to be equipped with a backup controller, which acts as a filter to the control actions resulting from the policy h. The function $B : X \times U \longrightarrow U$ maps the requested control action u_k taken in state x_k to a physical control action u_k^{phys}:

$$u_k^{phys} = B(\mathbf{x_k}, u_k, \theta), \quad (12)$$

with θ containing system specific information. In this case, θ contains $\underline{T}_k$ and $\overline{T}_k$, and $B(\cdot)$ is defined as:

$$B(\mathbf{x_k}, u_k, \theta) = \begin{cases} 1 & \text{if} \quad T_k \leq \underline{T}_k \\ 1 & \text{if} \quad T_k \leq \overline{T}_k \ \wedge \ u_{k-1} = 1 \\ u_k & \text{if} \quad T_k \leq \overline{T}_k \ \wedge \ u_{k-1} = 0 \\ 0 & \text{if} \quad T_k > \overline{T}_k \end{cases}, \quad (13)$$

3.3. Reward model

As discussed in the introduction, different applications can be considered to harvest the flexibility related to climate control of buildings. In dynamic pricing or energy arbitrage [16], i.e. responding to an external price vector λ, the reward function is defined as:

$$\rho\left(\mathbf{x_k}, u_k^{phys}, \lambda_k\right) = -P\Delta t \lambda_k u_{phys,k}, \quad (14)$$

with P is the average power consumption of the air conditioner in during the time interval Δt.

4. Model-Assisted Batch Reinforcement Learning

As discussed in Section 2.2, the control policy h^* is obtained using MABRL in combination with policy shaping, as illustrated in Figure 1. To this end, FQI is used to obtain an approximation $\widehat{Q}^*$ of the state-action value function Q^* from a batch of four tuples $\mathcal{F}^{\mathcal{L}}$, as detailed in [49]:

$$\mathcal{F}^{\mathcal{L}} = \left\{(\mathbf{x}_l, u_l, r_l, \mathbf{x}'_l),\ l = 1, ..., \#\mathcal{F}^{\mathcal{L}}\right\}, \quad (15)$$

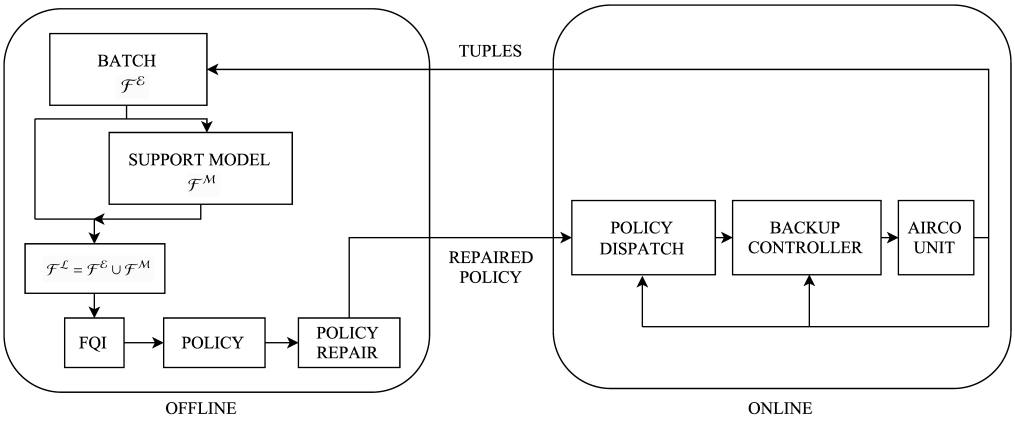


Figure 1: Overview of the information flow in this implementation of MABRL using FQI and policy repair.

where $\mathbf{x}'_l$ denotes the successive state in time to $\mathbf{x}_l$. As illustrated in Figure 1, $\mathcal{F}^{\mathcal{L}}$ is a combination of experimentally observed tuples $\mathcal{F}^{\mathcal{E}}$ and virtual tuples generated by a model $\mathcal{F}^{\mathcal{M}}$:

$$\mathcal{F}^{\mathcal{L}} = \mathcal{F}^{\mathcal{E}} \cup \mathcal{F}^{\mathcal{M}} . \tag{16}$$

From the resulting $\widehat{Q}^*(x, u)$ a control action u_k can be obtained following

$$u_k \in \arg\min_u \widehat{Q}^*(\mathbf{x}_k, u). \tag{17}$$

In the following subsections we detail the algorithms behind each block in Figure 1.

4.1. The support model: Artificial Neural Network

Following the approach presented in [24], an ANN is used to represent a support model. In this work, single-layer, single-output Extreme Learning Machines (ELMs) are trained to predict the change of internal temperature, ΔT. These are used as they allow for fast training of the weights of the network at the expense of reduced regression performance. The latter is partially mitigated by combining multiple ELMs in an ensemble [50].

The output of an ANN with p input neurons and n hidden neurons can be formulated as:

$$y(\mathbf{x}) = \sum_{i=1}^{n} \beta_i g_i(\mathbf{w_i} \cdot \mathbf{x},\ b_i) = \mathbf{G(x)}\boldsymbol{\beta} , \tag{18}$$

where $\mathbf{x} \in \mathbb{R}^p$ is the input vector, $\mathbf{w_i} \in \mathbb{R}^p \sim i.i.d.\, U(-1, 1)$ is the weight vector connecting the input nodes with the i-th hidden node, $b_i \in \mathbb{R} \sim U(0, 1)$ is the bias of the i-th hidden node, and $\beta_i \in \mathbb{R}$ is the output weight of i-th hidden node. $\boldsymbol{\beta} = [\beta_1 \ldots \beta_n]^T$ is the output weight matrix, while $\mathbf{G(x)}$ is the output matrix of the hidden nodes, where the nodes activation function g is a sigmoid.

Since the parameters of the hidden nodes are randomly generated, training an ELM corresponds to determining the output weight matrix $\boldsymbol{\beta}$ based on the least-squares solution of $\mathbf{G}\boldsymbol{\beta} = \mathbf{Y}$:

$$\boldsymbol{\beta} = \left(\frac{1}{C} + \mathbf{G}^T\mathbf{G}\right)^{-1} \mathbf{G}^T\mathbf{Y} , \tag{19}$$

where the regularisation term $\frac{1}{C}$ is used to enhance the robustness and the generalisation performance of the solution [51].

Finally, the ensemble model output $y_{ens}(\mathbf{x})$ is given by a weighted average of M individual ELMs outputs:

$$y_{ANN}(\mathbf{x}) = \frac{1}{M}\sum_{i=1}^{M} y_i(\mathbf{x}) . \tag{20}$$

The input vector at time k is $\mathbf{x_k} = (x_{t,k}, T_k, T_{o,k}, S_k, u_{k-1})$, and the output is $y_k = \Delta T_k = T_{k+1} - T_k$. For a deeper insight on ELMs, we refer to [52]. As the training process is fast, finding the appropriate number of hidden nodes is done using cross-validation.

4.2. Fitted Q-Iteration

A popular BRL technique that found its way into several practical implementations is FQI [49]. Typically, BRL techniques construct policies based on a batch of tuples. However, since in this context the reward function is known a priori and the resulting actions of the backup controller can be measured, Algorithm 1 uses tuples of the form $\left(\mathbf{x}_l, u_l, \mathbf{x}'_l, u_l^{phys}\right)$. Algorithm 1 shows how FQI [49] can be used in a demand response application when a forecast of the exogenous data is available. Here $\hat{\mathbf{x}}'_l$ denotes the successor state to $\mathbf{x}_l$. In Algorithm 1, the observed external temperature and solar radiance in $\mathbf{x}'_{ex,l}$ are replaced by their forecasted value $\hat{\mathbf{x}}'_{ex,l}$ (line 7 in Algorithm 1). As such, $\widehat{Q}^*$ becomes biased towards the provided forecast. Before constructing $\widehat{Q}^*$ a set $\mathcal{F}^{\mathcal{M}}$, containing at most n virtual samples is created with random state-action pairs. However, a randomly generated tuple is only accepted to $\mathcal{F}^{\mathcal{M}}$ if the nearest experimental tuple in $\mathcal{F}^{\mathcal{E}}$ falls outside a predefined radius r, following

Algorithm 1 Model-assisted Fitted Q-iteration using a forecast of the exogenous data

Input: $\mathcal{F}^{\mathcal{E}} = \left\{\mathbf{x}_l, u_l, \mathbf{x}'_l, u_l^{phys}\right\}_{l=1}^{\#\mathcal{F}^{\mathcal{E}}}$, $\hat{X}_{ex} = \left\{\hat{\mathbf{x}}_{\text{ex},k}\right\}_{k=1}^{T}$, y_{ANN}, r, n, C λ, θ

$\mathcal{F}^{\mathcal{M}}$ = generateSamples($\mathcal{F}^{\mathcal{E}}, y_{ANN}, r, n, C, \theta$)
$\mathcal{F}^{\mathcal{L}} = \mathcal{F}^{\mathcal{E}} \cup \mathcal{F}^{\mathcal{M}}$
let $\widehat{Q}_0$ be zero everywhere on $X \times U$
for $N = 1, \ldots, T$ **do**
  for $l = 1, \ldots, \#\mathcal{F}^{\mathcal{L}}$ **do**
    $r_l \leftarrow \rho\left(\mathbf{x_l}, u_l^{phys}, \lambda\right)$
    $\hat{\mathbf{x}}'_l \leftarrow \left(x'_{t,l}, T'_l, \hat{T}'_{o,l}, \hat{S}'_l, u'_{l-1}\right)$
    $Q_{N,l} \leftarrow r_l + \min_{u \in U} \widehat{Q}_{N-1}(\hat{\mathbf{x}}'_l, u)$
  end for
  use regression to obtain $\widehat{Q}_N$ from $\mathcal{T}_{\text{reg}} = \left\{\left((\mathbf{x}_l, u_l), Q_{N,l}\right), l = 1, \ldots, \#\mathcal{F}^{\mathcal{L}}\right\}$
end for
return $\widehat{Q}^* = \widehat{Q}_N$

Algorithm 2 generateSamples($\mathcal{F}^{\mathcal{E}}, y_{ANN}, r, n, C, \theta$)

Input: $\mathcal{F}^{\mathcal{E}}$, y_{ANN}, r, n, C, θ

$\mathcal{F}^{\mathcal{M}} = \{\emptyset\}$
$N \leftarrow 0$
while $\#\mathcal{F}^{\mathcal{M}} < n$ and $N < C$ **do**
  generate random state action sample $\{\mathbf{x}_k, u_k\}$
  $d = \min_{(\mathbf{x},u) \in \mathcal{F}^{\mathcal{E}}} \{\|\mathbf{x} - \mathbf{x}_k\| + \|u - u_k\|\}$
  if $d > r$ **then**
    $u_k^{phys} = B(\mathbf{x}_k, u_k, \theta)$
    $T'_{l,k} = y_{ANN}(\mathbf{x}_k, u_k^{phys}) + T_{l,k}$
    $\mathcal{F}^{\mathcal{M}} = \mathcal{F}^{\mathcal{M}} \cup \left\{(\mathbf{x}_k, u_k, \mathbf{x}'_k, u_k^{phys})\right\}$
  end if
  $N \leftarrow N + 1$
end while
return $\mathcal{F}^{\mathcal{M}}$

a distance metric Δ.[1] In order to keep Algorithm 2 tractable in the event of a dense set $\mathcal{F}^{\mathcal{E}}$, a computational budget C is added.

The distance metric is defined as: $\Delta\left((\mathbf{x}, \mathbf{x}'), (u, u')\right) = \|\mathbf{x} - \mathbf{x}'\| + \|u - u'\|$, where $\|\cdot\|$ is the Euclidean norm. Algorithm 2 uses the artificial network y_{ANN} to generate the virtual tuples as in (20)

Similarly as in [49], Algorithm 1 uses an ensemble of extremely randomized trees [53] as regression algorithm to estimate the Q-function. In principle, other regression algorithms, such as artificial neural networks or support vector machines can be applied.

4.3. Policy Repair

This section shows how to shape a policy h^* by using triangular Membership Functions (MFs) [22] and expert knowledge to enforce monotonicity in the policy.

[1] Note that this radius can also be defined based upon local inter-tuple distances. This is however to be explored in future work.

The centers of the triangular MFs are located on an equidistant grid with N_g MFs along each dimension of the state space. This partitioning leads to $(N_g^{|X_{phys}|})$ state-dependent MFs for each action.

The parameter vector θ_g^* that approximates the original policy can be found by solving the following least-squares problem:

$$\theta_g^* \in \arg\min_{\theta_g} \sum_{l=1}^{\#\mathcal{F}^{\mathcal{L}}} \left([F(\theta_g)](x_l) - h(x_l)\right)^2, \quad \text{s.t. expert knowledge} \tag{21}$$

where F denotes an approximation mapping of a weighted linear combination of triangular MFs and $[F(\theta_g)](x)$ denotes the policy $F(\theta_g)$ evaluated at state x. A more detailed description of how these triangular MFs are defined can be found in [22].

4.4. Policy dispatch

The policy dispatch block is in charge of operating the air conditioner according to the policy coming from the MABRL controller. The airco controller selects each action in any encountered state with nonzero probability (exploration), or dispatches the optimal control action by exploiting the acquired knowledge (exploitation) [40]. A common technique to balance the exploration-exploitation in RL is ε-greedy exploration:

$$u_k = \begin{cases} u \sim \text{Bernoulli}\,(0.5) & \text{if } \gamma \leq \varepsilon_j \\ [F(\theta_g^*)](\mathbf{x}_k) & \text{if } \gamma > \varepsilon_j \end{cases}, \tag{22}$$

where $\gamma \sim U(0, 1)$. In this study the exploration factor ε_j reduces by half every four days:

$$\varepsilon_j = \frac{\varepsilon_0 \varsigma}{\varsigma + j - 1}, \tag{23}$$

where the decay factor ς is 4, the initial probability ε_0 is 0.4, and the the day index is $j \in \{1, 2, 3, \ldots\}$.

5. Performance assessment

This section provides a qualitative and quantitative performance assessment of the controller discussed in Section 4.2. Section 5.1 presents a quantitative assessment in a setting where the online part in Figure 1 is simulated via an equivalent thermal parameter model [54] of the air conditioner and the building (further detailed in Appendix A). Section 5.2 provides a qualitative analysis of MABRL performance in a real climate control application.

5.1. Simulated environment

In order to test the convergence and the performance of the MABRL controller, a benchmark is required. As the experimental setup is a living lab, exact external conditions cannot be reproduced from day to day. Thus, the model described in Appendix A has been used to simulate the air conditioner and the lab room. An optimal solution to thermal scheduling obtained using MPC is taken as benchmark and is depicted in

Figure 2. This benchmark solution is explained in more detail in Appendix B.

The simulations have been performed using different exogenous information and price profiles from day to day [55]. In Figure 2 the top graph shows the cumulated cost of the different controllers: BRL, MABRL, Optimal (MPC), and default thermostatic control.

These results indicate that indeed the control approach as presented in Section 4 is able to find near optimal control policies in a learning time of approximately 20 days, after which the performance relative to a mathematical optimum is stable. However, adding virtual samples, as illustrated in the procedure in Section 4.2, has limited contribution, with a slight economic advantage of the MABRL approach over the BRL. The following subsection presents the policy computation on the basis of experimental data, together with its experimental validation in a living lab.

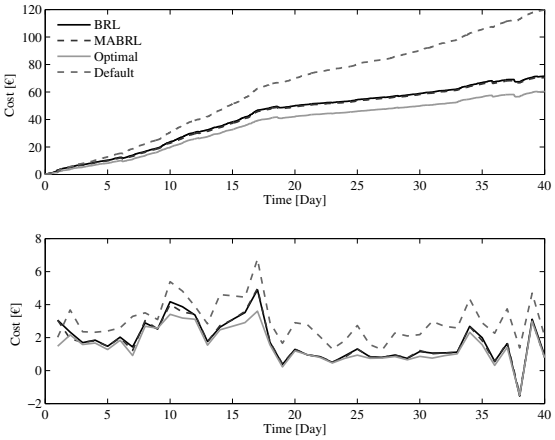


Figure 2: Cost performance of the different controllers: BRL (no virtual tuples), MABRL, Optimal (MPC), and Default (hysteresis). Top plot: cumulative electricity cost. Bottom plot: daily electricity cost.

