EPFL

Thèse n° 7926

Reliability and Performance of Model Predictive Control for Demand Response with Residential Heat Pumps

Présentée le 8 janvier 2021

à la Faculté des sciences et techniques de l'ingénieur
Groupe SCI STI FM
Programme doctoral en énergie

pour l'obtention du grade de Docteur ès Sciences

par

Frédéric Louis-Pierre Raphaël Marie AMBLARD

Acceptée sur proposition du jury

Prof. D. Dujic, président du jury
Prof. F. Maréchal, Prof. J. Page, directeurs de thèse
Prof. D. Finn, rapporteur
Prof. G. François, rapporteur
Dr N. Wyrsch, rapporteur

École polytechnique fédérale de Lausanne

2021

Un chocard survole l'arête enneigée,
L'air chaud de la vallée le fait virevolter et danser.
Il plane, s'élève, pique, mais quel artiste !
Le voilà qui repart pour un nouveau tour de piste.

Debout au sommet, un homme l'admire, fasciné.
Dans un éclat de rire, il se met à rêver.
Qu'il aimerait partir, s'envoler, lui aussi,
Déployer ses ailes s'il en a envie.

Emerveillé par la grâce de l'oiseau, sa voix chante
Une ode à la nature, vivante et puissante.
Car le jeune homme aime être là-haut,
Libre au grand air, il est heureux, il est beau.

« Allons-y » lui murmurent ses copains,
C'est ainsi qu'ils se remettent en chemin.
Quelle chance ont-ils d'avoir exploré ces joyaux
Surtout en compagnie du meilleur des amigo.
— La Clairette

To Adrien…

Acknowledgements

I would like to thank many people that have helped me during my PhD journey:

Prof. François Maréchal, for having believed in me all these years, for giving me the opportunity to work on this thesis and for recommending me to Prof. Jessen Page. Jessen, you gave me a lot of freedom and responsibility but were always there to give me confidence in my work when I doubted myself. Thank you also for the meticulous proof-reading of this document.

Prof. Donal Finn, Prof. Grégory François and Dr. Nicolas Wyrsch, for your interest in my work and for the very interesting comments and feedback.

All the partners of the Sim4block project, for the very interesting discussions and their valuable help. I would like to especially thank Aurélien for the great collaboration and for helping me with my simulation, David L. for hosting me at Neurobat and teaching me how to code better. My fellow PhDs Marcus and Gerard, for the great ideas shared alongside a few beers at the project meetings. Many thanks to Claudio, Ivan and Kilian for helping me with the pilot site.

All my colleagues and friends: Raman, the source of knowledge, available 7/7 from 10am to 2am. You would answer questions I didn't know I had. Pablo, you taught me the beauty of python and that "some know how to code, others how to climb". Francesca, the heat pump specialist, you always supported me and I will miss our discussions across the office and in the train. Lucas, you are the greatest and only student I ever had. David T, you put your whole career at risk to help me install sensors in dusty basements. Tristan, you asked me challenging questions and helped me achieve my first 4'000m peak.

All my friends at HES-SO in Sandy order: Alain, André, Baljeet, David M, Didier, Glory, Julian, Koen, Laurent, Natacha, Robin, Sandy, Sébastien, Tharsan, Tim, Tommaso, Vincent and Melvyn. Thank you for the adventures in the real world and beyond, the climbing sessions, the exciting discussions and the crazy theories developed at the cafeteria, at building G or even late in the train. I would wake up some days and travel 3 hours just so that I could see you. And although you forgot me in the bathroom of a hotel in Brussels, I won't forget you all!

Acknowledgements

All my friends at IPESE, especially Stefano my mentor, Paul my role model and Hür my friend.

All my great friends from EPFL and the big family of the SIE section Flo, Akio, Baptiste, Adri, Perrine, Zoran, Pierre, Elisa, Theo, Mehdi, Guillaume, Ivo, Gillian, Adrien, Ale, Ladina, Kaboul et al. Thank you for the fantastic adventures and parties we had together from 11 years ago up to tomorrow.

All my flatmates since 8 years, and especially "la Grande Famille Coloc", Intisar, Oscar, Solenne, Grégoire et Vivi with whom I share more than just a flat. Special thanks to Inti's room for hosting my private defense.

A mes parents Mireille et Philippe pour avoir cru le plus en moi. Malgré la distance vous m'avez toujours soutenu de votre amour, même lorsque je disais à 3 heures du matin que j'allais tout arrêter. A ma grande sœur Camille qui en me racontant ses journées après l'école m'a sans doute donné le goût d'apprendre. A mon cousin Paul pour m'avoir vanté les mérites de l'EPFL et de la Suisse avant de partir au Mexique. Und Danke an meine zweite Familie die Links.

Vivian, the love of my life, soldier at the front of the war of my emotions and motivation. Your incommensurate love, patience and support played a big role in the success of this thesis.

Lausanne, 10 December 2020 Frédéric

Abstract

Distributed energy resources like solar panels and wind turbines are increasingly replacing fossil fuel energy sources. Due to their natural fluctuations they cause imbalance in the power grid when their production does not overlap with demand. Thus, there is a need to re-think the power grid into a system where the demand could be constantly adjusted to follow the fluctuation of the electricity generation. Demand Response (DR) is a mechanism to shift the electricity consumption of end-use customers from times with a high wholesale market price or when the security of the system is threatened to other time periods through various economic incentives. In Europe more than 20% of the total energy consumption is used for heating, cooling, and domestic hot water of buildings. Heat pumps are predicted to become one of the main heating and cooling technologies in the future. Since heat pumps can be powered by electricity, they are regarded as a great source of flexibility, especially when combined with building inertia and heat storage to be used in DR. The goal of this thesis is to assess the benefits and challenges of implementing a flexible control of heat pumps. Although this question has been addressed in the literature, most studies use simulations. Here, we take an applied approach by implementing a flexible controller on a pilot site of inhabited residential buildings in Switzerland. We developed an interface using Model Predictive Control (MPC), that was tailored to the particularities of each of the pilot site's buildings. After validating the system using simulation, we have investigated the implementation of the necessary infrastructure to realise DR on a pilot site. The implementation highlights technical difficulties related to the integration on an existing and heterogeneous system, with limited monitoring capabilities. The pilot site allowed us to demonstrate that we can supply DR services to a third party electricity aggregator. Although the heat pumps were difficult to control because they have non-linear behaviour and long reaction times, we were able to react to DR calls from the aggregator. The work provides a comprehensive study of the effort required to implement MPC in existing residential buildings, and confirms that DR in combination with heat pumps is a promising set-up to provide flexibility in the future.

Key words: Demand Response; Heat Pumps; Buildings; Model Predictive Control; Pilot Site

Résumé

Les ressources énergétiques distribuées comme les panneaux solaires et les éoliennes remplacent progressivement les sources d'énergie fossiles. En raison de leurs fluctuations naturelles, elles provoquent un déséquilibre dans le réseau électrique lorsque leur production ne chevauche pas la demande. Il est donc nécessaire de repenser le réseau électrique en un système où la demande pourrait être constamment ajustée pour suivre la fluctuation de la production d'électricité. La réponse à la demande (DR) est un mécanisme qui permet de déplacer la consommation d'électricité de clients finaux quand le prix du marché est élevé ou la sécurité du système est menacée vers d'autres périodes, grâce à diverses incitations économiques. En Europe, plus de 20% de la consommation totale d'énergie est utilisée pour le chauffage, la climatisation et l'eau chaude sanitaire des bâtiments. Il est attendu que les pompes à chaleur deviendront l'une des principales technologies de chauffage et de refroidissement à l'avenir. Comme les pompes à chaleur peuvent être alimentées par l'électricité, elles sont considérées comme une grande source de flexibilité à utiliser dans des programmes de DR. L'objectif de cette thèse est d'évaluer les avantages et les défis d'un contrôle flexible des pompes à chaleur. Bien que cette question ait été abordée dans la littérature, la plupart des études utilisent des données simulées. Nous avons choisi d'aborder cette question en la mettant en œuvre sur un site pilote de bâtiments résidentiels habités en Suisse. Nous avons développé une interface utilisant la commande prédictive (MPC), qui a été adaptée aux particularités de chacun des bâtiments du site pilote. Après l'avoir validée sur des données simulées, nous avons équipé le site pilote de l'infrastructure nécessaire pour effectuer du DR. Nous avons rencontré des difficultés techniques dues au fait que nous devions équiper un système préexistant et hétérogène, avec des possibilités de surveillance limitées. Néanmoins, nous avons été en mesure d'effectuer des services de DR avec un acteur tiers des marchés de l'électricité, appelé un agrégateur. Bien que les pompes à chaleur aient été difficiles à contrôler car elles ont un comportement non linéaire et nécessitent de longs temps de réaction, nous avons pu réagir aux appels de DR de l'agrégateur. Le travail fournit une étude complète de l'effort nécessaire pour mettre en œuvre la MPC dans les bâtiments résidentiels préexistants, et confirme que le DR en combinaison avec les pompes à chaleur est une configuration prometteuse pour assurer, à l'avenir, la flexibilité du réseau électrique.