5.2. Experimental environment

5.2.1. Ability to integrate forecasts

A first analysis focuses on the ability MABRL to effectively take into account the forecast of exogenous information. To this end, Figure 3 shows several policies organised in a matrix for different predicted profiles of the external temperature (from left to right) and increasing size of the learning batch (from top to bottom). For clarity no virtual traces have been added. During the experiments control actions have been taken every 5 minutes.

Starting from the top row, policies are computed using experimental batches of: two days, eight days, and sixteen days. The black areas in the policies correspond to the control action *switch on*. Note that, recalling Eq. 13, once the heating system has been switched on it continues heating until the temperature upper bound is reached in order to avoid frequent switch events.

Consider Figure 3, from left to right only the profile of external temperature is changed. One can observe that the policies are highly influenced by the weather conditions: as soon as the external temperature is low, the policies correspond to basic thermostatic control (i.e. heat only when it is necessary). Finally, note that the policies in the first row resemble each others, regardless of the different forecast of the external temperature, due to the fact that the training tuples are few.

5.2.2. Policy shaping

Figure 4 shows the effect of shaping the policy as discussed in Section 4.2. Here triangular membership functions [22] have been used with the constraint that the policy needs to be strictly decreasing with increasing indoor temperature. This is a direct consequence from the physical understanding that a room at a higher temperature is subjected to higher losses to the environment. The results depicted in Figure 4 show that the policies are smoothed following the monotonicity constraints.

5.2.3. Effect of adding virtual samples

Figure 5 shows the impact of virtual tuples on the convergence of the policy. The policy is computed using experimental data and virtual tuples in three different proportions. The left chart in Figure 5 shows the computed policy with only few experimental data points (2 days) and no virtual tuples in support, which is called *early policy*.

The right chart shows the policy computed using a large set of experimental data, without virtual tuples (20 days), called *regime policy*. The middle chart shows the computed policy using few field data (2 days) and a large share of virtual tuples (400 samples) from the support model, which is trained on the basis of the few field data available. This latter is called *model-assisted policy*. One can recognise that the model-assisted policy is much more similar to the regime policy than the early policy, concluding that the support model enhances the policy convergence. Furthermore, adding virtual tuples results in more smooth policies.

5.2.4. Power profiles

Figure 6 shows the implementation of a regime policy (after 12 days). The bottom chart depicts the internal temperature and the control policy, where the black areas correspond to the control action *switch on* (recall that the heating is kept on until the temperature upper bound is reached, in order to avoid oscillations). The second last chart reports the power consumption of the air conditioning units power and the intraday balancing price. The second and first charts show the external and internal air temperature, respectively.

Figure 6 shows the implementation of a regime policy on a warmer day (16 days). In this case, since the thermal losses are moderate with respect to a cold day, the optimiser allows longer pre-heating in order to avoid high price zones. This doesn't occur in cold days, where the policy is much more similar to a simple thermostatic logic.

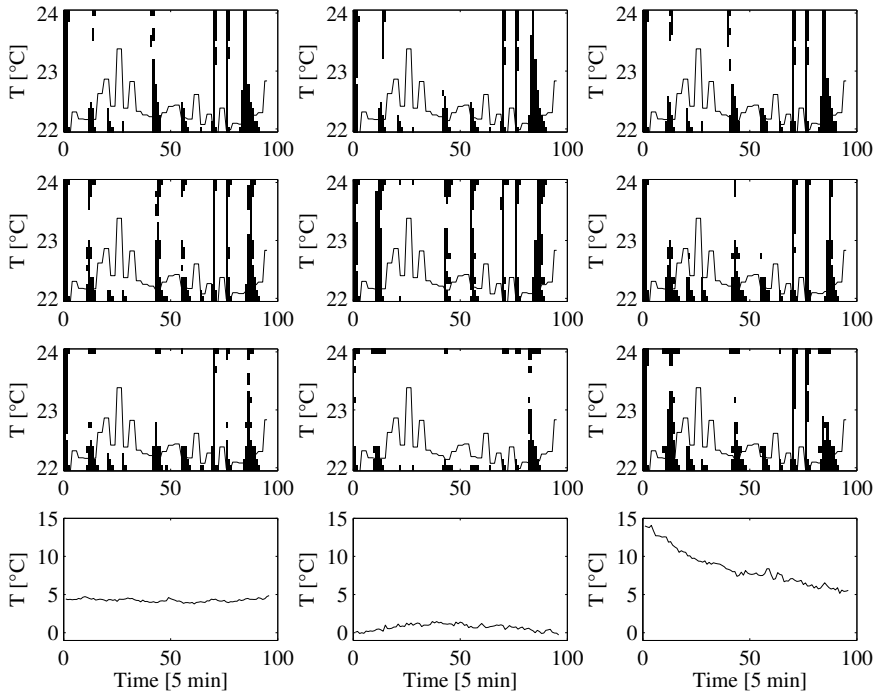


Figure 3: Policies projections obtained from experimental data for different forecast of the outside temperature, depicted in the lower row. From top to bottom data from 2, 8 and 16 days has been used. The lowest row indicates the forecasts of outside temperature for which the policies have been calculated.

6. Conclusions and Disussion

In this work model-assisted batch reinforcement learning has been deployed to harvest the flexibility related to building climate control. The results from both a quantitative analysis based upon a simulation show that a performance within 90% of a mathematical optimum is obtained within approximately 20 days. These results are confirmed qualitatively after deploying the control approach in a living lab. After collecting data for approximately 16 days the approach is able to generalise policies for different outside temperature forecasts. Adding virtual tuples from a support model result in a limited performance increase.

Future work will be aimed towards evaluating the presented algorithm for different objectives related to demand response. Furthermore, other types of models, such as gray-box models, will be used to add virtual tuples.

Acknowledgements

This work has been done under the FP-7 program Resilient. The authors would like to thank Peter Vrancx, Damien Ernst and Raphael Fonteneau for valuable discussions. The authors acknowledge the financial support of iPower, a project within the Danish Strategic Platform for Innovation and Research within Intelligent Electricity (www.ipower-net.dk).

Appendix A. ETP model

The data batch of the simulated experiments is provided by an Equivalent Thermal Parameter model (ETP) [54] that is fitted on experimental data from the living lab. The heat flow for the ETP model of a residential heating/cooling system is defined as follows:

$$\begin{aligned} \dot{T}_a &= \frac{1}{C_a}\left[\frac{1}{R_a}(T_o - T_a) + \frac{1}{R_m}(T_m - T_a) + A_s Q_s + A_c Q_{AC}\right] \\ \dot{T}_m &= \frac{1}{C_m}\left[\frac{1}{R_m}(T_a - T_m) + (1 - A_s) Q_s\right] \end{aligned}, \tag{A.1}$$

where C_m equals the thermal mass of the building envelope. T_o is the outside air temperature, T_a is the inside air temperature and T_m is the envelope temperature. R_m is the resistance between the inner air and the envelope, while C_a represents the thermal mass of the air. The heat flux into the interior air mass is given by a fraction A_s of the solar heat gain Q_s, a fraction A_c of the heat gains of the air conditioners Q_{AC}. The heat flux to the building envelope is given by the thermal exchange with the inner air and by a fraction A_s of the solar heat gain.

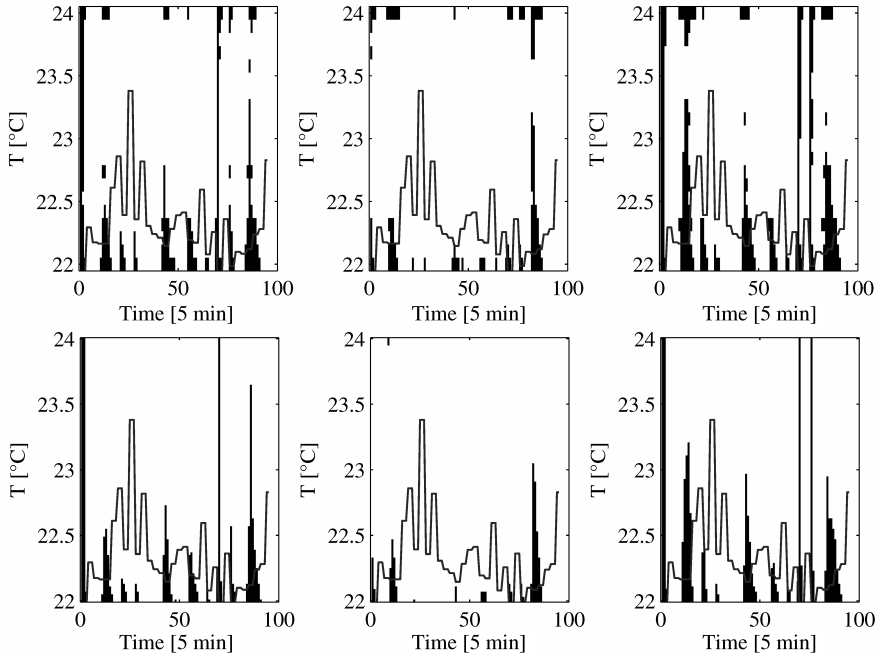


Figure 4: regime policies. Original form (row n. 3 from Figure 3), in the upper row, versus repaired policies, in the lower row.

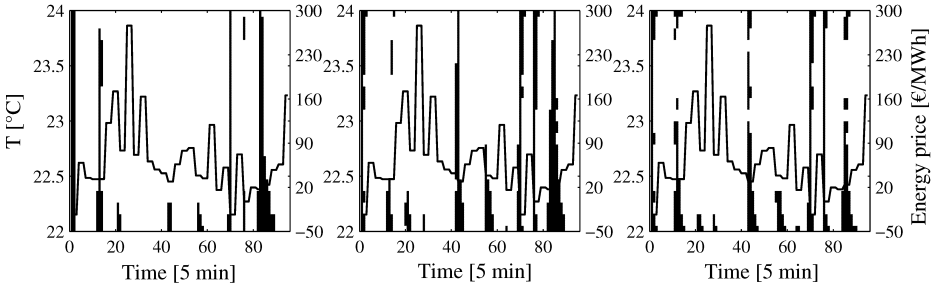


Figure 5: Computation of the closed-loop policies for different shares of virtual tuples over experimental data.

Table A.1: ETP model parameters

Parameter	Value
R_a	110 $°C/kW$
C_a	$2.5E+06$ $kWh/°C$
R_m	2000 $°C/kW$
C_m	$1.2E+07$ $kWh/°C$
A_s	0.5
A_c	1

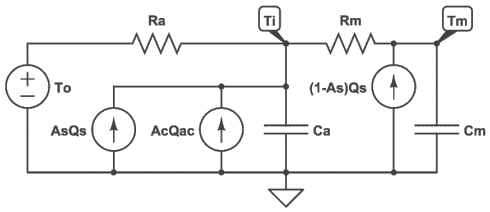


Figure A.8: Sketch of the equivalent thermal parameter model.

Appendix B. Benchmark

An optimal solution of the building heating scheduling can be found using mixed-integer linear programming. The objective

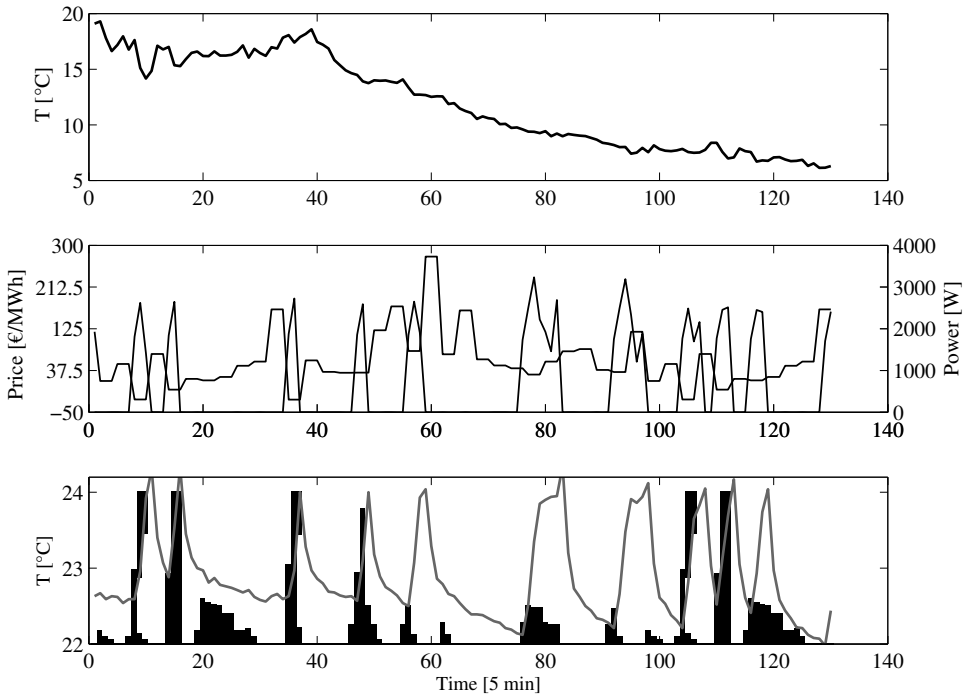


Figure 6: Experimental results 1. Top plot: outside temperature. Middle plot: consumed power by the HVAC and electricity price. Bottom plot: control policy obtained with MABRL and the resulting indoor air temperature.

of the HVAC controller is to minimize its electricity cost using known prices $\boldsymbol{\lambda} \in \mathbb{R}^T$:

$$\mathbf{P}^* = \sum_{t=1}^{T} P \lambda_t u_t^{phys} \Delta t \tag{B.1}$$

subject to:

$$x_{t+1} = f(x_k, u_k, w_k)$$
$$u_t^{phys} = B(x_k, u_k, \theta),$$

where the plant model f is defined by (A.1) and the backup controller B is defined by (13). Notice that the optimal controller knows the plant model, settings of the backup controller θ and future disturbances w. An optimal solution of this optimization problem was found by applying a mixed-integer linear programming solver using Gurobi [56].

References

[1] L. Pérez-Lombard, J. Ortiz, C. Pout, A review on buildings energy consumption information, Energy and buildings 40 (3) (2008) 394–398.

[2] F. Oldewurtel, D. Sturzenegger, M. Morari, Importance of occupancy information for building climate control, Applied Energy 101 (2013) 521–532.

[3] E. F. Camacho, C. B. Alba, Model predictive control, Springer Science & Business Media, 2013.

[4] R. Halvgaard, N. Poulsen, H. Madsen, J. Jorgensen, Economic model predictive control for building climate control in a smart grid, in: IEEE PES Innovative Smart Grid Technologies (ISGT), 2012, pp. 1–6.

[5] F. Oldewurtel, D. Sturzenegger, G. Andersson, M. Morari, R. S. Smith, Towards a standardized building assessment for demand response, in: IEEE 52nd Annual Conference on Decision and Control (CDC), IEEE, 2013, pp. 7083–7088.

[6] F. Oldewurtel, A. Parisio, C. N. Jones, M. Morari, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, K. Wirth, Energy efficient building climate control using stochastic model predictive control and weather predictions, in: American control conference (ACC), 2010, IEEE, 2010, pp. 5100–5105.

[7] J. Ma, J. Qin, T. Salsbury, P. Xu, Demand reduction in building energy systems based on economic model predictive control, Chemical Engineering Science 67 (1) (2012) 92–100.

[8] A. Afram, F. Janabi-Sharifi, Theory and applications of HVAC control systems–a review of model predictive control (MPC), Building and Environment 72 (2014) 343–355.

[9] P. Bacher, H. Madsen, Identifying suitable models for the heat dynamics of buildings, Energy and Buildings 43 (7) (2011) 1511 – 1522.

[10] S. F. Fux, A. Ashouri, M. J. Benz, L. Guzzella, EKF based self-adaptive thermal model for a passive house, Energy and Buildings 68 (2014) 811–817.

[11] M. Morari, C. E. Garcia, D. M. Prett, Model predictive control: theory and practice, in: Workshop on Model Based Process Control, 2014, pp. 1–12.

[12] K. Vanthournout, R. D'hulst, D. Geysen, G. Jacobs, A smart domestic hot water buffer, IEEE Transactions on Smart Grid 3 (4) (2012) 2121–2127.

[13] C. Zhang, D. Yi, N. C. Nordentoft, P. Pinson, J. Østergaard, Flech: A danish market solution for DSO congestion management through der flexibility services, Journal of Modern Power Systems and Clean Energy 2 (2) (2014) 126–133.

[14] M. D. Galus, S. Koch, G. Andersson, Provision of load frequency control

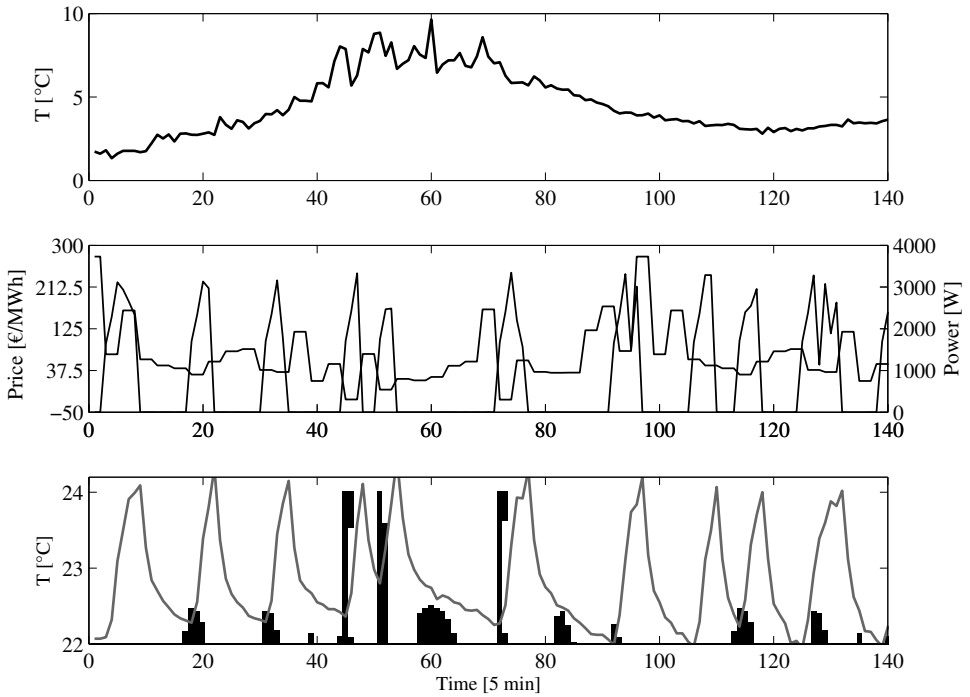


Figure 7: Experimental results 2. Top plot: outside temperature. Middle plot: consumed power by the HVAC and electricity price. Bottom plot: control policy obtained with MABRL and the resulting indoor air temperature.

by phevs, controllable loads, and a cogeneration unit, IEEE Transactions on Industrial Electronics 58 (10) (2011) 4568–4582.

[15] S. Vandael, T. Holvoet, G. Deconinck, S. Kamboj, W. Kempton, A comparison of two GIV mechanisms for providing ancillary services at the university of delaware, in: IEEE International Conference on Smart Grid Communications (SmartGridComm), 2013, pp. 211–216.

[16] J. Mathieu, M. Kamgarpour, J. Lygeros, G. Andersson, D. Callaway, Arbitraging intraday wholesale energy market prices with aggregations of thermostatic loads, IEEE Transactions on Power Systems 30 (2) (2015) 763–772.

[17] J. Širokỳ, F. Oldewurtel, J. Cigler, S. Prívara, Experimental analysis of model predictive control for an energy efficient building heating system, Applied Energy 88 (9) (2011) 3079–3087.

[18] N. Morel, M. Bauer, M. El-Khoury, J. Krauss, Neurobat, a predictive and adaptive heating control system using artificial neural networks, International Journal of Solar Energy 21 (2-3) (2001) 161–201.

[19] S. Liu, G. P. Henze, Experimental analysis of simulated reinforcement learning control for active and passive building thermal storage inventory: Part 1. theoretical foundation, Energy and Buildings 38 (2) (2006) 142 – 147.

[20] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al., Human-level control through deep reinforcement learning, Nature 518 (7540) (2015) 529–533.

[21] D. Ernst, M. Glavic, F. Capitanescu, L. Wehenkel, Reinforcement learning versus model predictive control: a comparison on a power system problem, Cybernetics, IEEE Transactions on Systems, Man, and Cybernetics 39 (2) (2009) 517–529.

[22] L. Busoniu, R. Babuska, B. De Schutter, D. Ernst, Reinforcement learning and dynamic programming using function approximators, Vol. 39, CRC press, 2010.

[23] R. Fonteneau, S. A. Murphy, L. Wehenkel, D. Ernst, Batch mode reinforcement learning based on the synthesis of artificial trajectories, Annals of operations research 208 (1) (2013) 383–416.

[24] T. Lampe, M. Riedmiller, Approximate model-assisted neural fitted Q-iteration, in: International Joint Conference on Neural Networks (IJCNN), IEEE, 2014, pp. 2698–2704.

[25] F. Ruelens, B. Claessens, S. Vandael, B. De Schutter, R. Babuska, R. Belmans, Residential demand response applications using batch reinforcement learning, arXiv preprint arXiv:1504.02125.

[26] E. Atam, L. Helsen, A convex approach to a class of non-convex building HVAC control problems: Illustration by two case studies, Energy and Buildings 93 (0) (2015) 269 – 281.

[27] J. Cigler, D. Gyalistras, J. Širokỳ, V. Tiet, L. Ferkl, Beyond theory: the challenge of implementing model predictive control in buildings, in: Proceedings of 11th Rehva World Congress, Clima, 2013.

[28] L. Ljung, System identification, Springer, 1998.

[29] P. Van Overschee, B. De Moor, A unifying theorem for three subspace system identification algorithms, Automatica 31 (12) (1995) 1853–1864.

[30] F. Oldewurtel, A. Parisio, C. N. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, M. Morari, Use of model predictive control and weather forecasts for energy efficient building climate control, Energy and Buildings 45 (2012) 15–27.

[31] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, P. Haves, Model predictive control for the operation of building cooling systems, IEEE Transactions on Control Systems Technology 20 (3) (2012) 796–803.

[32] D. Bondy, J. Parvizi, Modeling, identification and control for heat dynamics of buildings using robust economic model predictive control, Master's thesis, Masters thesis, Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark (2012).

[33] F. Sossan, X. Han, H. Bindner, Dynamic behaviour of a population of controlled-by-price demand side resources, in: PES General Meeting— Conference & Exposition, 2014 IEEE, IEEE, 2014, pp. 1–5.

[34] G. Costanzo, F. Sossan, M. Marinelli, P. Bacher, H. Madsen, Grey-box modeling for system identification of household refrigerators: A step toward smart appliances, in: 2013 4th International Youth Conference on

Energy (IYCE), 2013, pp. 1–5. doi:10.1109/IYCE.2013.6604197.
[35] A. M. Kosek, G. T. Costanzo, H. W. Bindner, O. Gehrke, An overview of demand side management control schemes for buildings in smart grids, in: Smart Energy Grid Engineering (SEGE), 2013 IEEE International Conference on, IEEE, 2013, pp. 1–9.
[36] S. Koch, M. Zima, G. Andersson, Active coordination of thermal household appliances for load management purposes, in: IFAC Symposium on Power Plants and Power Systems Control, Citeseer, 2009.
[37] A. H. Neto, F. A. S. Fiorelli, Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption, Energy and Buildings 40 (12) (2008) 2169–2176.
[38] R. Bellman, Dynamic Programming, Dover Publications, Incorporated., NY, 2003.
[39] M. Deisenroth, C. E. Rasmussen, J. Peters, Model-based reinforcement learning with continuous states and actions.
[40] A. G. Barto, Reinforcement learning: An introduction, MIT press, 1998.
[41] Z. Wen, D. O Neill, H. Maei, Optimal demand response using device-based reinforcement learning, IEEE Transactions on Smart Grid PP (99) (2015) 1–1. doi:10.1109/TSG.2015.2396993.
[42] M. Gonzalez Vaya, L. Rosello, G. Andersson, Optimal bidding of plug-in electric vehicles in a market-based control setup, in: Power Systems Computation Conference (PSCC), 2014, 2014, pp. 1–8. doi:10.1109/PSCC.2014.7038108.
[43] F. Ruelens, B. Claessens, S. Vandael, S. Iacovella, P. Vingerhoets, R. Belmans, Demand response of a heterogeneous cluster of electric water heaters using batch reinforcement learning, in: Power Systems Computation Conference (PSCC), 2014, pp. 1–7. doi:10.1109/PSCC.2014.7038106.
[44] S. Vandael, B. Claessens, D. Ernst, T. Holvoet, G. Deconinck, Reinforcement learning of heuristic ev fleet charging in a day-ahead electricity market, IEEE Transactions on Smart Grid PP (99) (2015) 1–1. doi:10.1109/TSG.2015.2393059.
[45] A. Faruqui, S. Sergici, Household response to dynamic pricing of electricity: a survey of 15 experiments, Journal of Regulatory Economics 38 (2) (2010) 193–225.
[46] D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, D. P. Bertsekas, Dynamic programming and optimal control, Vol. 1, Athena Scientific Belmont, MA, 1995.
[47] D. P. Bertsekas, J. N. Tsitsiklis, Neuro-dynamic programming (optimization and neural computation series, 3), Athena Scientific 7 (1996) 15–23.
[48] M. Scholz, M. Fraunholz, J. Selbig, Nonlinear principal component analysis: neural network models and applications, in: Principal manifolds for data visualization and dimension reduction, Springer, 2008, pp. 44–67.
[49] D. Ernst, P. Geurts, L. Wehenkel, Tree-based batch mode reinforcement learning (2005) 503–556.
[50] Z.-H. Zhou, Ensemble methods: foundations and algorithms, CRC Press, 2012.
[51] G.-B. Huang, Q.-Y. Zhu, C.-K. Siew, Extreme learning machine: theory and applications, Neurocomputing 70 (1) (2006) 489–501.
[52] E. Cambria, G.-B. Huang, L. L. C. Kasun, H. Zhou, C.-M. Vong, J. Lin, J. Yin, Z. Cai, Q. Liu, K. Li, et al., Extreme learning machines, IEEE Intelligent Systems 28 (6) (2013) 30–59.
[53] P. Geurts, D. Ernst, L. Wehenkel, Extremely randomized trees, Machine learning 63 (1) (2006) 3–42.
[54] W. Zhang, K. Kalsi, J. Fuller, M. Elizondo, D. Chassin, Aggregate model for heterogeneous thermostatically controlled loads with demand response, in: IEEE Power and Energy Society General Meeting, 2012, pp. 1–8.
[55] `www.elia.be`.
[56] Gurobi Optimization, Inc. Gurobi optimizer reference manual, `www.gurobi.com`, [Online: accessed 11-July-2015].

Electric Space Heating Scheduling for Real-time Explicit Power Control in Active Distribution Networks

Giuseppe Tommaso Costanzo*, Andrey Bernstein†, Lorenzo Reyes Chamorro‡,
Henrik W. Bindner*, Jean-Yves Leboudec† and Mario Paolone‡

*Center for Electric Energy and Power, Technical University of Denmark, Campus Risø
Frederiksborgvej 399, 4000 Roskilde, Denmark. email:{guco, hwbi}@elektro.dtu.dk.
†Laboratory for Communications and Applications (LCA), Ecole Polytechnique Federale de Lausanne,
Route Cantonale, 1015 Lausanne, Switzerland. email:{andrey.bernstein, jean-yves.leboudec}@epfl.ch.
‡Distributed Energy Systems Laboratory (DESL), Ecole Polytechnique Federale de Lausanne,
Route Cantonale, 1015 Lausanne, Switzerland. email:{lorenzo.reyes, mario.paolone}@epfl.ch.

***Abstract*—This paper presents a systematic approach for abstracting the flexibility of a building space heating system and using it within a composable framework for real-time explicit power control of microgrids and, more in general, active distribution networks. In particular, the proposed approach is developed within the context of a previously defined microgrid control framework, called COMMELEC, conceived for the explicit and real-time control of these specific networks. The designed control algorithm is totally independent from the need of a building model and allows exploiting the intrinsic thermal inertia for real-time control. The paper first discusses the general approach, then it proves its validity via dedicated simulations performed on specific case study composed by the CIGRE LV microgrid benchmark proposed by the Cigré TF C6.04.02.**

***Index Terms*—Demand-side Management, Explitic power control, Microgrids.**

I. Introduction

Today's trend of vast penetration of distributed generation in low and medium voltage power networks threats the energy/power equilibrium as well as the quality-of-supply of these electrical grids in a way that, in several countries, operational constraints are already attained. Additionally, their active contribution to the operation of transmission networks needs to be re-defined. In particular, the main concerns of power transmission networks operators refer to the definition of new control schemes able to evaluate, and schedule, the contributions of these active networks to the grid ancillary services (with particular reference to voltage and frequency controls, e.g. [1]). It this respect, a fully-composable framework for explicit and real-time control of network power flows has been introduced in [2], where real-time means with a typical response time of 0.1 sec. It is based on three main elements: (i) an abstract framework that applies to electrical subsystems and specifies their capabilities, expected behaviour and a simplified view of their internal state; (ii) the agents responsible for subsystems/resources capable to communicate with other agents by using a simple, yet powerful, protocol and (iii) the recursive property, namely it is possible to compose a set of interconnected elements into a simple entity that responds to protocol messages (see [2] for further details). This framework allows for abstracting devices status and consumers or producers operation capabilities regardless of their type and operation mode. Device abstraction and control is performed by the local resource controller, which exposes a standard interface toward the grid controller. The grid controller is totally generic and needs information only on the grid topology and its status, as well as of the devices underneath, to compute the power setpoints for all the resources, thus guarantee the network safe operation. Such framework, called COMMELEC [2], is based on request/response communication protocol between the devices controllers (Resource Agents, RAs) and the grid controller (Grid Agent, GA). The RAs export generic and standardized information about devices status to the GA that computes, in real-time, the optimal power injections at the grid nodes that the various RAs are requested to realize. Thus, the design of a Resource Agent is specific to device type, whereas the GA is totally generic and can be adapted to any network.

Within the context of the COMMELEC framework, there is the interest in investigating whether the flexibility related to this type of load can be exploited. [2] has already presented the design of RAs for induction motors, PV plants and micro hydro turbines, while the building consumption has been treated as a non-controllable load. This paper goes further: it proposes a systematic approach for abstracting and representing the intrinsic flexibility related to building thermal inertia within the COMMELEC framework. The design of a RA for electric space heating system, called Load Agent (LA), is here presented. It uses the measured air temperature in different building sectors to compute the system flexibility and the other information sent to the GA.

The structure of the paper is the following: Section II summarizes the information exchange protocol and the COMMELEC framework by making use of a case study composed by the Cigré LV microgrid benchmark of the TF C6.04.02 [3]. In this first case study, the building space heating is a non-controllable load. The simulation reported in this section shows the positive impact on the network performances that COMMELEC has with respect to the standard droop-based

control approaches. From this starting point, the design of the Load Agent for electric space heating is fully explored in Section III. A simulation study using the building as flexible load is presented in Sec. IV. The simulation results quantify the positive impact of the proposed LA in improving the microgrid performances while increasing the production from renewables, all without any modification of the GA that does not need to be aware of the load under control of the LA. Sec. V outlines the final conclusions and the next research steps.

II. The COMMELEC Framework

In this section, we briefly discuss the COMMELEC framework for real-time control of electrical grids using explicit power setpoints introduced in [2]. This framework uses a hierarchical representation of *software agents*, each associated with a single device or an entire subsystem (including a grid and/or a number of devices). Thanks to a standardized interface between the agents, it is possible to completely disregard devices details in the overall network control and easily upscale (or downscale) the entire architecture. Each agent speaks for and controls the subsystem under its responsibility, which can consist of other electrical grids or resources (loads, generators and storage devices), by using explicit power setpoints.