Mots clefs : Réponse à la demande ; Pompe à chaleur ; Bâtiments ; contrôle prédictif ; Site pilote

Contents

List of Figures

List of Tables

1 Introduction

1.1 Growing need for flexibility

1.1.1 Intermittent Electricity production

In power grids, demand and supply need to be balanced at all times to avoid frequency variations, which in extreme cases could lead to black-outs. The traditional approach of power suppliers to tackle this problem is top-down grid architecture, where the power generation is controlled to fulfil the energy demand curves. Today, power grid systems are facing the great challenge of reducing the fossil fuel energy usage to mitigate the unequivocal climate change impacts. The proliferation of Distributed Energy Resources (DER) like solar panels and wind turbines in today's electric grid is helping to reach this goal. However, these DERs increase the difficulty to achieve the power balance due to their natural fluctuations introducing significant uncertainty. There is a need to re-think the power grid as a system where the demand could be constantly adjusted to follow the fluctuation of the electricity generation. This leads to an increased need for flexibility on the demand side (Kirschen et al., 2012; Lund et al., 2015) and for new storage capacity (Denholm et al., 2010; Weitemeyer et al., 2015). One attractive solution identified for future power systems is to manage not only the energy supply but also the demand via Demand Response (DR) programs.

1.1.2 Demand Response in residential buildings

Demand response is defined by the Strategic Energy Technologies Information System (SETIS) of the European Commission as: "*an intentional modification of normal consumption patterns by end-use customers in response to incentives from grid operators*" (Giodano et al., 2013). DR is a mechanism to increase the flexibility of the energy market and enable optimal use of networks by reducing curtailment problems and the need for new infrastructure. The simple principle behind it is to use various economic incentives to shift electrical loads from times with a high wholesale market price or when the security of the system is threatened to other time periods.

The concept of DR is not new, but was already discussed at the beginning of last century with the deployment of electricity grids, especially with respect to time-differentiated electricity rates. DR started to become popular in the 1970s thanks to two concepts, (1) interruptible load management, mainly for industrial customers, and (2) direct load control for residential customers. Recent papers that explore concepts like the microgrid (Lasseter & Paigi, 2004) and the EnergyHub (Geidl et al., 2006) have put forward the idea of the prosumer, who not only consumes but also produces and participates in the efficiency and efficacy of the grid, while playing an active role in the electricity market. As has been pointed out in (Ma et al., 2017), there are predominantly two types of DR programs for residential buildings, Explicit DR (also called incentive-based) and Implicit DR (also called price-based). In Implicit DR, a price signal is sent to the prosumers such that they can voluntarily modify their consumption and production patterns in accordance with the signal. Explicit DR involves the participation of a third party, who takes action on behalf of a customer by sending a control signal such that the system behaviours are directly modified.

With nearly 40% of the total energy consumption coming from buildings in North America and in Europe, out of it more than 50% comes from thermal energy consumption (Cao et al., 2016), there is definitely a role that buildings can play in DR programs. However, as pointed out by Lawrence et al. (2016), "*the adoption and integration of newer control and communication technologies into buildings can be problematic with older legacy Heating, Ventilation and Air Conditioning (HVAC) and building control systems*". There has been general improvement towards the installation of control devices, communication and control protocols as well as efforts towards standardization but DR is not yet used on a wide scale because regulatory/policy/market framework conditions still need to be defined. Two main barriers that are standing in the way of widespread use of DR in Europe were identified in (Giodano et al., 2013) as being:

1. The absence of clear rules for the rights and obligations of the actors with access to the markets and for the technical validation of flexible demand response transactions on those markets.

2. The need to gain the customer's trust and to encourage their participation in projects by ensuring sensible data protection, robustness of the solutions proposed (no loss of comfort) and providing financial improvement (incentives, lower electricity prices) without putting the customer at financial risk or discomfort.

1.1.3 Definition and usage of flexibility

The term "Energy Flexibility" is used in many different domains related to electric power systems, but it is understood differently by physicists, thermodynamicists or control engineers, which leads to almost as many definitions as papers published. For this reason, Reynders et al. (2018) review and evaluate existing definitions and quantification methodologies applied to flexible energy in buildings. They find that even within this field, flexibility is studied for

different contexts, e.g. the context of energy infrastructure or electricity price. Nevertheless, they find that flexibility is generally defined as the ability to shift energy. A similar definition is proposed by Jensen et al. (2017): "*The Energy Flexibility of a building is the ability to manage its demand and generation according to local climate conditions, user needs, and energy network requirements.*"

The capacity of residential buildings to intentionally modify their consumption patterns to shift energy is affected by several human, technical and external factors: (1) the inhabitant's behaviour and their comfort requirements, (2) the building characteristics such as the thermal capacity and the insulation, (3) the heat distribution system including Power to Heat (P2H) devices (e.g. heat pumps, electric radiators), storage tanks and heat exchangers, (4) their control system and the degree of monitoring, and (5) the weather conditions (Reynders & Wetter, 2015). The thermal demand of buildings changes in response to the external temperature.

A comprehensive overview of how much Energy Flexibility buildings may offer to future power grids is lacking. Based on this observation, a panel of experts initiated the IEA EBC Annex 67 "Energy Flexible Buildings" in 2016 (Jensen et al., 2017). This document will serve in future research projects and commercial applications as a basis to define metrics and methodologies to characterise the Energy Flexibility in buildings.

1.2 Demand Response with Heat Pumps

1.2.1 Basic principle of heat pumps

Heat pumps are devices that transfer heat from a low-temperature "heat source" to another reservoir at higher temperature called the "heat sink" by downgrading a source of a higher exergetic value: either electricity or a higher temperature heat source. In conventional heat pumps using a vapor compression cycle, the compression is performed by mechanical work, i.e. electricity. The thermodynamic cycle of a vapor compression heat pump consists of four components: (1) an evaporator, (2) a compressor, (3) a condenser and (4) an expansion valve (see Figure 1.1). In the thermodynamic cycle, a transfer fluid (refrigerant) circulates to transport the heat from the low temperature source to a higher temperature sink. The refrigerant typically circulates in a closed-loop but in a few systems the refrigerant also circulates in a distribution network (open loop). In the evaporator, the refrigerant in contact with the source evaporates. The mechanical compressor, using electricity in most systems, compresses the refrigerant gas to raise its temperature to the required level of the sink. The refrigerant, that is now a gas at high temperature and pressure, enters the condenser, where it cools down and becomes liquid again. This phase change releases energy which is transferred to the heat sink, raising its temperature. The liquid refrigerant exiting the condenser is still under pressure and needs to pass through an expansion valve to reduce its pressure before entering back into the evaporator. In rare applications, compression can also be driven by gas combustion engines. In adsorption heat pumps, compression is performed by an adsorbent material

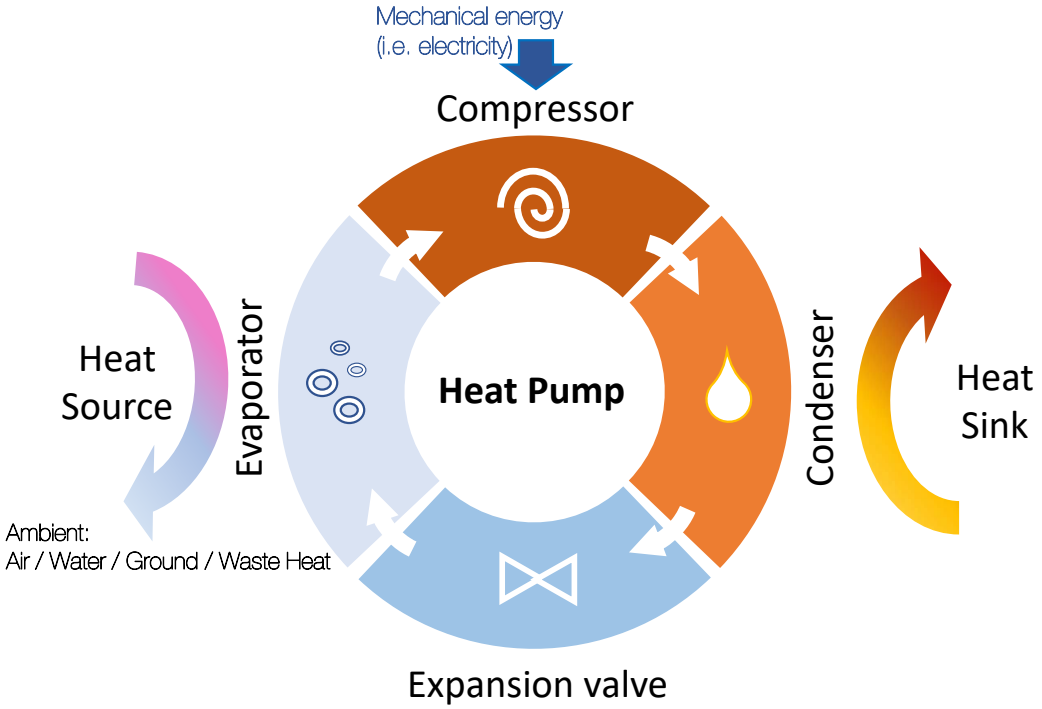


Figure 1.1 – Vapour compression cycle of a heat pump

packed or coated on an adsorbent bed. Details on the working principle can be found in (Dias & Costa, 2018). Heat pumps working on alternative thermodynamic cycles exist for example trans-critical heat pumps.