An agent can be assigned the role of *leader* of one or more other agents, which we term the *followers* of that leader. The roles follow the topological/voltage hierarchy of power networks. An example of agents hierarchy is shown in Fig. 1, where there are two Grid Agents: one for the MV grid (called MVGA), and one for the LV grid (called LVGA). The LVGA is the leader of a group of Resource Agents (RAs).

The considered network model is mainly taken from [3], where a $0.4kV$ benchmark microgrid is proposed. In [2] we have assumed to connect the low voltage microgrid with a simplified 20kV medium voltage grid. The whole network is considered islanded. As typically used in a microgrid setup, we assume that power electronic devices are used to interface the generation/storage units with the grid and, as presented in [2], the distributed generation is composed by (i) photovoltaic plants (PVi) and a hydraulic microturbine (μH), (ii) two battery-based energy storage systems (ESS and ESS1), (iii) uncontrollable loads (ULi) and (iv) controllable loads modelled as water boilers (WBi) capable of deploying explicit power control setpoints (in the following, called just setpoints).

The agents communicate with each other by using a simple Advertisement/Request protocol, and using some simplified quantitative information about their devices capabilities and internal states. In particular, the follower agents periodically advertise to their leader the following three elements: (i) *PQt profile*, that is a region in PQt space of setpoints that the subsystem is willing to implement; (ii) *Virtual cost function*, that acts as proxy for the internal state of the system, the preferred setpoints, and how close the subsystem is to its operational constraints; and (iii) *Belief function*, representing all the possible values of (P, Q) that the subsystem *may implement in practice* when it receives a certain setpoint. Note that the latter is a *set-valued function* from $\mathbb{R}^2$ to sets over $\mathbb{R}^2$. Observe the difference between PQt profiles and belief functions: the former indicates the setpoints that this subsystem is willing to receive, whereas the latter indicate all the possible operating conditions that may result from applying a received setpoint. The PQt profile is used by the grid agent to compute an "optimal" setpoint, whereas the belief function is used to determine the region of the safe operation of the grid.

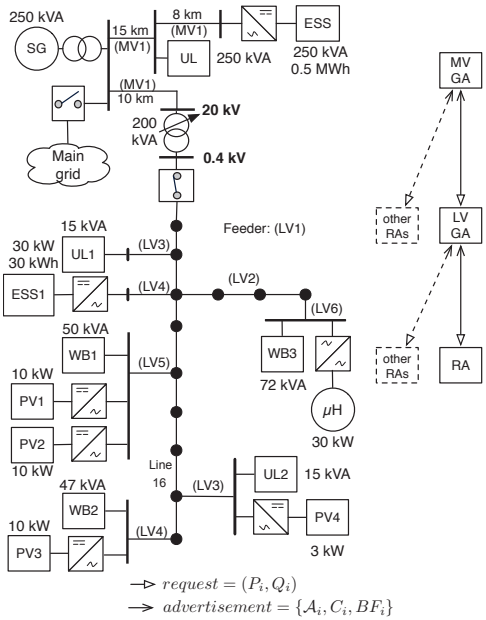


Fig. 1: Reference distribution network. The LV part of the grid is composed by the CIGRE LV microgrid benchmark proposed by TF C6.04.02 [3]. The MVGA is the leader of one or more LVGA, which themselves are leaders of one or more RAs (dotted arrows).

All agents monitor in real time the state of the grid device(s) and/or the agents (the GAs) under their responsibility and compute the general messages to send to their leader agent. When receiving a new request, each agent projects the power setpoint to the current updated PQt boundaries in order to ensure its individual safety. We refer the reader to [2] for further details on the protocol and the computation of the setpoints of the GA.

In the following we show the operation of COMMELEC and its positive impact in a $0.4kV$ LV microgrid in terms of: (i) minimization of renewables curtailment, (ii) local power balance, (iii) automatic management of the line congestions and voltage control in a case study where three different control methods are compared: COMMELEC, droop with only primary frequency control (DP) for all resources and primary voltage control only in the slack resource (in our case the ESS), and Droop with additional secondary frequency control

(DPS) at the slack resource. The focus here is on the dynamic short-term behaviour, so that results are presented over a time horizon of 1600 seconds.

As it can be seen from Fig. 2, COMMELEC allows higher energy production from PV with respect to both droop control strategies. Figure 3b shows COMMELEC performance in

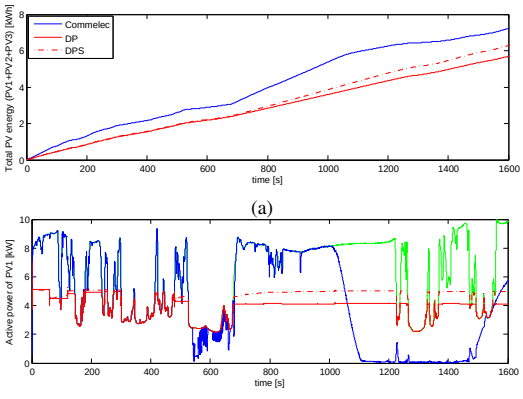


Fig. 2: Comparison between COMMELEC and Doop control strategies: total PV production (a) and detailed production of PV1 (b) (the green line is the ideal production).

controlling the local water boiler WB2 for compensating the local PV3 production, while the line current and node voltage are kept within the safe bounds (Fig. 3d). This case study shows some of the benefits introduced by COMMELEC, as the maximization of production from renewables and local power balance at LV level. Observe that in the network topology shown in Fig. 1 there is an uncontrollable load of 15kVA (UL2) connected at the end of the feeder. In the following study we replace this load with a 8kW residential load consisting of building space heating. We next focus on the design of the Load Agent for this resource, while the grid configuration, the GA and the other agents remain unchanged.

III. The Load Agent

The algorithm, which design is here presented, is the core of the building Resource Agent, henceforth called Load Agent (LA), which generality and model-independent design this section is intended to enlighten. The presented LA can be adapted to buildings of any size and type provided that the heating system is electric (resistive heaters) and air temperature measurements are available in each control volume. Note that, since this study is based on simulations, the LA communicates with a building emulator (Fig. 4), which model is detailed in [4].

As introduced in Sec. II, the information the LVGA sends to RAs is an explicit power setpoint. RAs, as acknowledgement, send to the GA the *PQt* profile, the *belief function* and the *virtual cost function*, which calculation is presented in the following three subsections. The fourth subsection presents

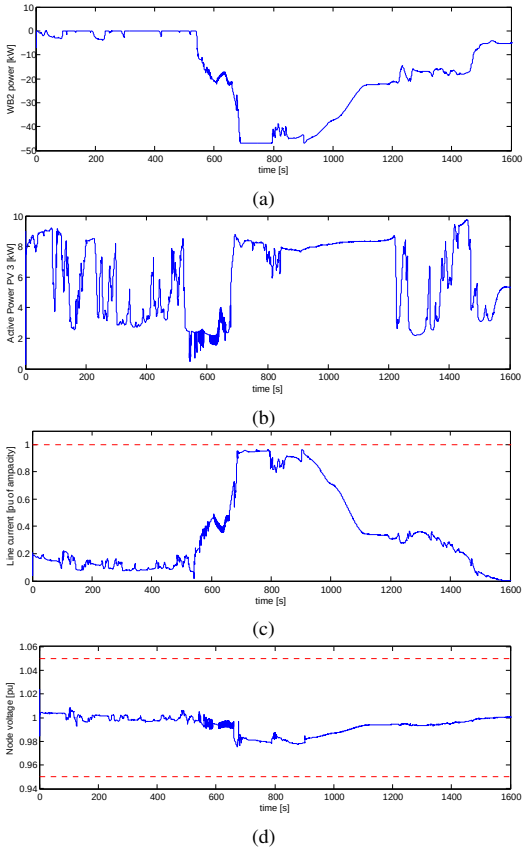


Fig. 3: COMMELEC control strategy: WB2 power consumption (a), PV3 production (b), line current (c) and node voltage (d).

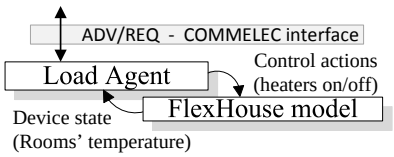


Fig. 4: Interaction between the Load Agent and the building simulator.

the control algorithm used in the LA to implement the power setpoint requested by the GA.

A. *PQt profile*

The Load Agent advertises to the GA the region in the (P, Q) plane of the load admissible active and reactive power setpoints at time t, where *negative power means consumption*. Being $\mathcal{X}$ the set of all *possible* power setpoints $x = (P, Q) \in \mathbb{R}^2$, the PQt profile sent at time t_0 is the collection: $\mathcal{A}(t) \subseteq \mathcal{X}$, where t is the time instant at which the PQt profile is advertised. Since in this study the controlled system consists of resistive space heaters, only pure active power consumption is considered in a way that $\mathcal{A}$ is in this case an interval in P and zero in Q.

Even if the notion of the load flexibility is associated to both the PQt profile and the cost function, we here link it to the former. Namely, the load is *flexible* if it can be interrupted and re-started without causing user discomfort or device failure. Specifically, a single room flexibility is determined by the fact that the heater can be switched on or off without compromising the user comfort as long as the room temperature is within the comfort bounds. Conversely, if the room temperature is below the comfort bound (T_{LB}) the heater *must* operate and if the temperature is above (T_{UB}) the heater *must not* operate (Fig. 5).

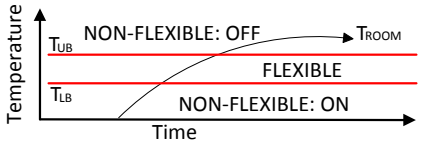


Fig. 5: The flexibility of operation of a heating system is related to the system state itself.

The flexibility of the whole system (building) is determined by its subsystems (rooms) flexibilities. Let $\mathcal{R}$ denote the set of rooms in the building and $j \in \mathcal{R}$ the room index. Define $\mathcal{F}_j$ as the flexibility of room j in terms of upper and lower active power consumption:

$$\mathcal{F}_j = \begin{cases} [P_j^*, P_j^*] & : \quad T_j < T_{LB,j} \\ [P_j^*, 0] & : \quad T_{LB,j} \leq T_j \leq T_{UB,j} \\ [0, 0] & : \quad T_j > T_{UB,j} \end{cases} , \qquad (1)$$

where P_j^* is the rated active powers of the heating system of room j.

The building *PQt* profile is obtained by performing the Minkowski sum of $\mathcal{F}_j$ over $\mathcal{R}$:

$$\mathcal{A} = \sum_{j \in \mathcal{R}} \mathcal{F}_j \, . \qquad (2)$$

The rooms whose temperature is below T_{LB} determine the building minimum power consumption, while the maximum consumption is limited by those rooms whose temperature is above T_{UB}. The (P, Q) flexible region is determined by those rooms which temperature is within the comfort requirements $[T_{LB}, T_{UB}]$.

Figure 6 presents the computation of the flexibility region based on rooms temperatures in case of a linear variation of external temperature. It is possible to observe the effect of this disturbance on the PQt profile.

The example provided in Fig. 6 aims at showing how the upper and lower bounds of set $\mathcal{A}$ evolve with respect to the system state and how external disturbances impact the system flexibility. It can be seen that the bounds of $\mathcal{A}$ coincide as soon as the rooms are outside the comfort zone (from time 0 to 0.2 hours) and load consumption constrained to $8kW$. Analogous situation occurs when, due to an external disturbance in the outside temperature, the rooms' temperature rises above T_{UB}, causing the load consumption to be constrained to zero.

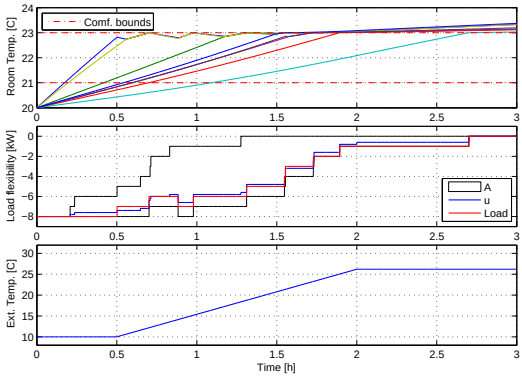


Fig. 6: Evolution of rooms' temperature (first sub-plot), *PQt* profile computation (second sub-plot) and external temperature (third sub-plot).

In the same figure, the blue line is the power setpoint from the GA and the red line is the load consumption. The LA implements the requested power setpoint by means of Laod Admission Control (LAC), which details are deferred to Sec. III-D.

B. Belief function

The *belief function*, henceforth denoted by B, describes the ability of the Load Agent to map a requested power setpoint, u, coming from the Grid Agent, to actual load consumption. Note that B is a set-valued function from $\mathbb{R}^2$ to sets over $\mathbb{R}^2$ and is designed to account for model uncertainties and quantization effects in the load.

In this study, the building is composed by eight rooms with one electric space heater each. The heaters can be either switched on or off, therefore the load can implement only discrete active power setpoints from $0kW$ to $8kW$ in steps of $1kW$. Hence the *belief function* simplifies to the *nearest integer* function:

$$B\left(u\right) = \left\{nint\left(u\right)\right\} , \qquad (3)$$

where u is a power setpoint received from the GA.

Note that, thanks to the LAC algorithm (Sec. III-D), the LA achieves tracking a power setpoint from the GA with an accuracy of $\pm 0.5kW$.

C. Virtual cost function

The cost function serves as a proxy for load internal conditions and it associates a cost to every operating point $x \in \mathcal{A}$ and it is a map: $C : \mathcal{A} \mapsto \mathbb{R}$, that is formally defined as:

$$C\left(x\right) = \sum_{j \in \mathcal{R}} \left[U\left(j\right) c_h\left(T_j\right) - \left(1 - U\left(j\right)\right) c_h\left(T_j\right)\right] , \qquad (4)$$

where the terms $U(j)$ are the components of the heaters activation vector $\overrightarrow{U}$. This vector, given a certain state of the system, is computed with respect to a specific operating point x by means of LAC (Sec III-D): $\overrightarrow{U} = LAC\ (\mathcal{M}, x)$.

In Eq. (4) the terms $c_h\left(T_j\right)$ are computed as:

$$c_h\left(T_j\right): \begin{cases} 1+\left(T_j-T_{UB_j}\right)^2 & : T_j \geq T_{UB_j} \\ \dfrac{T_j-T_{M_j}}{T_{UB_j}-T_{M_j}} & : T_{LB_j} < T_j < T_{UB_j} \\ -1-\left(T_{LB}-T_j\right)^2 & : T_j \leq T_{LB_j} \end{cases} \quad (5)$$

and they express the cost associated to heating a room j; T_j is the room internal temperature, T_{UB_j} and T_{LB_j} are the comfort upper and lower bounds and $T_{M_j}=\left(T_{UB_j}-T_{LB_j}\right)/2$.

Equation (5) returns zero when the room temperature is in the middle of the comfort band, T_M, positive when it is above T_M and negative when it is below. This design choice follows the fact that the building exhibits the maximum flexibility for longer time if all the rooms temperatures are in the middle of their comfort zone. The cost is linearly defined within the comfort zone in order to not excessively penalize deviations from T_M and it is always related to the action of *heating* ($U\left(j\right)=1$). It follows that the action of *not heating* ($U\left(j\right)=0$) generates opposite cost in Eq. (4). Therefore the total cost expressed in Eq. (4) is obtained by summing up the costs related to each room control action. Figure 7 shows examples of virtual costs in correspondence of different times

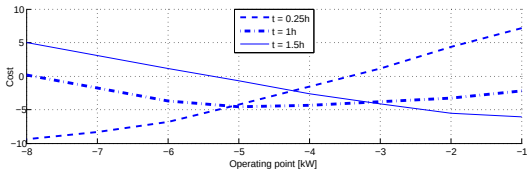


Fig. 7: Example of cost function computed for different system states ($t=0.25h$, $t=1h$, $t=1.5h$)

Note that in this case study only active power operating points are considered and the LA does not need any information about the building model to compute the cost. Every time the Load Agent computes the cost function, it uses the information on load state (temperatures) to compute the terms c_h and it calls the LAC algorithm, presented in the next subsection, which uses the same temperature and comfort information to compute the $U\left(j\right)$.