In this thesis, we only consider the conventional non reversible thermo-electric heat pump type, as they are an efficient technology to upgrade a low temperature source into heat and are increasingly used for residential heat supply. As 75% of the heat produced by the heat pumps is taken from the environment (e.g. air, water and ground), heat pumps can produce more heat with the same amount of primary energy than boilers or electric resistances, leading to a reduction of CO_2 emissions. Thus, there has been a sustained effort to replace boilers and other thermal devices with heat pumps. More details explaining the integration of heat pumps as one of the main component of Europe's future energy system can be found in (Novak, 2018).

1.2.2 Market development in Europe and Switzerland

The European market for heat pumps is expanding quickly with a 12% annual average growth since 2015, as reported by the European Heat Pump Association (EHPA). By the end of 2018, a total of 11.80 million heat pumps were in service in the 21 countries evaluated. In the Swiss market, the popularity of heat pumps has grown from less than 25'000 units present in 1990 to more than 267'000 units in 2018 as presented in Figure 1.3. In 2019, 23'800 heat pumps were installed, representing 48.8% of the total heat-generation devices sold that year in Switzerland

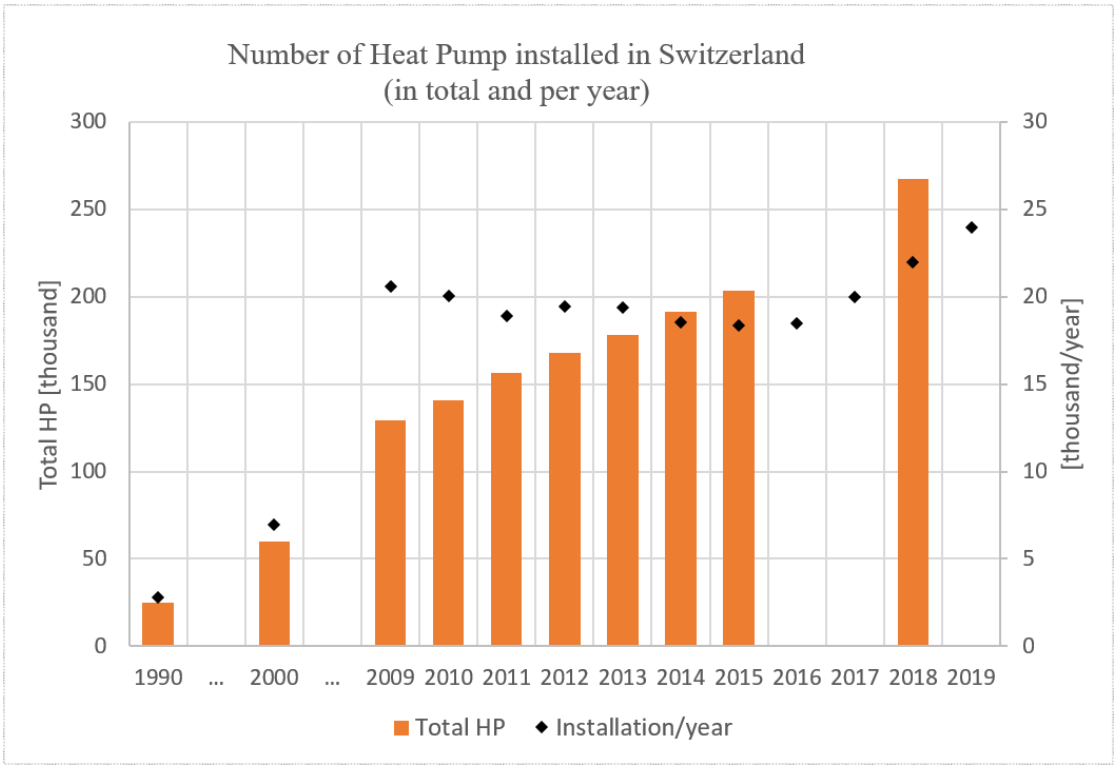


Figure 1.2 – Statistic of the number of heat pumps and their installation rate in Switzerland. Statistics are taken from the Swiss Federal Statistical Office, 2009-2015, 1990, 2000 and from Fachvereinigung Wärmepumpen Schweiz, 2010-2019. Missing statistics for 2016 and 2017

(Peterhans et al., 2019). However, to achieve the climate objectives set by the Swiss federation, around 40'000 heat pumps per year should be installed.

The Swiss Federal Statistical Office (FSO) has reported that in 2000, heat pumps represented 2% of the main energy source for Space Heating (SH) in buildings[I]. In 2017, they represented 17.9 % and this percentage is predicted to exceed 50% by 2050. Heat pumps' recent development is not only limited to SH but also to Domestic Hot Water (DHW) production. As buildings have better and better thermal envelops, more energy is consumed for DHW than for SH. In 2000, the Swiss heating sector was highly dominated by fossil resources (51.2%), followed by electric resistances (39%) and heat pumps were only marginally used (1.8%). In 2017, despite a small increase of natural gas, the share of fossil fuels, as well as electric resistances, has declined to 43.4% and 32.7% respectively, while the share of heat pumps increased to 13.1%

While the type of heat pumps installed in buildings varies depending on the usage, a vast majority of the heat pumps in Europe are installed to provide SH. The exception is countries like Spain and Portugal, where the majority of the heat pumps are installed to satisfy cooling demand. In this thesis, we focus only heat pumps providing SH and DHW and not cooling.

[I] https://www.bfs.admin.ch/bfs/en/home/statistics/construction-housing/buildings/energy-field.html (Online accessed: 2020-07-26)

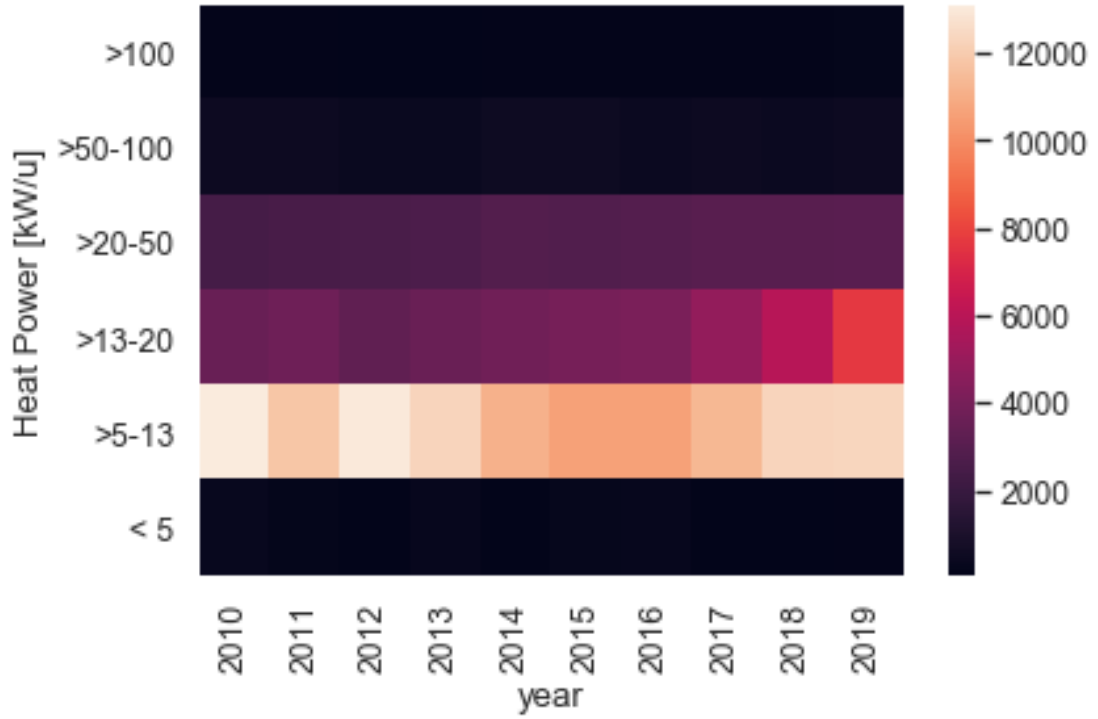


Figure 1.3 – Number of heat pumps installed per year and per range of heat power. Data from Fachvereinigung Wärmepumpen Schweiz, 2010-2019

1.2.3 Heat pump size and properties

Figure 1.3 presents the number of heat pumps installed in Switzerland over the past 10 years (i.e. 2010-2019) as a function of their net heat power. This heat map shows that small systems dominate (i.e. between 5 to 13 kW_{th}), which are mostly installed in single-family houses. This means that important aggregation efforts are needed to reach the power necessary to enter into energy markets. It also highlights a significant increase of heat pumps with a heat power ranging from 13 to 20 kW_{th}, which the typical system size encountered in small multi-family dwellings (less than four residential units).

There exists a wide variety of residential heat pump system layouts, but they can generally be characterised according to three main features: (1) the heat pump itself, (2) the heat supply and (3) the heat storage. Figure 1.4 presents the main possible types for these features in residential systems.

The heat pump properties

The main properties of a heat pump are determined by the type of heat source and the heat sink, and more particularly their temperature difference, as this directly influences the Coefficient Of Performance (COP) of the heat pump. The COP is used to describe the ratio of useful heat movement (i.e. heating or cooling) per work input (e.g. electricity). Variation in

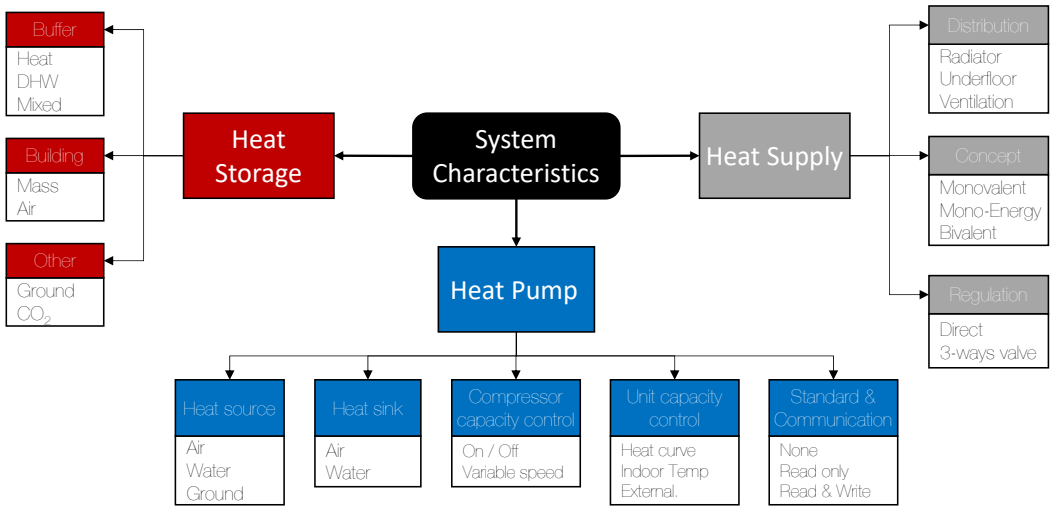


Figure 1.4 – Main distinctive features of residential heat pump systems. Schema adapted from Fischer and Madani, 2017.

the temperature difference between the source and the sink will affect the amount by which the heat pump's electricity consumption can be increased or decreased by ramping the heat pump up or down. Ground and water heat pumps are less affected by seasonal temperature fluctuations than heat pumps operating with air as heat source, and therefore their COP is less affected by the seasons. Figure 1.5 details the heat pump sales in Switzerland over the last ten years in function of their energy source and sink. The majority of heat pumps use air as their heat source, despite the lower efficiency, because air is everywhere. Brine-Water (e.g. ground source) and Water-Water heat pumps represent less than a third of the sales. However, they often have a higher nominal power. In the future, the development of low temperature heating network in cities could result in a higher use of decentralised Water-Water heat pumps.

The compressor capacity control also strongly affects the ramp-up and the maximum run time of the heat pumps. Most existing heat pumps work with ON/OFF compressors, which limits their operation flexibility. Heat pumps have an internal controller that usually steers the system either as a function of a heat curve (see 3.4.2) or by reacting to indoor temperature deviation. In more rare occasions, the heat pump is connected to a building management system with its own control logic. The internal control logic of the heat pump is often not easily accessible, which can play a role in how well a heat pump can respond to DR calls. Last but not least, heat pumps are not equally equipped for communication. The accessibility to the internal controller of the heat pump and the use of specific communication standards determine the control options available. When no communication is available, the only solution is to directly cut the the power supply of the heat pump. With the "`read & write`" option, more complex control solutions for DR can be implemented.

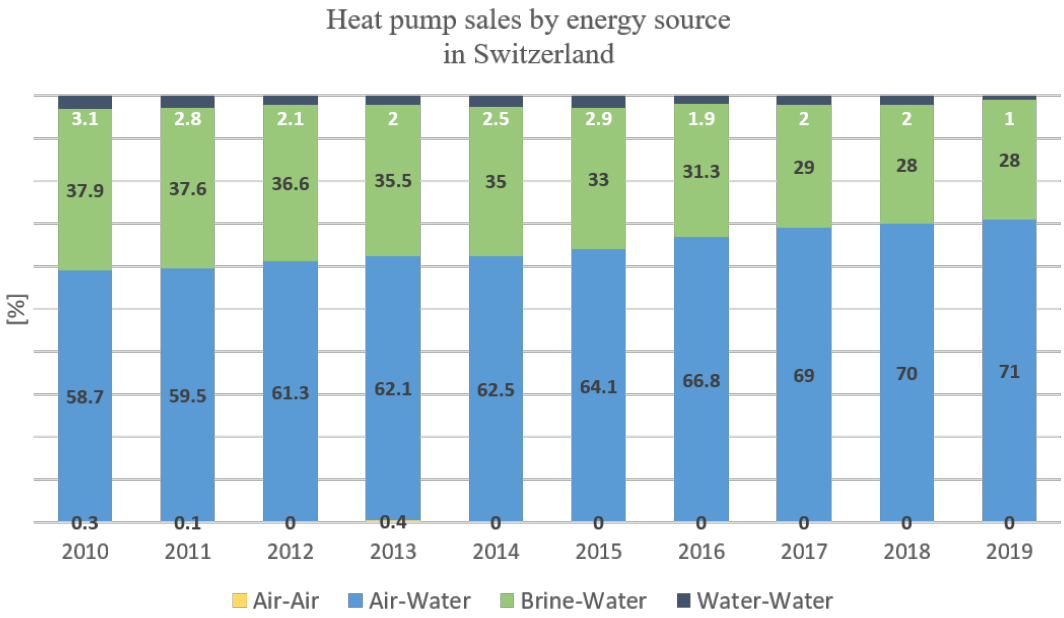


Figure 1.5 – Heat pump yearly sales per energy source. Statistic taken from Fachvereinigung Wärmepumpen Schweiz, 2010-2019.

The heat supply system

The type of heat distribution system is linked to the type of heat sink. Water is used for the heating of DHW, radiators and floor heating systems, whereas air is mostly used in ventilation systems. In residential buildings, the most common distribution systems associated with heat pumps are radiators and floor heating. The operation temperatures of the heat pumps differ depending on the supply system. The highest temperatures are encountered in systems providing DHW (up to 65°C), closely followed by radiators (up to 55°C), while the temperatures for floor heating rarely exceed 35°C. The systems requiring a high maximum temperature also generally display the highest range of temperature and therefore the highest power consumption variations. While the power consumption of heat pumps providing floor heating can be assumed to be fairly constant, the power consumption of heat pumps for DHW or radiators can vary by up to a factor of two. When the entire heat demand of the building is fulfilled by only a heat pump, the system is called a monovalent system. In a bivalent system, the heat pump is supplemented by an extra heat source (e.g. gas, wood). When the additional heat source is electricity, the system is called mono-energy. Finally, the type of regulation of the heat supply will impact how the energy is transferred from the heat pump to the building.

Storage type and size

The storage type and size determines how much energy can be stored and for how long. The energy is stored as heat. The type of storage is strongly linked to the type of heat sink and heat distribution system. The most common storage options are SH and DHW water tanks, (Fischer

et al., 2014) or the thermal inertia of the air or the mass of the buildings directly (Reynders et al., 2013; van der Heijde et al., 2017). In terms of heat capacity, the mass of the building is several orders of magnitude higher than the air or even than the water tanks. Other options exist, including CO_2 networks or seasonal storage options in boreholes (De Ridder et al., 2011), but they are still marginal.

1.3 Overview of building and heat pump control

Various control algorithms, methods and guidelines for building energy systems have been described in the literature. The ASHRAE handbook (ASHRAE, 2019) is a reference of the existing control methodologies for building energy systems and is regularly updated. The simplicity of the design and the low computational complexity required to determine control actions have long limited the range of possible technologies. However, in recent years the progress in communication and computational technologies and the development of small and affordable monitoring devices have enabled the development of new algorithms and control methods for buildings. These methods include techniques such as reinforcement learning with data driven control (Claessens et al., 2018; Costanzo et al., 2016), or more commonly Model Predictive Control (MPC), as found in (Cigler, Siroky, et al., 2013; Killian & Kozek, 2016; Oldewurtel et al., 2012; Vrettos et al., 2013).

Traditionally, the main control task of a HVAC system consists of supplying thermal energy to meet the comfort requirement. Advanced control methods try to perform this task while minimising the cost of operation or maximising comfort. With the development of smart grids, HVAC systems are expected to offer more services than just comfort for inhabitants. Fischer and Madani (2017) present a thorough review of the role heat pumps will play in future power grids. The authors highlight the fact that smart grids will influence the way heat pumps are controlled. Future heat pumps are expected be able to:

- Plan and schedule their operation in advance for maximising auto-consumption (when associated with PV panels or for offering service systems such as DR)
- Change their operation in real time as a reaction to an external signal (e.g. peak shaving, frequency control)

1.3.1 Non-predictive control

Conventionally, the HVAC systems are operated using Rule-Based Controllers (RBC). RBC, is based on rules like “if-then-else” and often manages a specific part of the system. RBC’s often are operated separately with no supervision nor optimisation from an upper layer. This is due to the fact it is impractical to generalise all their rules at the level of building. In old buildings, RBC like dead-band controllers are very common due to their low computational complexity.