D. Load Admission Control (LAC) algorithm

This subsection presents the algorithm for load Direct Power Control (DPC) via Load Admission Control (LAC). DPC is introduced in [5] and consists of run-time control of a DER for the tracking of a power setpoint. It differs from the Power Schedule by the fact that future values of the power setpoint are not known to the DER. The interested reader can refer to [6] and the references therein for an overview on control policies for energy consumption in buildings.

Algorithm 1 presents the Load Admission Control, an heuristic-based search algorithm for load online scheduling that was first introduced in [7] for demand response applications and provision of *PowerMax* service [5] . However, the algorithm presented in [7] is here modified in order to track a power reference instead of limiting the peak power.

Define the triple γ_j associated to room j as:

$$\gamma_j=\left\langle h_j, P_j^*, p_j\right\rangle \ , \quad (6)$$

where P_j^* is the heater rated active power consumption, h_j is a priority factor defined as:

$$h_j=\begin{cases} 1 & : T_j \leq T_{LB_j} \\ \dfrac{T_j-T_{LB,j}}{T_{UB_j}-T_{LB,j}} & : T_{LB_j} < T_j < T_{UB_j} \\ 0 & : T_j \geq T_{UB_j} \end{cases} \ , \quad (7)$$

and p_j is boolean value representing the load preemptiveness (interruptibility). A device is interruptible if its operation state can be modified from ON to OFF. Heaters in rooms which temperature is below T_{LB} are non-preemptible.

Define $(\mathcal{M},\leq)$ a totally ordered set of pending consumption requests:

$$\mathcal{M}=\bigcup_{j\in\mathcal{R}}\gamma_j \ , \quad (8)$$

where $\gamma_k \leq \gamma_j$, $\forall k>j$ and $(\leq)$ is a *Lexicographical order*.

Given a power setpoint from the GA, u, the control vector $\overrightarrow{U}$ for the heaters is computed via LAC: $\overrightarrow{U}=LAC\left(\mathcal{M},u\right)$, where $LAC\left(\cdot\right)$ is detailed in Algorithm 1.

Algorithm 1 Load Admission Control for DLC.

Variables:

- u: active power setpoint from GA ($u\in\mathcal{A}$)
- $\mathcal{M}$: set of pending power requests
- $\mathcal{L}$: set of accepted power requests
- ρ: cumulative power of accepted requests
- $\overrightarrow{U}$: control decision vector

Require:

Initialize: $\mathcal{L}=\emptyset$, $\rho=0$, $U=O$
for all $\gamma_j\in\mathcal{M}$ **do**
  if $p_j=$ false **then**
    $\mathcal{M}=\mathcal{M}\setminus\gamma_j$, $\mathcal{L}=\mathcal{L}\cup\gamma_j$, $\rho=\rho+P_j^*$
  end if
end for
for all $\gamma_j\in\mathcal{M}$ **do**
  if $|u-\rho-P_j^*|<=|u-\rho|$ **then**
    $\mathcal{M}=\mathcal{M}\setminus\gamma_j$, $\mathcal{L}=\mathcal{L}\cup\gamma_j$, $\rho=\rho+P_j^*$
  end if
end for
set $U\left(j\right)=1\ \forall\gamma_j\in\mathcal{L}$
return $\overrightarrow{U}$

IV. Simulation

This section presents two simulations studies that show the building response to signals from the LVGA when: i) six rooms out of eight are initialized below the comfort zone (Sec. IV-A), ii) all the rooms are initialized within the comfort zone (Sec. IV-B). Note that the LVGA here was not modified with respect to the simulation presented in Sec. II.

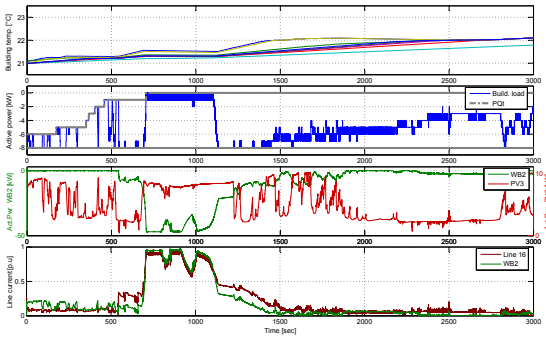


Fig. 8: Simulation with LA - rooms initialized below the comfort zone. From the top chart: temperature of building rooms, building power consumption, WB2 consumption and PV3 production, building and connecting line (Line 16) currents.

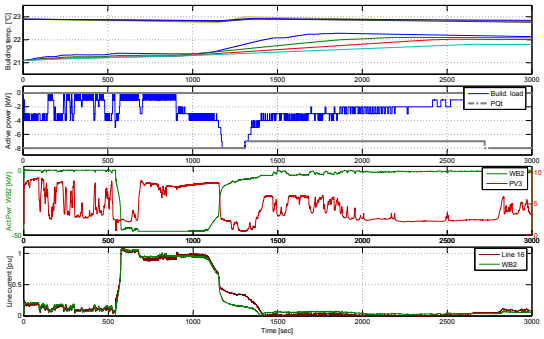


Fig. 9: Simulation with LA - rooms initialized within the comfort zone. From the top chart: temperature of building rooms, building power consumption, WB2 consumption and PV3 production, building and connecting line (Line 16).

A. *First scenario: load initialized below the comfort zone*

In this simulation, the PQt profile of the LA is initially between $-6kW$ and $-8kW$, as six rooms out of eight are below the comfort zone (Fig. 8). The LA is able to track the explicit power setpoint coming from the LVGA and the building flexibility increases as soon as all the rooms enter the comfort zone (upper chart). At this time, the LVGA fully exploits the load flexibility when dealing with local resource contingencies and line constraints.

One can observe that, given a combination of such events as: increased PV production, increased building flexibility and pending WB2 operation, the LVGA allows the WB2 to charge (by increasing the WB2 power setpoint) and limits the building consumption (by decreasing the LA power setpoint) in order to avoid the congestion of the line connecting the water boiler, the building and the two PV installations to the distribution feeder (Line 16 in Fig. 1). As intrinsic property of composability of the system, this behaviour emerges without the need of any predefined control policy or manual intervention.

The building is used as a virtual electrical storage resource by the LVGA to optimize the operation of the microgrid, as the LA is requested to partially follow the local PV production (PV2 receives the same irradiation of PV3). At the same time, as the rooms approach the midpoint of the comfort zone, due to the virtual cost function, the LVGA reduces the heating input (Fig. 8, last part of the simulation).

B. *Second scenario: load initialized within the comfort zone*

Due to the initialization of the internal temperature in all the rooms within the comfort zone, the PQt profile is constrained between $0kW$ and $-8kW$ (Fig. 9), in a way that the building exhibits the maximal flexibility to the LVGA from the beginning.

Figure 9 shows that only the rooms in the lower part of the comfort zone are heated, while the LVGA operates the water boiler WB2 in advance with respect to the previous simulation. When WB2 is switched on, the LA is requested to reduce the building consumption in order to not overload the feeder. At about 1200s, the decreased PV production brings the LVGA to reduce the WB2 charging power and restore the building heating so that all the rooms steer toward the middle of the comfort zone (i.e. to the minimum virtual cost).

V. Conclusions and Future Work

This paper presented a systematic approach to abstract the flexibility of a building heating system and interface it with the composable framework for real-time grid control, COMMELEC, i.e. to be able to use a building's thermal inertia for real-time control.

The presented simulations show the efficacy of the LA to operate the load within the user comfort requirements and communicate generic information to the GA to perform explicit power control. The LA designed for the building space heating system is independent from the building model and it was plugged in the COMMELEC framework and tested in the same setup presented in [2]. In this context, the next research steps foresee the implementation of the presented LA in a real-time control platform and test with real loads in a microgrid setup currently developed at EPFL.

References

[1] K. A. Papadogiannis and N. Hatziargyriou, "Optimal allocation of primary reserve services in energy markets," *Power Systems, IEEE Transactions on*, vol. 19, no. 1, pp. 652–659, Feb 2004.

[2] A. Bernstein, L. Reyes-Chamorro, J.-Y. Le Boudec, and M. Paolone, "A composable method for real-time control of active distribution networks with explicit power setpoints," 2014, arXiv:1403.2407.

[3] S. Papathanassiou, N. Hatziargyriou, K. Strunz *et al.*, "A benchmark low voltage microgrid network," in *Proceedings of the CIGRE Symposium: Power Systems with Dispersed Generation*, 2005, pp. 1–8.

[4] D. E. M. Bondy and J. Parvizi, "Modeling, identification and control for heat dynamics of buildings using robust economic model predictive control," Master's thesis, Technical University of Denmark, 2012.

[5] O. Gehrke and F. R. Isleifsson, "An aggregation friendly information model for demand side resources," in *2010 IEEE 35th Conference on Local Computer Networks*, pp. 1019–1023.

[6] A. Kosek, G. Costanzo, H. Bindner, and O. Gehrke, "An overview of demand side management control schemes for buildings in smart grids," in *Smart Energy Grid Engineering (SEGE), 2013 IEEE International Conference on*, Aug 2013, pp. 1–9.

[7] G. Costanzo, J. Kheir, and G. Zhu, "Peak-load shaving in smart homes via online scheduling," in *Industrial Electronics (ISIE), 2011 IEEE International Symposium on*, June 2011, pp. 1347–1352.

APPENDIX B

Insights

B.1 Overview on system identification via Maximum Likelihood Estimation

Mathematical modelling is a crucial topic in many disciplines of engineering and science. It is a procedure that, by trading off complexity with accuracy, strives at finding a tractable, but rather accurate, mathematical description of a system behaviour or a physical phenomenon. The purposes are diverse, at glance we find: simulation, analysis, and control systems design. System identification consists in applying numerical methods to find the best values for the model parameters according to the observed system behaviour.

Maximum Likelihood Estimation is a procedure leads to the set of parameters that maximises the probability of obtaining the observed data under a chosen parametric model. Assuming that the observations are independent and identically-distributed, their joint probability function is:

$$f\left(x_1 \dots x_n | \, \theta\right) = \prod_{i=1}^{n} f\left(x_i | \, \theta\right) \, ,$$

where $f\left(\cdot \, | \, \theta\right)$ is a parametric model that generates the data. This model is unknown, but it belongs to a family of distributions $\{f\left(\cdot \, | \, \theta\right), \, \theta \in \Omega\}$, where θ is a vector of parameters for this family. Denoting θ^* the *true* parameters vector, we have that $f^* = f\left(\cdot \, | \, \theta^*\right)$. The aim is to obtain an estimator $\hat{\theta}$ that is as close as possible to θ_0.

Let $\mathcal{L}$ denote the *likelihood function*, with a fixed set of *parameters* x_i, $i \in \{1 \dots n\}$, and the function variable θ. If $\mathcal{L}$ is defined as:

$$\mathcal{L}\left(\theta; x\right) = f\left(x_1 \dots x_n | \, \theta\right) \, ,$$

maximising $\mathcal{L}$ is equivalent of maximising the probability of obtaining the observed data, given a set of parameters θ, under the parametric model $f\left(\cdot \, | \, \theta\right)$. In practise it is common to maximise the logarithm of $\mathcal{L}$, since $ln\left(\cdot\right)$ is a monotonic-increasing function.

Denoting the *log-likelihood* as:

$$\ell\left(\theta; x\right) = \ln \mathcal{L}\left(\theta; x\right) \, ,$$

the method of MLE finds an estimate for θ^*, called $\hat{\theta}$, that maximises $\ell(\hat{\theta}; x)$:

$$\hat{\theta} \subseteq \left\{ \arg\max_{\theta \in \Theta} \ell(\theta; x) \right\} .$$

MLE is based on the assumption that the model parameters are unknown variables with Gaussian distribution, therefore their estimation is provided in terms of mean and variance. This is particularly useful for prediction in terms of expected system behaviour and its variance.

For a deeper insight of MLE, the interested reader can refer to [Mad08], and its application to identification of grey-box models in [KMJ04]. The tool used in this research for parameters estimation is Continuous-Time Stochastic Modeling (CTSM-R), an open source toolbox for R developed by DTU Compute at Technical University of Denmark [Dan13]. [REF: CTSM1 and CTSM2] present a closer insight on the CTSM operation and mathematics.

B.2 A gentle introduction to Model Predictive Control

A predictive controller has an internal model that is used to predict the behaviour of the process, starting from the current time, over a future prediction horizon. Given a system described by:

$$\frac{dx}{dt} = f(x, u) , \tag{B.1}$$

where x is the process state, u is the control input, and f is the process model (often non-linear), a predictive controller strives at minimising a function of the form:

$$V(x, u, t) = \int_0^T \ell(x(t), u(t), t)\, dt + F(x(T)) , \tag{B.2}$$

in which ℓ is function of the the system state x, the control input u, and the time t, and F is function of the final state of the system. In general terms, every optimal control problem can be expressed in this form and be solved by solving the Hamilton-Jacobi-Bellmann equation:

$$\frac{\partial}{\partial t} V^0(x, t) = \min_{u \in U} H\left(x, u, \frac{\partial}{\partial x} V^0(x, t)\right) , \tag{B.3}$$

where $H(x, u, \lambda) = \ell(x, u) + \lambda f(x, u)$, with the boundary condition $V^0(x, T) = F(x)$ [Mac02]. The optimal solution is therefore the following *feedback* control law, which depends on actual time and actual state:

$$u^0(x, t) = \arg\min_{u \in U} H\left(x, u, \frac{\partial}{\partial x} V^0(x, t)\right) . \tag{B.4}$$

Unfortunately, no analytical solution has yet been found for u^0, except for a restricted class of problems, such as fuel minimisation for trip planning [Mac02]. However, under the linearity assumption, it is possible to reduce problem B.4 from an optimisation over a set of functions to an optimisation over a finite set of parameters, which are indeed the values of the control input over a finite time horizon.

Let's start with considering the discrete time model of the system for which we are interested in designing the MPC controller:

$$S : \begin{cases} x_{k+1} = A_d x_k + B_d u_k + E_d d_k \\ y_k = C_d x_k + D_d u_k \end{cases} , \tag{B.5}$$

where k is the current time step, x is the system state space, y is the system output, A_d, B_d E_d C_d D_d are the state space matrices, u is the control input, and d is the disturbance vector. The cost function of a predictive controller can be designed to pursue different objectives. For instance we may be interested in following a reference signal, limiting the control effort and the cost of control action:

$$V(k) = \sum_{j=k}^{k+N} \left\| y_{k+j|k} - r_{k+j|k} \right\|^2_{Q(j)} + \sum_{j=0}^{N-1} \left\| \Delta u_{k+j|k} \right\|^2_{R(j)} + \sum_{j=0}^{N-1} \rho_{k+j} u_{k+j|k} \ . \tag{B.6}$$

In B.6 N is the prediction horizon and k is the current time step, at which the MPC optimisation is performed. The first term penalises the deviation of the system output, y, from a given reference, r; this is used for output reference tracking (i.e temperature in a building, state of charge of a battery, etc..). The second therm penalises the derivative of the control signal in order to avoid having hard control effort on the system (in this way the control signal is smoothened out). Finally, the third term penalises the control signal itself by a weighting factor (i.e. energy price if u_j represents electrical power). Matrices Q and R in Eq. B.6 are used to assign different weights to the squared norms: $\|x\|^2_Q = x^T Q x$.

Note that MPC does not provide an analytical expression of the feedback control law, therefore it is necessary to recompute the control vector at every time step using the updated system state. This concept is known as *receding horizon* (Fig.B.1), and at every MPC computation only the first value of the obtained control vector is applied to the physical system.

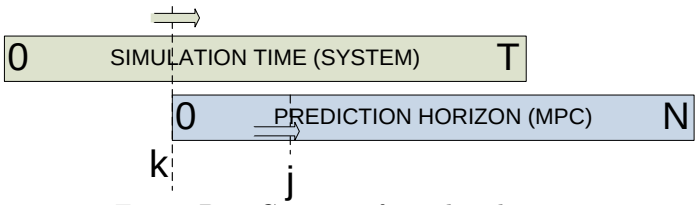


Figure B.1: Concept of receding horizon.