Traditionally, heat pumps are seldomly controlled remotely. Studies such as Motegi et al.

(2007) that look at control strategies for thermal devices (e.g. heat pumps, boilers etc.) have identified approaches such as Global Temperature Adjustment (GTA), Passive Thermal Mass storage within the buildings and Supply Air Temperature Increase. Most of these strategies are performed using feedback looping and Proportional-Integral-Derivative (PID) controllers or similar mechanisms with little to no predictive input based on the electricity pricing or weather conditions.

1.3.2 Model predictive control

Model Predictive Control (MPC) is a well-established method for constrained control which was initially developed in the late seventies and early eighties in the process industries (oil refineries, chemical plants, etc.), and has been successfully applied in many areas, both within the research community and in industry (Borrelli et al., 2017; Maciejowski, 2002). One main strength of MPC and a reason for its growing attention in the field of building control is that it easily handles multi-variable control problems with complex objectives, and is able to take actuator limitations into account (Mayne, 2014; Serale et al., 2018).

MPC can be formulated as a set of optimisation problems. The optimisation problem consists of a model describing the dynamics of a system, a set of inequalities accounting for the constraints on the state and the input variables, and an objective to minimise or maximise. The model used in the optimisation captures the dynamics of a system (e.g. its states x). The dynamics can be affected by disturbances (e.g. weather) or by controlled inputs u (e.g. heat supply). The model is used to predict all possible reactions of the system to future disturbances and controlled inputs over a time horizon that starts at the current discretised instant k and spans a prediction horizon Hp. From all these possible solutions, a solver finds the optimal sequence of controlled inputs of length p by minimising or maximising an objective. Contrary to an open-loop controller, the MPC controller only applies the first optimal input u_k over a control horizon Hu, and at the end of it, solves again a new updated optimisation problem to find the next best sequence of input of length p. This principle is called the "receding horizon" and is illustrated in Figure 1.6. The top panel represents the results of the optimisation performed at instant k. The red line represents the history (full line) and the prediction (dashed line) of the state x. The blue line is the history (full line) and the best future sequence (dashed line) of inputs u. The bottom panel represents the next iteration. The instant $k+1$ becomes k and the new state x_k is assessed. Due to errors in the model or uncertainties in prediction, the estimated state $\hat{x}_{k+1}$ can be different than the new state x_k. The previous solution is shown in light grey.

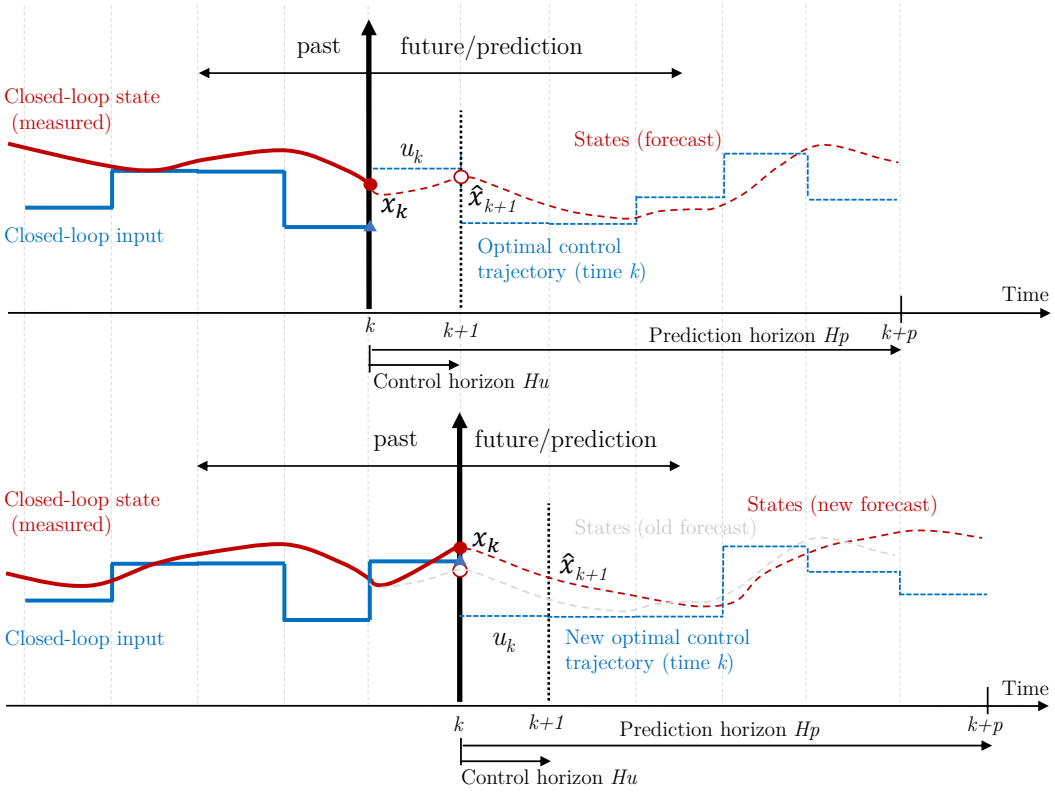


Figure 1.6 – Schematic of Model Predictive Control and the principle of the receding Horizon. The difference between the top and the bottom panel is one iteration of the MPC optimisation

1.4 Real world application

For some time now, Demand Response (DR) has been promoted as a solution that, by providing flexibility, allows to improve the efficiency of the electricity grid with minimal investment on infrastructure. However, DR is currently still rarely implemented in the commercial and residential sector in Europe.

Serale et al. (2018) reveals that only a fourth of the 161 papers on MPC in buildings they review consider residential buildings. Only a bit more than a fifth compared experimental cases to simulated cases.

D'hulst et al. (2015) quantifies the flexibility of residential smart appliances (e.g. washing machines, dishwashers or DHW heaters) based on three years of measurements in 186 households from a large-scale research and demonstration site in Belgium (LINEAR pilot). They show that the flexibility potential is highly asymmetric over a day. More specifically, they show that maximal power increase always strongly exceeds the maximal power decrease of most appliances. Of all wet appliances monitored, DHW buffers present the highest potential for DR services.

The Couperus project is one of the few projects including real test cases for load shifting provided by residential buildings (300 flats) heated by individual heat pumps. Wijbenga et al. (2014) demonstrates within this project that multiple objectives (comfort, user incentives and trade dispatch incentives) can be achieved. By shifting the operation of the heat pumps while strictly respecting thermal comfort constraints, up to 21% of the power used by the heat pump was made flexible to enable wind power absorption.

Kohlhepp et al. (2019) performed a thorough review of recent international field studies testing the application of DR in practice. They reviewed the 16 DR projects including field test and demonstration of DR from around the world. Only 4 projects had more than 100 households, a size large enough to represent load diversity and test resource competition. The authors also highlight the difficulties of comparing the results between projects, as goals and the environment differ and there are no test case benchmarks. In their opinion, advanced control systems are still lacking in large field studies. They concluded that a breakthrough in control systems or market design is not yet to be seen and that field studies with more real buildings are required.

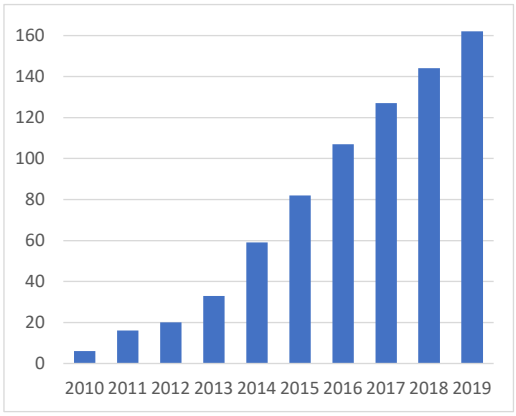
1.5 Knowledge gaps and research questions

The topic of DR within buildings is very relevant and considerable knowledge has been built over the last ten years. Using the search engine "Web of Science", we counted the number of scientific papers published in increasingly specific fields (Figure 1.7). The keywords search in panel (a) are "Demand Response" and "Commercial/Residential buildings". We see a strong increase of interest, with a total of 756 papers published between 2010 and 2019. In panel (b) and (c) we added the keywords "Model Predictive Control" and "Heat Pump", respectively. Although the trend is increasing, the number of publications is very much restricted for both topics. In panel (d) we combine all keywords and were left with a total of only ten papers. This means that the field is new and a lot of things are yet to be discovered. In this introduction we identified several research gaps. We summarise them here and explain how they will be addressed in this thesis.

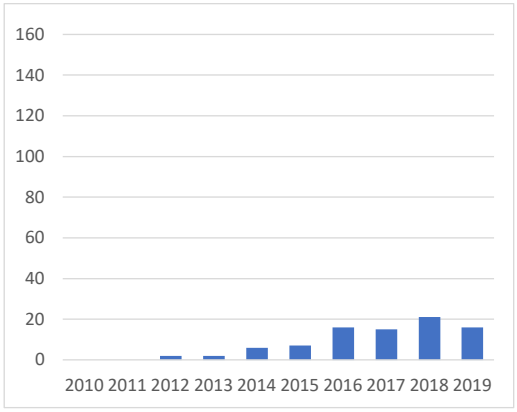
It is known that several factors affect the capacity of buildings to intentionally modify their consumption patterns to shift energy (e.g. user's behaviour, thermal capacity and the heat distribution system). However, it is unclear how important each of these factors are and which need to be further addressed. **Research question:** How important is the influence of factors such as user behavior and physical building properties on the functioning of DR?

The effect on the demand of daily or long term seasonality of external temperature is widely recognised. However, only few studies evaluate this effect for the application of DR services. **Research question:** What is the effect of daily and long term external temperature fluctuations on the functioning of DR?

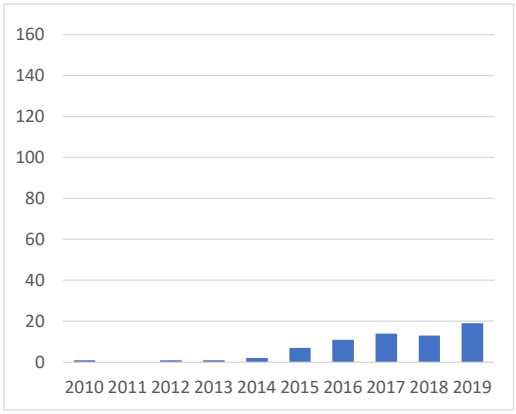
Most research for DR is done for commercial and educational buildings since they are often



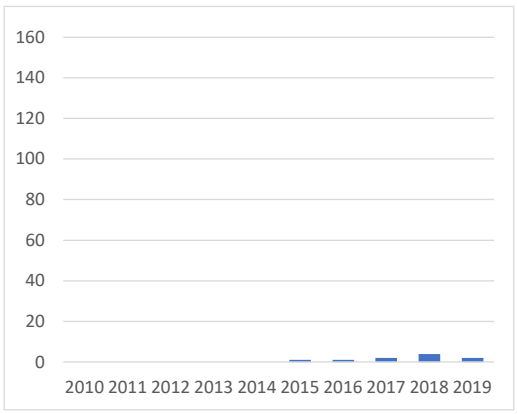
(a) TS_1 = (Building AND "Demand Response" AND (Residential OR Commercial))



(b) TS_2 = TS_1 AND "model predictive control"



(c) TS_3 = TS_1 AND "heat pump"



(d) TS_4 = TS_1 AND "heat pump" AND "model predictive control"

Figure 1.7 – Analysis of number of papers published using the search engine Web of Science and querying the topics given by TS.

well-equipped and have relatively predictable and large power consumption. Residential buildings are usually less well-equipped, individually consume less energy and therefore seem less interesting. However, together residential buildings consume a sufficient amount of energy to be interesting for DR. It is not well studied if DR can be implemented in existing residential buildings. **Research questions:** Can DR be implemented in existing residential buildings? Can residential buildings be aggregated to provide enough flexibility for the participation in an energy market?

Much of the research done for DR in buildings investigates the flexibility potential offered by electric resistances, as they are simple to model and offer more flexibility. However, heat pumps are foreseen as the main heating system devices of the future. Thus, there is a need to analyse their behaviour under smart grid operations. However, their characteristics and control option strongly depends on their usages, their types and their manufacturers. **Research question:**

Can the heterogeneous characteristics of heat pumps be modeled and can heat pumps be remotely controlled?

The flexibility in residential buildings is not free and is a trade-off with inhabitants comfort. It is not clear if the revenue can cover the additional cost of the devices allowing to connect the building to DR programs. **Research question:** Is DR on residential buildings profitable economically?

In the following thesis, we will address our research questions using a pilot site of inhabited residential buildings in Switzerland, which is equipped with Water-Water heat pumps. We will develop and implement our DR framework including hardware and software on the existing installation of different types of Water-Water heat pumps. Using simulated data, monitored data for two full years that covers all seasons, and specific tests, we will assess if DR in combination with heat pumps is a promising set-up to provide flexibility in the future.

2 Model Predictive Control for residential Heat Pumps

In any experimental case study, simulations (even if using simplified models) are necessary during the controller design phase to properly set up the controller parameters (e.g. models and constraints) and ensure reliable performance under different weather conditions or initial states. In order to tap the full potential of model predictive control for exploiting the energy flexibility in residential buildings with heat pumps, the following challenges should be addressed:

- Trade-offs between model accuracy and computational efficiency should be considered in the choice of component models for buffer tanks, the thermal behaviour of buildings and the associated heat and ventilation systems.

- Occupant thermal comfort is a fundamental constraint, but indoor temperatures in residential buildings are not always available, or they are measured in unoccupied spaces. Models of room temperatures cannot be validated directly.

- In residential buildings, only central heat pump and buffer tank control can be acted upon. Local control in apartments cannot be overruled and is largely unknown.

Section 2.1 presents the simulation environment used to properly set up the controller parameters (e.g. models and constraints) and test the reliability and the performance of the algorithm under different conditions (e.g. weather, initial state). Section 2.2 details the framework and the main elements of model predictive control that we will study in the rest of the thesis. The models and the constraints developed for the buildings and the heat pump systems are detailed in Section 2.3. The goals and terms of the objective function are presented in Section 2.4. Section 2.5 describes the modeling of the main sources of disturbances affecting the load patterns of residential buildings. Finally, Section 2.6 summarises the results on flexibility characterisation presented at two conferences: IBPSA 2019 and BAUSIM 2020.

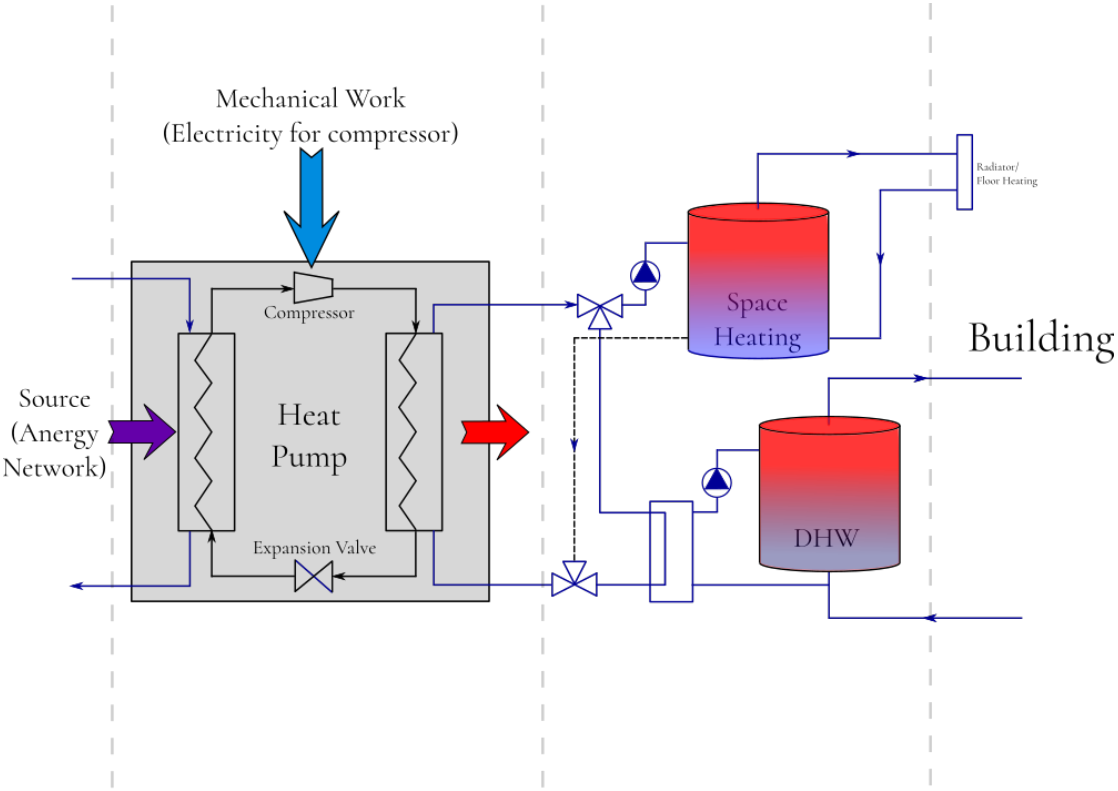


Figure 2.1 – Schematic of the analyzed system at building level.

2.1 Simulation

Simulation environments are often used to test control algorithms. Compared to physical experiments, they have the advantage of being faster, cheaper and safer. Furthermore, different controls can be compared in identical testing conditions (Haves & Xu, 2007; Wetter, 2011).