Assuming $k = 0$, without loss of generality, the MPC problem is stated as:

$$\min_{u_j,\theta_j} V : \tfrac{1}{2}\left[\sum_{j=1}^{N}\|y_j - r_j\|_Q^2 + \sum_{j=0}^{N-1}\|\Delta u_j\|_R^2 + \sum_{j=0}^{N-1}\rho_j u_j + \sum_{j=1}^{N}\eta_j\theta_j\right]$$

$$\begin{aligned} s.t. \quad & x_{j+1} = A_d x_j + B_d u_j + E_d d_k \\ & y_j = C_d x_k + D_d u_k \\ & y_{\min} - \theta_j \le y_j \le y_{\max} + \theta_j \\ & u_{\min} \le u_j \le u_{\max} \\ & \Delta u_{\min} \le \Delta u_j \le \Delta u_{\max} \\ & \theta_j \ge 0 \end{aligned} \tag{B.7}$$

By massaging the equations, Problem B.7 can be casted to the following least-square minimisation problem and solved using standard commercially-available and open-source QP solvers:

$$\begin{aligned} \min_x \phi : \tfrac{1}{2}x^T H x + f^T x \\ s.t. : \quad Ax \le b \end{aligned} \tag{B.8}$$

where the matrices H, f, A, and b are computed as shown in the following. Let's consider the system in Eq. 4.14. Its state and output at a generic time k can be written as:

$$x_k = A^k x_0 + \sum_{i=0}^{k-1} A^{k-1-i} B u_i + \sum_{i=0}^{k-1} A^{k-1-i} E d_i \quad ; \quad y_k = C x_k$$

The evolution of y in the horizon N can be written in matrix form (e.g. $N = 3$):

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{Y} = \underbrace{\begin{bmatrix} CA \\ CA^2 \\ CA^3 \end{bmatrix}}_{\Phi_x} x_0 + \underbrace{\begin{bmatrix} H_1^u & 0 & 0 \\ H_2^u & H_1^u & 0 \\ H_3^u & H_2^u & H_1^u \end{bmatrix}}_{\Gamma_u} \underbrace{\begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix}}_{U} + \underbrace{\begin{bmatrix} H_1^d & 0 & 0 \\ H_2^d & H_1^d & 0 \\ H_3^d & H_2^d & H_1^d \end{bmatrix}}_{\Gamma_d} \underbrace{\begin{bmatrix} d_0 \\ d_1 \\ d_2 \end{bmatrix}}_{D}, \tag{B.9}$$

with:

$$H_i^u = CA^{i-1}B, \quad H_i^d = CA^{i-1}E\,.$$

Therefore, system 4.14 is equivalent to:

$$Y = \Phi_x x_0 + \Gamma_u U + \Gamma_d D\,. \tag{B.10}$$

Recall the first therm of the cost function in Problem B.7, $\sum_{j=1}^{N}\|y_j - r_j\|_Q^2$; considering Eq. B.10, it can be written as:

$$\|\Phi x_0 + \Gamma_u U + \Gamma_d D - R\|_Q^2 = \|\Gamma_u U - (R - \Phi x_0 - \Gamma_d D)\|_Q^2 = \|\Gamma_u U - m\|_Q^2\,, \tag{B.11}$$

where R denotes the vector $\begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}^T$. By developing the norm in Eq. B.11, we obtain:

$$\frac{1}{2}\left\|\Gamma_u U - m\right\|_Q^2 = \frac{1}{2}U^T\Gamma_u^T Q\Gamma_u U - \left(\Gamma_u^T Q m\right)^T U + \frac{1}{2}m^T Q m \,. \tag{B.12}$$

The second therm of the cost function in Problem B.7 is: $\frac{1}{2}\sum_{j=0}^{N-1}\left\|\Delta u_j\right\|_R^2$, where $\Delta u_j = u_j - u_{j-1}$. In vector notation (e.g. $N=3$):

$$\frac{1}{2}\begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix}^T \underbrace{\begin{bmatrix} 2S & -S & 0 \\ -S & 2S & -S \\ 0 & -S & S \end{bmatrix}}_{H_s} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} + \underbrace{\begin{bmatrix} -Su_{-1} \\ 0 \\ 0 \end{bmatrix}^T}_{M_{u_{-1}}} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} + \frac{1}{2}u_{-1}Su_{-1} \,,$$

or:

$$\frac{1}{2}U^T H_s U + M_{u_{-1}}U + \frac{1}{2}u_{-1}Su_{-1} \,. \tag{B.13}$$

The third and fourth therm of the cost function in Problem B.7 are linear, and therefore they can be directly written as:

$$\sum_{j=0}^{N-1}\rho_j u_j + \sum_{j=1}^{N}\eta_j\theta_j = \underbrace{\begin{bmatrix} \rho_0 \\ \rho_1 \\ \rho_2 \end{bmatrix}^T}_{\Psi} U + \underbrace{\begin{bmatrix} \eta_0 \\ \eta_1 \\ \eta_2 \end{bmatrix}^T}_{\Upsilon} \underbrace{\begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}}_{\Theta} = \Psi U + \Upsilon\Theta \,. \tag{B.14}$$

Therefore, the matrices H and f, and the vector x in Problem B.8 are:

$$H = \begin{bmatrix} \Gamma_u^T Q\Gamma_u + H_s & O \\ O & O_N \end{bmatrix}$$

$$f = \begin{bmatrix} \Gamma_u^T Q\left(\Phi x_0 + \Gamma_d D - R\right) + M_u u_{-1} + \Psi \\ \Upsilon \end{bmatrix}$$

$$x = \begin{bmatrix} U \\ \Theta \end{bmatrix}$$

Note that the constant therms in Eq. B.12 and B.13 are offsets to the optimisation problem, and therefore are discarded.

We continue with showing the calculations for matrices A and b. Recall the system dynamics in Eq. B.10. We can express the first three constraints of Problem B.7 as:

$$\begin{matrix} \Gamma_u U + \left(\Phi_x x_0 + \Gamma_d D\right) \leq & Y_{\max} + \Theta \\ -\Gamma_u U - \left(\Phi_x x_0 + \Gamma_d D\right) \leq & -Y_{\min} + \Theta \end{matrix} \,,$$

or in matrix form:

$$\begin{bmatrix} \Gamma_u & -I \\ -\Gamma_u & -I \end{bmatrix}\begin{bmatrix} U \\ \Theta \end{bmatrix} \leq \begin{bmatrix} Y_{\max} - \Phi_x x_0 - \Gamma_d D \\ -Y_{\min} + \Phi_x x_0 + \Gamma_d D \end{bmatrix} \tag{B.15}$$

Considering $N = 3$ as example, we can express the fifth constraint of Problem B.7 in matrix form:

$$\begin{bmatrix} \Delta u_{\min} + u_{-1} \\ \Delta u_{\min} \\ \Delta u_{\min} \end{bmatrix} \leq \begin{bmatrix} I & O & O \\ -I & I & O \\ O & -I & I \end{bmatrix} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} \leq \begin{bmatrix} \Delta u_{\max} + u_{-1} \\ \Delta u_{\max} \\ \Delta u_{\max} \end{bmatrix}$$

which, extracting the dependency of u_{-1}, can be expressed with the following two constraints:

$$\underbrace{\begin{bmatrix} \Delta u_{\min} \\ \Delta u_{\min} \end{bmatrix}}_{\Delta U_{\min}} \leq \underbrace{\begin{bmatrix} -I & I & O \\ O & -I & I \end{bmatrix}}_{\Lambda} \begin{bmatrix} u_0 \\ u_1 \\ u_2 \end{bmatrix} \leq \underbrace{\begin{bmatrix} \Delta u_{\max} \\ \Delta u_{\max} \end{bmatrix}}_{\Delta U_{\max}} \tag{B.16}$$

$$\Delta u_{\min} + u_{-1} \leq u_0 \leq \Delta u_{\max} + u_{-1} \, . \tag{B.17}$$

Equation B.16 is re-written as:

$$\begin{bmatrix} \Lambda & O \\ -\Lambda & O \end{bmatrix} \begin{bmatrix} U \\ \Theta \end{bmatrix} \leq \begin{bmatrix} \Delta U_{\max} \\ -\Delta U_{\min} \end{bmatrix} \tag{B.18}$$

The fourth and sixth constraints of Problem B.7 are straightforward:

$$\begin{bmatrix} I & O \\ -I & O \\ O & -I \end{bmatrix} \begin{bmatrix} U \\ \Theta \end{bmatrix} \leq \begin{bmatrix} U_{\max} \\ -U_{\min} \\ O \end{bmatrix} \, , \tag{B.19}$$

where $U_{\max}$ and $U_{\min}$ are defined taking into account Eq. B.17:

$$\begin{aligned} U_{\max} &= \begin{bmatrix} \min\{u_{\max}, \Delta u_{\max} + u_{-1}\} & u_{\max,1} & \dots & u_{\max,N-1} \end{bmatrix}^T \\ U_{\min} &= \begin{bmatrix} \max\{u_{\min}, \Delta u_{\min} + u_{-1}\} & u_{\min,1} & \dots & u_{\min,N-1} \end{bmatrix}^T \end{aligned} \, .$$

By combining Eq. B.15, B.18, and B.19, matrices A and b in Problem B.8 are therefore defined as follows:

$$A = \begin{bmatrix} \Gamma_u & -I \\ -\Gamma_u & -I \\ \Lambda & O \\ -\Lambda & O \\ I & O \\ -I & O \\ O & -I \end{bmatrix} \quad , \quad b = \begin{bmatrix} Y_{\max} - \Phi_x x_0 - \Gamma_d D \\ -Y_{\min} + \Phi_x x_0 + \Gamma_d D \\ \Delta U_{\max} \\ -\Delta U_{\min} \\ U_{\max} \\ -U_{\min} \\ O \end{bmatrix} \, . \tag{B.20}$$

Concerning a Bakery unit, for example a dishwasher, it has a predefined power pattern and an earliest starting time and latest finishing time defined by the user.

the MPC Problem B.7 needs therefore to be reformulated as follows:

$$\begin{aligned} &\min_{u_j, d_j} V : \sum_{j=0}^{N-1} u_j \rho_j \\ &s.t. \quad u_i = P_i \quad \forall i \in \{k \dots k+\tau-1\} \\ &\qquad k \in \{\max(0, T_{earliest}) \dots \min(T_{latest} - \tau, N)\} \\ &\qquad P_i \in \mathcal{P} \\ &\qquad \rho_j > 0 \end{aligned} \tag{B.21}$$

Problem B.21 is reformulated as a Mixed-Integer Quadratic Program (MIQP):

$$\begin{aligned} &\min_x V : x^T H x + f^T x \\ &s.t. : \quad Ax \leq b \\ &\qquad\quad A_{eq} = b_{eq} \end{aligned} \tag{B.22}$$

where:

$$x = [\, u_0 \dots u_{N-1} |\, d_0 \dots d_{N-\tau-1}\,] \;,\; x \in \mathbb{R} \;,\; d \in \{0,1\}$$

$$H = O_{2N-\tau} \quad , \quad f = [\, \rho_0 \dots \rho_{N-1} |\, O_{1,N-\tau}\,]^T$$

and τ is the process running time (in time steps), N is the MPC prediction horizon, ρ is the energy price, and d is a slack variable indicating the process starting time. Clearly, u is the control signal. We show the computation of matrices A and b in the case of $\tau = 2$ and $N = 2$:

$$A = \left[\begin{array}{ccccc|cccc} -1 & 0 & 0 & 0 & 0 & P_1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & P_2 & 0 & 0 & 0 \\ \hline 0 & -1 & 0 & 0 & 0 & 0 & P_1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & P_2 & 0 & 0 \\ \hline 0 & 0 & -1 & 0 & 0 & 0 & 0 & P_1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & P_2 & 0 \\ \hline 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & P_1 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & P_2 \end{array}\right] \;,\; x \left[\begin{array}{c} u_0 \\ u_1 \\ u_2 \\ u_3 \\ \hline u_4 \\ d_0 \\ d_1 \\ d_2 \\ d_3 \end{array}\right] \;,\; b = O_{2N-\tau,1}$$

where P_i is the average power consumption in the time slot i. The equality constraints are used to ensure that the process is run at least once (in the following, we impose just once):

$$A_{eq} = [\, O_{1,N} |\, I_{1,N-\tau+1}\,] \quad , \quad b_{eq} = 1$$

B.3 Dual Decomposition.

This method is a variant of the Dual Ascent method that applies to decomposable objective functions and allows solving the N optimisation problems separately, and

concurrently. The basic concept is simple: to the primal problem is associated a dual problem, which is also separable and convex. The dual problem uses an additional variable that serves to coordinate the distributed solution of the N subproblems. Once found the solution to the dual subproblems, assuming that the strong duality holds [1], the solution to the primal problems is easily reconstructed.

Consider the problem:

$$\begin{array}{l} \min\limits_{x} \ f(x) \\ s.t: \ Ax = b \end{array} ,$$

the associated Lagrangian function (assuming $A = [A_1 \dots A_N]$) is:

$$L(x,y) = f(x) + y^T (Ax - b) \ = \sum_{i=1}^{N} \left(f_i(x_i) + y^T A_i x_i - y^T b N^{-1} \right) ,$$

and the variable y is the Lagrange multiplier, called also *dual variable*. Let $g(y)$ denote the *dual function*:

$$g(y) = \inf_x L(x,y) = -f^* \left(-A^T y \right) - b^T y ,$$

where f^* is the complex conjugate of f [2]; the dual problem is:

$$\min_y g(y) \ .$$

Denoting the solution to the dual problem with y^*, the solution to the primal problem, x^* is given by:

$$x^* = \arg\min_x \ L(x, y^*) \ ,$$

provided that there is a unique minimiser of $L(x, y^*)$ (which is actually the case since g is convex and smooth). Therefore the N subproblems can be iterativey solved in parallel with the following algorithm:

$$\begin{array}{l} x_i^{k+1} := \arg\min\limits_{x_i} L_i \left(x_i, y^k \right) \\ y^{k+1} := y^k + \alpha \nabla g(y) \end{array} ,$$

where $\alpha > 0$, $\nabla g(y) = \left(Ax^{k+1} - b \right)$, and for which a stopping criterion can be

$$\left\| y^{k+1} - y^k \right\|^2 < \varepsilon$$

[1] Strong duality holds when the primal and the dual solutions are equivalent or, in other words, the duality gap is zero. A sufficient condition for strong duality in convex optimisation problems is given by the Slater's condition, which states that the feasible region must have an interior point.

[2] Let $f : \mathbb{R}^n \mapsto \mathbb{R}$. The function $f^* : \mathbb{R}^n \mapsto \mathbb{R}$ is called *conjugate function* and is defined as: $f^*(y) = \sup\limits_{x \in dom\ f} \left(y^T x - f(x) \right)$. We see immediately that f^* is a convex function, since it is the pointwise supremum of a family of convex (indeed, affine) functions of y. [Boy+11]

Note that the x-minimisation step is carried independently and in parallel. Once the x_i^{k+1} are computed, they are sent to the coordinator. This last carries the y-minimisation step and broadcasts the resulting y^{k+1}. In this context, the dual variable can be thought as a price variable. An implementation of DMPC based on DD for grid congestion management is present by Biegel in [Bie+12].

The Dual decomposition works if g is differentiable, and therefore $\nabla g(y) = \left(Ax^{k+1} - b\right)$ holds. In this case, the convergence is monotone, $g\left(y^{k+1}\right) > g\left(y^k\right)$, and x^k and y^k converge respectively to an optimal point and a dual optimal point. In some cases in which g is not differentiable, the choice of α^k is different and the algorithm may not converge or the convergence may not be monotone. For example, if the decision variables are binary, the Dual Decomposition does not converge.

B.4 Alternating Direction Method of Multipliers.

Augmented Lagrangian methods were developed in part to bring robustness to the dual ascent method, and in particular, to yield convergence without assumptions like strict convexity or finiteness of f. The **Alternating Direction Method of Multipliers (ADMM)** is an algorithm that blends the decomposability of dual ascent with the superior convergence properties of the method of multipliers [Boy+11]. Without going into details of ADMM, we present below the algorithm.

Consider a problem of type:

$$\begin{aligned} &\min\ f(x) + g(z) \\ &s.t.:\ Ax + Bz = c \end{aligned}$$

where the only assumption is that f and g are non-strictly convex. Note the difference with general linear equality constraint problem: the original variable x has been split in x and z, with the objective function being separable across splitting. The Lagrangian function is presented below, which has been augmented with a quadratic therm according to the method of augmented multipliers:

$$L_\rho(x, z, y) = f(x) + g(z) + y^T(Ax + bz - c) + (\rho/2)\|Ax + bz - c\|_2^2 \ ,$$

where $\rho > 0$ is called *penalty parameter*. The augmented Lagrangian is introduced to enhance the robustness of Dual Decomposition and to yield convergence without the assumption of strictly convexity of f and g [Boy+11]. The ADMM algorithm follows:

$$\begin{aligned} x^{k+1} &:= \arg\min_x L_\rho\left(x, z^k, y^k\right) \\ z^{k+1} &:= \arg\min_z L_\rho\left(x^{k+1}, z, y^k\right) \\ y^{k+1} &:= y^k + \rho\left(Ax^{k+1} + Bz^{k+1} - c\right) \end{aligned}$$

When implementing ADMM, x and y are local variables to the subsystem i, whereas z is the global variable. Boyd et al. in [Boy+11] point out the main features needed to implement ADMM: *mutable state:* each subsystem must be able to

store the current values of x_i and y_i (local variables); *local computation:* each subsystem must be able to solve a *small* convex problem; *global aggregation:* there must be a system for collecting and averaging the local variables, and broadcast the result back to each system. Note that a message passing (MP) implementation is also possible in ADMM for general consensus problems, where the z is not centrally stored and the nodes communicate between themselves the local variables; *synchronisation:* all the local variables must be updated before performing global aggregation, and the local updates must use the latest global variable.

A general convergence result for ADMM consists in the following two conditions: 1) the functions f and g are closed, proper, and convex, and 2) the augmented Lagrangian function has a saddle point. Under these two conditions, ADMM leads to [Boy+11]: *Residual convergence*, i.e. the iterates approach feasibility, *Objective convergence*, i.e. the objective function of the iterates approaches the optimal value, and *Dual variable convergence*, i.e. the dual variable converges to a dual optimal point. The interested reader can refer to [Boy+11] for an extensive review on distributed optimisation techniques, the ADMM method, and its application to a variety of problems.

B.5 Extremely Randomized Trees.

A tree is a data structure organised as a type of graph, where nodes and edges follow a hierarchical structure. If a node generates other branches it is called *split node*, otherwise it is a terminal node and it is called *leaf node* (or simply *leaf*). Most of trees applications for classification and regression use binary trees, where the split nodes have only two outgoing edges. Note that all nodes have exactly one incoming edge and, in contrast with general graphs, trees do not contain loops [CSK11]. The Extremely Randomized Trees (ERT) are a type of regression trees with some peculiar characteristics.

Let's consider the problem of learning a general non-linear function:

$$f : \mathbb{R}^m \mapsto \mathbb{R} \,,$$

where the domain $\mathbb{R}^m$ is often referred as *feature space* and the function output is called *target value*. The idea behind regression trees is that the estimation of the target value can be performed by sequentially evaluating different properties of the feature space. Which property to evaluate at which step depends on the previous evaluation outcome (true or false). The regression based on trees uses the tree as path to these evaluation steps until a leaf node is reached, where an estimation of the target value is available.

Trees training is a supervised-learning task, in which the training dataset consist of tuples of input-target points. The goal is to train the regressor in order to emulate the input-output relationship hidden in the experimental data coming from the DSR. Trees and ERT are trained by taking one training point at time. Starting from the root, a feature (i.e. dimension of the feature space), called split feature, is randomly

selected and a random number (split value) is drawn from a uniform distribution which bounds are specific to each feature and correspond to the minimum and maximum observed value (this needs a pre-processing the training data set). In ERT, both split feature and split value are chosen randomly, while in other approaches based on trees, such as Random Forests (RF), the split feature and split value are optimised based on a given information criterion (e.g. information Enthropy) [CSK11]. The split feature and the split value are stored in the current node, called *split node*, and are used to divide the training set in two subsets. Each point having the value of the split feature grater than the split value is sent to the right branch, while the remaining points are sent to the left branch. This divide-and-conquer process is iterated recursively for each node, until the number of train points in a node reaches n_{min}. At this point the branch growth is stopped and the node becomes a *leaf node*.

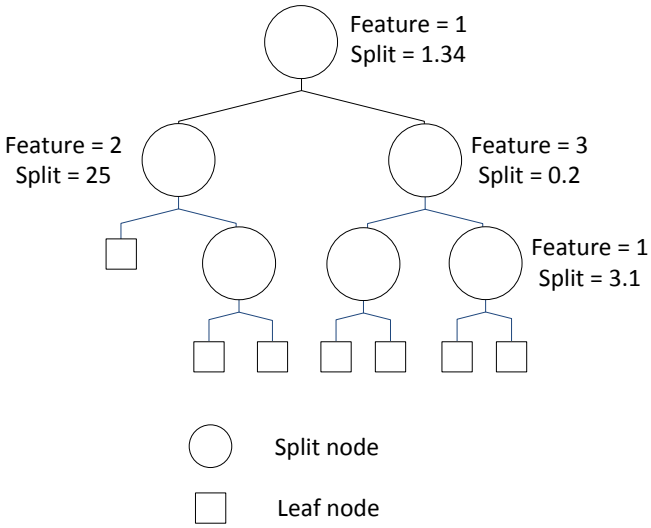


Figure B.2: Example of tree training

Note that at every leaf node there is a collection of train points, which includes the target values. Then, input data with unknown target is tested according to the tree and, when reached a leaf node, the target value is predicted by averaging the target values of the train points present in that leaf. Being a regression tree (or ERT in this case) a sort of piece-wise linear approximator of the training dataset, it is a low-accuracy regressor (weak regressor).

Ensemble methods for ERT. A way of improving the regression accuracy is to combine multiple weak regressors (base learners) in ensembles, instead of using a single sophisticated regressor to solve the same problem [Zho12].

Most of ensemble learning approaches use a single technique to produce homogeneous sets of learners (regressors), even though in some applications heterogeneous

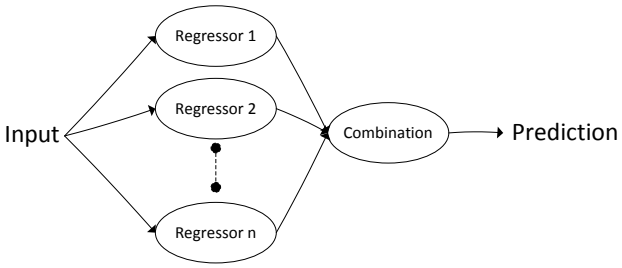


Figure B.3: Example of ensemble architecture.

sets of learners can be used (e.g. decision trees, neural networks, SVMs, etc...). The beauty and appeal of ensemble methods stay in: i) the generalisation capabilities of an ensemble, which are often much stronger than a single strong learner and ii) the ability of boosting weak learners, which are slightly better than a random guess, into strong learners, which makes accurate predictions. Furthermore, the computational time required to train weak regressors is often few orders of magnitude smaller than strong regressors.

In this research, a number of ERTs are fitted on various sub-samples of the learning dataset and then combined in a meta estimator by means of averaging in order to improve the predictive accuracy and control the over-fitting. The software used to perform such task is the Scikit Leran library for Python, which is distributed under open-source license [SCI]. We refer the interested reader to [Bis+06] and [Has+05] for a good insight on many machine learning techniques for linear and non-linear signal analysis and data mining, to [CSK11] and [Bus+10] (appendix) for a closer look at classification and regression trees, and to [Zho12] for an overview on ensemble methods.

B.6 Value-iteration Dynamic Programming

The final goal of DP is to obtain a control policy that maximises the cumulative rewards along a trajectory starting from any initial state x_0, consequently the reward obtained by the controller in the long run. Recalling Eq. 5.6 and specialising it for the finite-horizon case, the *Q-function* associated to a generic policy h is:

$$Q^h(x,u) = \rho(x,u) + \sum_{k=1}^{T} \gamma^k \rho(x_k, h(x_k)) , \tag{B.23}$$

where T is the finite control horizon. The optimal Q-function is defined as the best Q-function that can be obtained by any policy:

$$Q^*(x,u) = \max_h Q^h(x,u) . \tag{B.24}$$

The Bellman optimality equation characterising Q^* states that the optimal value (reward) of an action u taken in state x corresponds to the immediate reward plus the discounted return obtained by the best action in the next state [Bus+10]:

$$Q^*(x,u) = \rho(x,u) + \gamma \max_{u'} Q^*(f(x,u), u') \tag{B.25}$$

The algorithm used in this section consists in iteratively computing the Q-function (from which the name **Q-iteration**) until Q^* is obtained. Then, the optimal policy h^* is derived from Q^* with:

$$h^*(x) \in \arg\max_{u \in U} Q^*(x,u). \tag{B.26}$$

The proof of convergence for the Q-iteration algorithm is provided in [Bus+10], and it is based on the fact that the update of the Q-function:

$$Q_{l+1} = \Psi(Q_l) \ , \tag{B.27}$$

is given by the map $\Psi(Q)$:

$$[\Psi(Q)](x,u) = \rho(x,u) + \gamma \max_{u'} Q(f(x,u), u') \ , \tag{B.28}$$

which is a contraction. Therefore it follows that:

$$Q^* = \Psi(Q^*) \ , \tag{B.29}$$

and equation B.27 is iterated until $Q_l = Q_{l-1}$ or another stopping criterion is met.

We are interested in obtaining from the DP controller a closed-loop policy that contains the optimal control actions for any internal temperature at any time within the optimisation time horizon. Therefore the backward induction approach is suitable. The following python code shows how the implemented DP controller works, which optimises the control action from the end of the control horizon backward to the beginning, at every state (measured temperature).

```
for t in np.arange(tmax,-1,-1):
    print 'time:'+str(t)
    CostToGoNew = np.zeros(np.shape(CostToGo));
    policyMap   = np.zeros(np.shape(CostToGo));
    for stateC in np.arange(0,np.shape(stateArray)[0],1):
        # Select a specific state to analyze
        state = stateArray[stateC];
        temp  = [];
        controlArrayLocal = controlArray;
        for control in controlArrayLocal:
            data_input        = Test_set1[t,1:]
            localPrice        = control*priceSignal[t]
            data_input[0]     = state # T(i)
```

```
14              data_input[1]       = control # u(i)
15              out                 = MODEL0.predict(data_input) # T(i+1)
16              Vnext               = np.interp(out,stateArray,CostToGo);
17              temp                = np.append(temp,localPrice+Vnext);
18          indexMin                = temp.argmin();
19          minner                  = temp[indexMin]
20          CostToGoNew[stateC]     = minner;neurons
21          policyMap[stateC]       = indexMin;
22      CostToGo                    = CostToGoNew;
23      CostToGo[stateArray<20.2]   = 1e9; # deal with comfort
24      CostToGo[stateArray>21.8]   = 1e9; # deal with comfort
25      policyMat                   = np.vstack([policyMat,policyMap]);
26      VMat                        = np.vstack([VMat,CostToGoNew]);
```

The main loop goes from the end of the optimisation horizon (tmax) to the beginning of time, the computed policy is stored in the variable policyMap, and the value function in the CostToGo. Synthetically, this DP algorithm determines, for every time step, for every internal temperature, the best control action from all the possible control actions.

If we intended to perform DP optimisation for Problem 5.4, the previous code should have been extend with an additional nested loop, that for each measured internal temperature had to determine the action to be taken at each $T_i(t-1)$. This would have led the problem dimension to grow by one order of magnitude, resulting in a matrix of policies, where one policy of the type shown in Fig.5.9 is associated to every possible value of $T_i(t-1)$.

Bibliography

[AS] Dong Enrgy A/S. *Power Hub - an important part of the Smart Grid*.

[Baa+97] M Baadsgaard et al. "Estimation in stochastic differential equations with a state dependent diffusion term". In: *SYSID*. 1997.

[Bar98] Andrew G Barto. *Reinforcement learning: An introduction*. MIT press, 1998.

[BC88] Steven L Beuerman and Edward J Coyle. "The delay characteristics of CSMA/CD networks". In: *Communications, IEEE Transactions on* 36.5 (1988), pages 553–563.

[Bel03] R. Bellman. *Dynamic Programming*. NY: Dover Publications, Incorporated., 2003.

[Bel09] Bellifemine, F.L. et al. "Smart Grid: Energia & ICT". In: *Notiziario tecnico Telecom Italia* 3 (December 2009), pp. 15–32.

[Ber+14a] Andrey Bernstein et al. "A Composable Method for Real-Time Control of Active Distribution Networks with Explicit Power Setpoints". In: *arXiv preprint arXiv:1403.2407* (2014).

[Ber+14b] Andrey Bernstein et al. "A Composable Method for Real-Time Control of Active Distribution Networks with Explicit Power Setpoints". In: *arXiv preprint arXiv:1403.2407* (2014).

[Ber+95] Dimitri P Bertsekas et al. *Dynamic programming and optimal control*. Volume 1. 2. Athena Scientific Belmont, MA, 1995.

[Bie+12] Benjamin Biegel et al. "Congestion management in a smart grid via shadow prices". In: *8th IFAC symposium on power plant and power system control*. 2012, pages 518–523.

[Bis+06] Christopher M Bishop et al. *Pattern recognition and machine learning*. Volume 1. springer New York, 2006.

[Bli+10a] Frits Bliek et al. "PowerMatching City, a living lab smart grid demonstration". In: *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*. IEEE. 2010, pages 1–8.

[Bli+10b] F. Bliek et al. "PowerMatching City, a living lab smart grid demonstration". In: *Innovative Smart Grid Technologies Conference Europe (ISGT Europe), 2010 IEEE PES*. October 2010, pages 1–8. DOI: 10.1109/ISGTEUROPE.2010.5638863.

[BM11] Peder Bacher and Henrik Madsen. "Identifying suitable models for the heat dynamics of buildings". In: *Energy and Buildings* 43.7 (2011), pages 1511–1522.

[Boy+11] Stephen Boyd et al. "Distributed optimization and statistical learning via the alternating direction method of multipliers". In: *Foundations and Trends® in Machine Learning* 3.1 (2011), pages 1–122.

[BP12] DEM Bondy and J Parvizi. "Modeling, identification and control for heat dynamics of buildings using robust economic model predictive control". PhD thesis. Master's thesis, Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark, 2012.

[Bus+10] Lucian Busoniu et al. *Reinforcement learning and dynamic programming using function approximators*. Volume 39. CRC press, 2010.

[Cam+13] Erik Cambria et al. "Extreme Learning Machines." In: *IEEE Intelligent Systems* 28.6 (2013), pages 30–59.

[CGK10] Peter Cappers, Charles Goldman, and David Kathan. "Demand response in US electricity markets: Empirical evidence". In: *Energy* 35.4 (2010), pages 1526–1535.

[Cle96] Scott H Clearwater. *Market-based control: A paradigm for distributed resource allocation*. World Scientific Publishing Co., Inc., 1996.

[Coa14] Smart Energy Demand Coalition. *Mapping Demand Response in Europe Today*. Technical report. April 2014.

[Cos+12a] Giuseppe Tommaso Costanzo et al. "An experimental study on load-peak shaving in smart homes by means of online admission control". In: *Innovative Smart Grid Technologies (ISGT Europe), 2012 3rd IEEE PES International Conference and Exhibition on*. IEEE. 2012, pages 1–8.

[Cos+12b] G.T. Costanzo et al. "A System Architecture for Autonomous Demand Side Load Management in Smart Buildings". In: *Smart Grid, IEEE Transactions on* 3.4 (December 2012), pages 2157–2165. ISSN: 1949-3053. DOI: 10.1109/TSG.2012.2217358.