Dynamic simulations of the cluster of buildings are carried out with MATLAB ® SIMULINK ® using a modified version of the Carnot toolbox (Wemhöner et al., 2000). In order to limit simulation time, buildings are reduced to a one zone model with four nodes: the external mass $T^{w_{ext}}$, the internal mass $T^{w_{in}}$, the floor T^{f} (for floor heating) and the air inside the buildings T^{in}. Space Heating (SH) is delivered to the buildings by radiators or floor heating. A building's heat distribution consists of three loops, which link the building with the buffer tank, the buffer tank with the heat pump and the heat pump with a heat source as displayed in Figure 2.1. This is a typical design for buildings supplied by heat pumps. The storage tank for DHW is linked to the heat pump through an external heat exchanger. The storage tanks for DHW and SH are both modelled with multi layers to account for thermal stratification. The dynamic behaviour of the heat pumps is computed using piece-wise linearisation of the source and sink thermal power and electrical power given by manufacturer data. The outlet temperatures of the condenser and the evaporator are calculated by two differential equations.

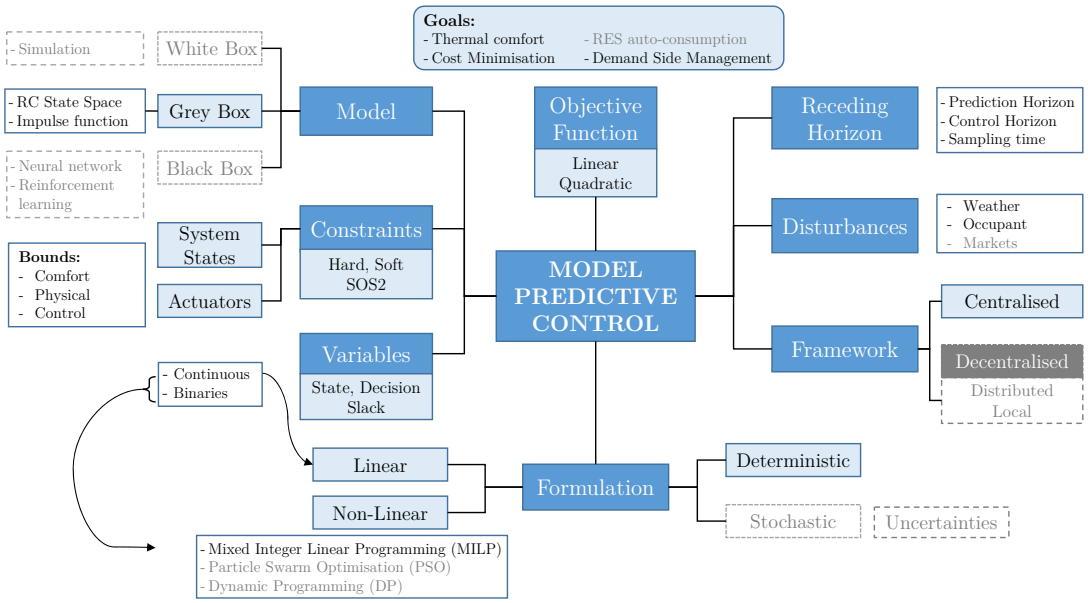


Figure 2.2 – Framework and main elements of Model Predictive Control applied to buildings adapted from Figure 3 in (Serale et al., 2018). The boxes in blue denote the elements that directly influence the MPC. The box or text in grey are not evaluated during this thesis.

2.2 Model predictive control

Model Predictive Control (MPC) is a well-established method for constrained control. One main strength of MPC and a reason for its growing attention in the field of building control are that it easily handles multi-variable control problems with complex objectives, to take actuator limitations into account. Most importantly, MPC is able to explicitly enforce constraints and operate closer to them, allowing to increase performance.

Another advantage of MPC is that it allows for the anticipation and optimisation of the behaviour of a system while taking into account the future disturbances acting upon it (e.g. weather) and its requirements (e.g. thermal comfort). This is possible by using a receding horizon control strategy. This control strategy can be described as the three repeated following steps:

1. Update the state, the parameters and disturbances
2. Solve the optimisation problem defined over a prediction horizon H_p (closed-loop) and obtain the sequence over the N planned control actions, where $N \leq H_p$
3. Implement only the first control action on the system over the control horizon h_u

A general framework of MPC, adapted from the work of (Serale et al., 2018), is presented in Figure 2.2. It details the main elements that usually compose MPC formulations for the

thermal energy management of buildings. There exists a wide variety of models, formulations with different levels of complexity, implementation frameworks or even goals for the objective function, that could not possibly be covered by this thesis. We chose to study a centralised, mixed-integer linear deterministic formulation of a cluster of buildings, including both their thermal behaviour and their heating system. The optimisation problem to solve at each time step is formulated as a Mixed Integer Linear Programming (MILP) problem.

MILP problems consist in the minimisation or maximisation of a linear or quadratic objective function subject to linear inequality and equality constraints, whereby decision variables can be continuous or confined to integer values. Due to its versatility and the availability of efficient solvers, MILP is used in a wide range of applications in engineering and management (Bradley et al., 1977; Jünger et al., 2009). MILP can, for instance, be used to balance energy demand and supply with multiple energy carriers in an optimal way (Evins et al., 2014), and to simultaneously optimise the design and operation of energy conversion components in buildings (Schütz et al., 2017). For MILP formulations, different levels of complexity can be envisioned and require a trade-off between accuracy and computing time.

Four main challenges have to be taken into account when implementing MPC. They are intensively researched (Borrelli et al., 2017).

1. **Feasibility**: The optimisation problem may become infeasible at some future time step, i.e. there does not exist a control sequence satisfying all the constraints
2. **Stability**: Closed-loop stability, i.e. convergence, is not guaranteed
3. **Robustness**: Closed-loop systems are not robust against uncertainties or disturbances, i.e. the optimisation problem is feasible but disturbances drive the system outside the feasible limits
4. **Implementation**: The optimisation problem needs to be solved in "real-time", i.e. within the sampling time of the system

2.3 Model and constraints

2.3.1 Building model

We assume that the building to be controlled can be described by a grey box model using a state space formulation:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_u u(t) + B_z z(t) + Gw(t) \\ y(t) &= Cx(t) + D_u u(t) + D_z z(t) + v(t), \end{aligned} \tag{2.1}$$

where, x is the vector of the state variables (e.g. the building node temperatures, u the vector of control variables (e.g. heat pump power), z is the vector of measured disturbances affecting

the system (e.g. weather, internal gains) and y is the vector of the outputs. The outputs cannot always be measured (e.g inner wall temperature). Equation 2.1 incorporates the uncertainty in the system. The vector w denotes the unmeasured disturbances (i.e. the combined effect of model uncertainty and unmeasured exogenous variables) and the vector v stands for measurement noise. The terms A, B_u, B_z, G, C, D_u and D_z are, respectively, the state matrix, the input matrix, the measured disturbances matrix, the matrix of the effects of unmeasured disturbances on the states, the output matrix and the feedthrough matrix for manipulated and measured disturbances. In a building system, the output y is not a function of the manipulated inputs nor the measured disturbances, which means that the matrices D_u and D_z can be set to zero. In this thesis we chose a deterministic formulation by disregarding uncertainty affecting the system, thus we do not include the stochastic terms $Gw(t)$ and $v(t)$.

The state space model can be discretised, assuming that the inputs u and disturbances z stay constant over the sampling time t_s (i.e. zero-order hold assumption). For the matrices below, the subscripts t and d represent the discrete time steps and discretised version respectively.

$$\begin{aligned} \dot{x}_{t+1} &= A_d x_t + B_{u_d} u_t + B_{z_d} z_t \\ y_t &= C_d x_t, \end{aligned} \tag{2.2}$$

where

$$A_d = e^{A t_s} \tag{2.3a}$$

$$B_d = A^{-1}(A_d - I)\left[B_u \;\middle|\; B_z\right] \tag{2.3b}$$

$$C_d = C \tag{2.3c}$$

In order to keep the formulation as simple as possible while accounting for the main thermal dynamics of the buildings, a single zone model with two capacities C_i and C_m is used. The second-order state space model is used to represent the thermal inertia within the building. Contrary to first-order models, it is able to account for the two main time constants of the building: the fast dynamics of the air $\dot{T}^i$, used for assessing the comfort, and the slow dynamics of the internal mass $\dot{T}^m$, used for assessing the main thermal storage capacity. A physical-based formulation with the main physical parameters describing the dynamics of the building and of the disturbances acting upon it is given in Equation 2.4.

$$\underbrace{\begin{bmatrix} \dot{T}^m \\ \dot{T}^i \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} \frac{-\kappa_{iw}A_{iw}}{C_m} & \frac{\kappa_{iw}A_{iw}}{C_m} \\ \frac{\kappa_{iw}A_{iw}}{C_i} & \frac{-\kappa_{iw}A_{iw}-q_{ie}}{C_i} \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} T^m \\ T^i \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} \frac{1-\alpha}{C_m} \\ \frac{\alpha}{C_i} \end{bmatrix}}_{B_u} \underbrace{\begin{bmatrix} Q^{heat} \end{bmatrix}}_{u} + \underbrace{\begin{bmatrix} 0 & \frac{A_{winds}g}{C_m} & 0 \\ \frac{\kappa_{iw}A_{iw}+q_{ie}}{C_i} & 0 & \frac{1}{C_i} \end{bmatrix}}_{B_z} \underbrace{\begin{bmatrix} T^{ext} \\ P^{sol} \\ G^{in} \end{bmatrix}}_{z} \tag{2.4}$$