[Cos+13a] G.T. Costanzo et al. "A coordination scheme for distributed model predictive control: Integration of flexible DERs". In: *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES*. October 2013, pages 1–5. DOI: 10.1109/ISGTEurope.2013.6695474.

[Cos+13b] G.T. Costanzo et al. "Grey-box modeling for system identification of household refrigerators: A step toward smart appliances". In: *Energy (IYCE), 2013 4th International Youth Conference on*. June 2013, pages 1–5. DOI: 10.1109/IYCE.2013.6604197.

[CSK11] Antonio Criminisi, Jamie Shotton, and Ender Konukoglu. "Decision forests: A unified framework for classification, regression, density estimation, manifold learning and semi-supervised learning". In: *Foundations and Trends® in Computer Graphics and Vision* 7 (2011), pages 81–227.

[Dan13] Danish Technical University - Department of Informatics and Mathematical Modeling. *Continuous Time Stochastic Modeling*. 2013.

[Dou+11] Philip J Douglass et al. "Demand as frequency controlled reserve: Implementation and practical demonstration". In: *Innovative Smart Grid Technologies (ISGT Europe), 2011 2nd IEEE PES International Conference and Exhibition on*. IEEE. 2011, pages 1–7.

[DSS06] Thomas Degner, Jürgen Schmid, and Philipp Strauss. *Dispower: Distributed Generation with High Penetration of Renewable Energy Sources: Final Public Report*. ISET, 2006.

[EGW05] Damien Ernst, Pierre Geurts, and Louis Wehenkel. "Tree-based batch mode reinforcement learning". In: *Journal of Machine Learning Research*. 2005, pages 503–556.

[Ele09] Electric Power Research Institute. *Assessment of Achievable Potential from Energy Efficiency and Demand Response Programs in the U.S. (2010-2030)*. Technical report. 2009.

[Ern+09] Damien Ernst et al. "Reinforcement learning versus model predictive control: a comparison on a power system problem". In: *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on* 39.2 (2009), pages 517–529.

[Eur13] European Climate Foundation. *Roadmap 2050: a practical guide to a prosperous, low-carbon Europe*. Technical report. December 2013.

[FS10] Ahmad Faruqui and Sanem Sergici. "Household response to dynamic pricing of electricity: a survey of 15 experiments". In: *Journal of Regulatory Economics* 38.2 (2010), pages 193–225.

[GI10] O. Gehrke and F. Isleifsson. "An aggregation friendly information model for demand side resources". In: *Local Computer Networks (LCN), 2010 IEEE 35th Conference on*. October 2010, pages 1019–1023. DOI: 10.1109/LCN.2010.5735674.

[Gio+12] Vincenzo Giordano et al. *Smart Grid projects in Europe: Lessons learned and current developments*. Technical report. European Joint Research Center), 2012.

[Ham+07] D J Hammerstrom et al. "Pacific Northwest GridWise Testbed. Demonstration Projects Part I. Olympic Peninsula Project." In: *Contract* (2007), page 157.

[Has+05] Trevor Hastie et al. "The elements of statistical learning: data mining, inference and prediction". In: *The Mathematical Intelligencer* 27.2 (2005), pages 83–85.

[HN84] William J Hausman and John L Neufeld. "Time-of-day pricing in the US electric power industry at the turn of the century". In: *The RAND Journal of Economics* (1984), pages 116–126.

[HR11] Roland Hafner and Martin Riedmiller. “Reinforcement learning in feedback control”. In: *Machine learning* 84.1-2 (2011), pages 137–169.

[HZS06] Guang-Bin Huang, Qin-Yu Zhu, and Chee-Kheong Siew. “Extreme learning machine: theory and applications”. In: *Neurocomputing* 70.1 (2006), pages 489–501.

[IEA11] IEA. “World energy outlook 2011.” In: (2011).

[Ins13] Electric Power Research Institute. *IntelliGrid program, 2013 annual review*. Technical report. 2013.

[JF] Bingnan Jiang and Yunsi Fei. “Smart Home in Smart Microgrid: A Cost-Effective Energy Ecosystem with Intelligent Hierarchical Agents”. In: ().

[Jor+11] JM Jorgensen et al. “Ecogrid EU?A prototype for european smart grids”. In: *Power and Energy Society General Meeting, 2011 IEEE*. IEEE. 2011, pages 1–7.

[Juh+13] Rune Juhl et al. “CTSM-R User Guide”. In: *Technical University of Denmark* 2 (2013).

[Kar+12] Emre Can Kara et al. “Using smart devices for system-level management and control in the smart grid: A reinforcement learning framework”. In: *Smart Grid Communications (SmartGridComm), 2012 IEEE Third International Conference on*. IEEE. 2012, pages 85–90.

[KGB12] Daniel Kullmann, Oliver Gehrke, and Henrik Bindner. “Asynchronous control of Distributed Energy Resources using behaviour descriptions”. In: *Computing, Networking and Communications (ICNC), 2012 International Conference on*. IEEE. 2012, pages 216–220.

[KMJ04] Niels Rode Kristensen, Henrik Madsen, and Sten Bay Jørgensen. “Parameter estimation in stochastic grey-box models”. In: *Automatica* 40.2 (2004), pages 225–237.

[Kos+13] Anna Magdalena Kosek et al. “An overview of demand side management control schemes for buildings in smart grids”. In: *Smart Energy Grid Engineering (SEGE), 2013 IEEE International Conference on*. IEEE. 2013, pages 1–9.

[KVM10] Koen Kok, Gerben Venekamp, and Pamala Macdougall. “Market-based control in decentralized electrical power systems”. In: *First International Workshop on Agent Technologies for Energy Systems, ATES2010, Toronto*. 2010.

[LA05] Fernando L. and Alvarado. “Controlling power systems with price signals”. In: *Decision Support Systems* 40.3?4 (2005). Challenges of restructuring the power industry, pages 495–504. ISSN: 0167-9236. DOI: 10.1016/j.dss.2004.05.012. URL: http://www.sciencedirect.com/science/article/pii/S0167923604001241.

[LR14] Thomas Lampe and Martin Riedmiller. “Approximate model-assisted Neural Fitted Q-Iteration”. In: *Neural Networks (IJCNN), 2014 International Joint Conference on*. IEEE. 2014, pages 2698–2704.

[Ma+12] Yudong Ma et al. “Model predictive control for the operation of building cooling systems”. In: *Control Systems Technology, IEEE Transactions on* 20.3 (2012), pages 796–803.

[Mac02] Jan Marian Maciejowski. *Predictive control: with constraints*. Pearson education, 2002.

[Mad08] H. Madsen. *Time Series Analysis*. Texts in statistical science. Chapman & Hall/CRC, 2008. ISBN: 9781420059670.

[Mar+12] F. Marra et al. “Demand profile study of battery electric vehicle under different charging options”. In: *Power and Energy Society General Meeting, 2012 IEEE*. 2012, pages 1–7. DOI: 10.1109/PESGM.2012.6345063.

[ML83] J.S. Meditch and C.-T.A. Lea. “Stability and Optimization of the CSMA and CSMA/CD Channels”. In: *Communications, IEEE Transactions on* 31.6 (June 1983), pages 763–774. ISSN: 0090-6778. DOI: 10.1109/TCOM.1983.1095881.

[Mor+01] Nicolas Morel et al. “Neurobat, a predictive and adaptive heating control system using artificial neural networks”. In: *International Journal of Solar Energy* 21.2-3 (2001), pages 161–201.

[NF08] Alberto Hernandez Neto and Flávio Augusto Sanzovo Fiorelli. “Comparison between detailed model simulation and artificial neural network for forecasting building energy consumption”. In: *Energy and Buildings* 40.12 (2008), pages 2169–2176.

[Nor+13] N. C. Nordentoft et al. *Development of a DSO-Market on Flexibility Services*. Technical report. iPower, 2013.

[Nor13] Niels Christian Nordentoft. *Development of a DSO-market on Flexibility Services*. Technical report. iPower, 2013.

[NSN11] Per Nørgaard, Fabrizio Sossan, and Henrik Nielsen. “INDIRECT REGULATION OF MANY DER UNITS THROUGH BROADCASTED DYNAMIC PRICE SIGNAL”. In: *21st International Conference on Electricity Distribution*. 2011.

[Old+10] Frauke Oldewurtel et al. “Energy efficient building climate control using stochastic model predictive control and weather predictions”. In: *American control conference (ACC), 2010*. IEEE. 2010, pages 5100–5105.

[Old+12] Frauke Oldewurtel et al. “Use of model predictive control and weather forecasts for energy efficient building climate control”. In: *Energy and Buildings* 45 (2012), pages 15–27.

[ONe+10] Daniel O'Neill et al. "Residential demand response using reinforcement learning". In: *Smart Grid Communications (SmartGridComm), 2010 First IEEE International Conference on*. IEEE. 2010, pages 409–414.

[P+05] Stavros Papathanassiou, Nikos Hatziargyriou, Kai Strunz, et al. "A benchmark low voltage microgrid network". In: *Proceedings of the CIGRE Symposium: Power Systems with Dispersed Generation*. 2005, pages 1–8.

[Ped+14] Rasmus Pedersen et al. "Aggregation and Control of Supermarket Refrigeration Systems in a Smart Grid". In: *Proceedings of the 19th IFAC World Congress*. 2014.

[Pee+09] Eefje Peeters et al. "ADDRESS: scenarios and architecture for active demand development in the smart grids of the future". In: *Electricity Distribution-Part 1, 2009. CIRED 2009. 20th International Conference and Exhibition on*. IET. 2009, pages 1–4.

[Pet+13] M.K. Petersen et al. "A taxonomy for modeling flexibility and a computationally efficient algorithm for dispatch in Smart Grids". In: *American Control Conference (ACC), 2013*. June 2013, pages 1150–1156. DOI: 10.1109/ACC.2013.6579991.

[Pin+12] Pierre Pinson et al. *INDIRECT CONTROL BY PRICES ? BASIC CONCEPTS, APPLICATIONS AND REQUIREMENTS*. Technical report. iPower, 2012. URL: http://www.ipower-net.dk/~/media/iPower/Documents/iPower-FlexInterface-IndirectControl.ashx.

[PL06] Paolo Piagi and Robert H Lasseter. "Autonomous control of microgrids". In: *Power Engineering Society General Meeting, 2006. IEEE*. IEEE. 2006, 8–pp.

[PM09] V. Pothamsetty and S. Malik. *Smart Grid: Leveraging Intelligent Communications to Transform the Power Infrastructure*. Technical report. CISCO Systems Whyte Paper, February 2009.

[Prí+11] Samuel Prívara et al. "Model predictive control of a building heating system: The first experience". In: *Energy and Buildings* 43.2 (2011), pages 564–572.

[Ras06] Eric Rasmusen. *Games and Information: An Introduction to Game Theory*. Wiley, 2006. ISBN: 9781405136662. URL: http://books.google.dk/books?id=5XEMuJwnBmUC.

[RK04] Reuven Y Rubinstein and Dirk P Kroese. *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer Science & Business Media, 2004.

[Rue+14] Frederik Ruelens et al. "Demand Response of a Heterogeneous Cluster of Electric Water Heaters Using Batch Reinforcement Learning". In: *status: accepted* (2014).

[Rue+15] Frederik Ruelens et al. “Residential Demand Response Applications Using Batch Reinforcement Learning”. In: *arXiv preprint arXiv:1504.02125* (2015).

[Rus80] B.D. Russell. “Communication Alternatives for Distribution Metering and Load Management”. In: *Power Apparatus and Systems, IEEE Transactions on* PAS-99.4 (July 1980), pages 1448–1455. ISSN: 0018-9510. DOI: 10.1109/TPAS.1980.319568.

[SB98] Richard S Sutton and Andrew G Barto. *Introduction to reinforcement learning*. MIT Press, 1998.

[SBM09] Michael Sullivan, Josh Bode, and Mangasarin. *2009 Pacific Gas and Electric Company SmartAC Ancillary Services Pilot*. Technical report. Freeman, Sullivan & Co., 2009. URL: http://www.pge.com/includes/docs/pdfs/mybusiness/energysavingsrebates/demandresponse/cs/2009_pacific_gas_and_electric_company_smartAC_ancillary_services_pilot.pdf.

[SBN14] Fabrizio Sossan, Henrik W. Bindner, and Per Bromand N rg rd. “Indirect control of flexible demand for power system applications.” PhD thesis. 2014.

[Sca09] Riccardo Scattolini. “Architectures for distributed and hierarchical model predictive control–a review”. In: *Journal of Process Control* 19.5 (2009), pages 723–731.

[Sch+80] F.C. Schweppe et al. “Homeostatic Utility Control”. In: *Power Apparatus and Systems, IEEE Transactions on* PAS-99.3 (May 1980), pages 1151–1163. ISSN: 0018-9510. DOI: 10.1109/TPAS.1980.319745.

[SCI] SCIKIT. *Scikit Learn*. http://scikit-learn.org/stable/index.html.

[SML08] Maria Sebastian, Juan Marti, and Peter Lang. “Evolution of DSO control centre tool in order to maximize the value of aggregated distributed generation in smart grid”. In: *SmartGrids for Distribution, 2008. IET-CIRED. CIRED Seminar*. IET. 2008, pages 1–4.

[Sos+13] F. Sossan et al. “Scheduling of domestic water heater power demand for maximizing PV self-consumption using model predictive control”. In: *Innovative Smart Grid Technologies Europe (ISGT EUROPE), 2013 4th IEEE/PES*. October 2013, pages 1–5. DOI: 10.1109/ISGTEurope.2013.6695317.

[Sos+15] Fabrizio Sossan et al. “Grey-box Modelling of a Household Refrigeration Unit for Energy Consumption Prediction and Optimization Using Time Series Data”. In: *Submitted to: Sustainable Energy, Grids and Networks* (2015).

[Sus11] Susanne Langsdorf. *EU Energy Policy: From the ECSC to the Energy Roadmap 2050*. Technical report. December 2011.

[Tha08] Anders Thavlov. "Dynamic optimization of power consumption". PhD thesis. Technical University of Denmark, DTU, DK-2800 Kgs. Lyngby, Denmark, 2008.

[The13] The Ministry of Climate, Energy and Building. *The Danish Climate Policy Plan: towards a low carbon society*. Technical report. August 2013.

[The75] The Council of European Communities. "Council Resolution of 17 September 1974 concerning a new energy policy strategy for the Community." In: *Official Journal C 153 , 09/07/1975* (1975), page 2.

[THL10] Jacopo Torriti, Mohamed G Hassan, and Matthew Leach. "Demand response experience in Europe: Policies, programmes and implementation". In: *Energy* 35.4 (2010), pages 1575–1583.

[US13] Daniel Urieli and Peter Stone. "A learning agent for heat-pump thermostat control". In: *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*. International Foundation for Autonomous Agents and Multiagent Systems. 2013, pages 1093–1100.

[Van+12] Koen Vanthournout et al. "A smart domestic hot water buffer". In: *Smart Grid, IEEE Transactions on* 3.4 (2012), pages 2121–2127.

[Vin11] Peter Vinter. *Introduction to Power Hub, Confidential*. Technical report. DONG Energy, November 2011.

[Wel+09] Adriaan van der Welle et al. *FENIX deliverable D3.3 Report: FINANCIAL AND SOCIO-ECONOMIC IMPACTS OF EMBRACING THE FENIX CONCEPT*. Technical report. Energy research Centre of the Netherlands (ECN), 2009.

[Zha+12] Wei Zhang et al. "Aggregate model for heterogeneous thermostatically controlled loads with demand response". In: *Power and Energy Society General Meeting, 2012 IEEE*. IEEE. 2012, pages 1–8.

[Zho12] Zhi-Hua Zhou. *Ensemble methods: foundations and algorithms*. CRC Press, 2012.

[ZM12] Hai-xiang Zhao and Frédéric Magoulès. "A review on the prediction of building energy consumption". In: *Renewable and Sustainable Energy Reviews* 16.6 (2012), pages 3586–3592.