The control input u of the system is Q^{heat} and the disturbances z are: the external temperature T^{ext}, the solar irradiance P^{sol} and the internal gains G^{in}. The parameters C_i and C_m are the

Table 2.1 – Main parameters of the building thermal state space model

Description	term	unit
Thermal capacity of the mass (e.g. wall, furniture)	C_m	kJ/K
Thermal capacity of the air in the building	C_i	kJ/K
Overall heat transfer coefficient between indoor air and walls	κ_{iw}	kW/(m^{2}K)
Overall heat transfer coefficient of external walls	κ_{ew}	kW/(m^{2}K)
Overall heat transfer coefficient of windows	κ_{wind}	kW/(m^{2}K)
Area of transfer between indoor air and indoor walls	A_{iw}	m^2
Area of transfer between indoor air and external walls	A_{ew}	m^2
Windows surface area	A_{wind}	m^2
Windows surface area equivalent facing south	A_{winds}	m^2
Fresh air change rate from infiltration (passive)	η	1/h
Percentage of heat recovery from ventilation	rec	-
Solar Heat Gain Coefficient (SHGC)	g	-
The share of heat directly supplied to the air	α	-

total thermal capacity of the air and the mass of the building, respectively. The parameters κ correspond to the overall heat transfer coefficients between the components of the system. They account for the coupled effect of convection and conduction happening between the different components. The parameter $q_{ie} = \kappa_{ew} + \kappa_{wind} + C_i\eta\frac{(1-rec)}{3600}$ is the overall heat flux in [kW/K] between the inside and outside of the building, accounting for the conduction and the passive and forced convection. The parameter α corresponds to the share of heat supplied directly to the air as opposed to the mass. An α value of 1 corresponds to an air ventilation system, while a value of 0 corresponds to a floor heating system. The details of the model parameters can be found in Table 2.1.

The parameters of the RC models are deduced from physical properties of the building, as insufficient monitoring data were available before the third year of the thesis to perform model calibration. However, physics based models tends to perform poorly for low order models. Another solution is to simulate a high order model and to use an order reduction method on it to create a lower order model. In this case, we do not estimate the physical parameters but the elements of the matrices A, B_u and B_z. Therefore we create a "white-box" model in Simulink using the Carnot toolbox. With this simulation of the buildings, we generate time series of temperatures and heat consumption. We then rely on the estimation method "n4sid", combined with the model order reduction method "balred" (on Matlab®), to generate a second order model for each building in the MPC formulation.

2.3.2 Thermal Energy Storage

Here, thermal Energy Storages (TES) refers to water storage tanks present in buildings for Space Heating (SH) and Domestic Hot Water (DHW) purposes. Storage tanks can be modelled as perfectly mixed volumes with homogeneous temperature or as a set of several layers, either

with constant volumes and varying temperatures or with constant temperatures and varying volumes (Schütz, Streblow, et al., 2015; Steen et al., 2015). Schütz, Harb, et al. (2015) compared different models for TES and heat pump combinations to be used in a MILP formulation. They compared a simple TES model (e.g. fully mixed tank) with a model capable of dealing with thermal stratification. The stratified TES model is modeled as four homogeneous layers with passive and active convection among them and losses with the ambient. They found that stratified models led to no comfort violation but their computing time was more than a hundred times higher than for simple TES model. Operating costs were also marginally improved (i.e. less than 4%). Simple TES model offer a good compromise between accuracy and computing time, which is why we chose them for the MPC.

Thermal energy storages for SH are modelled as a fully mixed tank model. This means that the temperature in the tank is assumed to be homogeneous. The change in SH temperature T^{sto} per time step dt is calculated as the difference between the heat supply by the heat pump $\dot{Q}^{hp,sh}$ and the heat given to the building $\dot{Q}^{house}$ through the distribution system. Thermal losses to the storage room T^{amb} are also accounted for. The resulting energy balance is given in equation 2.5:

$$C_n^{sto} \frac{T_{n,t+1}^{sto} - T_{n,t}^{sto}}{dt} = hA_n^{sto}\left(T_n^{amb} - T_{n,t}^{sto}\right) + \dot{Q}_{n,t}^{hp,sh} - \dot{Q}_{n,t}^{house}, \tag{2.5}$$

where C^{sto} is the heat capacity of the storage, h the heat transfer coefficient in [W/(m^{2}K)] and A^{sto} the surface area of the tank. n and t are the building index [0, ..., N] and time index [0, ..., h_p], respectively. The buffer's temperature is bound by a lower value ($T^{sto,min}$) and an upper value ($T^{sto,max}$), which are defined as hard constraints. The lower bound ensures thermal comfort of the heating system. The upper bound represents the maximum temperature of the buffer, which often corresponds to the maximum temperature that the heat generator can reach. The lower bound $T^{sto,min}$ is defined for each building n. It is also time dependent, as thermal comfort requirements vary within days or seasons (e.g. night setback, winter to summer transition).

TES for Domestic Hot Water (DHW) are also modelled as a fully mixed tank model. This is a strong assumption, as DHW buffers are in reality strongly stratified with a thermocline. A thermocline is a thin but distinct layer in which temperature changes more drastically with depth than it does in the layers above or below. This results from a high temperature difference between the water in the storage and the cold water entering at the bottom. The energy balance for the DHW storage tank is given in equation 2.6:

$$C_n^{dhw} \frac{T_{n,t+1}^{dhw} - T_{n,t}^{dhw}}{dt} = hA_n^{dhw}\left(T_n^{amb} - T_{n,t}^{dhw}\right) + \dot{Q}_{n,t}^{hp,dhw} + \dot{Q}_{n,t}^{el} - \dot{Q}_{n,t}^{dhw}, \tag{2.6}$$

where C^{dhw} is the heat capacity of the storage, h the heat transfer coefficient in [W/(m^{2}K)] and A^{dhw} the surface area of the DHW storage. $\dot{Q}^{hp,dhw}$ is the heat supply by the heat pump to the DHW storage tank. In this case we neglect the losses in the heat exchanger. $\dot{Q}^{el}$ is the

heat supply by the back-up electrical heater while $\dot{Q}^{dhw}$ is the DHW consumption. The DHW storage temperature is also bound by a lower value ($T^{dhw,min}$) and an upper value ($T^{dhw,max}$). The lower value is often set to a smaller value than the typical DHW comfort requirements (e.g. > 45°C) as the average temperature of the storage tank is smaller than the monitored one due to the stratification effect. As opposed to the lower bound for the SH buffer, the one for the DHW storage is not subjected to time variation.

2.3.3 Cyclic constraints

Terminal constraints and costs are often used to ensure the stability of the system being controlled. Cyclic constraints account for the high daily seasonality of building systems (e.g. weather disturbances). It ensures that the optimisation does not benefit from "free energy" by bounding the new optimised next state ($t = 1$) with the terminal state ($t = p$). It can be either a percentage of the state defined as a hard constraint (see Equations 2.7 and 2.8) or it can be associated with a terminal cost. In this case, it is formulated as a soft constraint (Equation 2.9) with a slack variable $W^{term} > 0$.

$$T^{sto}_{n,t=p} \geq 0.85\ T^{sto}_{n,t=1} \tag{2.7}$$

$$T^{dhw}_{n,t=p} \geq 0.85\ T^{dhw}_{n,t=1} \tag{2.8}$$

$$T^{m}_{n,t=p} \geq T^{m}_{n,t=1} - W^{term}_{n} \tag{2.9}$$

However, these constraints can sometimes be counter-productive by forcing the solver to discard valid solutions. This is especially the case when the prediction horizon is shorter than the seasonal period. If the solver has trouble converging within the time constraints imposed or if the problem becomes infeasible, we relax these constraints by either suppressing them or reducing the multiplication factor (here equal to 0.85).

2.3.4 Heat pump models and constraints

Heat pumps are more complex to model than electric resistances for two main reasons. First, most of them can not be operated at variable speed (i.e. the compressor is either ON or OFF). This introduces binary variables which convert a Linear Problem (LP) into a MILP problem. MILP problems are considered to be NP-complete. Second, their electricity consumption and thermal output can vary with different operating temperature. Some authors disregard the second issue by considering the Coefficient Of Performance (COP) of a heat pump to be constant (Madsen et al., 2012). This can be a valid assumption if operating temperatures are fairly constant over the prediction horizon. This is the case when the heat pump uses a heat source at constant temperature (e.g. the ground or a low temperature thermal network) and has a heat sink at a fairly constant temperature (e.g. floor heating). In this case, the nominal values for the heat sink $Q^{sink,nom}$ and electricity consumption $P^{hp,nom}$ are defined *a priori*

and used in the following equations:

$$\forall n, t: \quad P^{hp}_{n,t} = S^{hp}_{n,t}\, P^{hp,nom}_{n,t} \tag{2.10}$$

$$\forall n, t: \quad Q^{hp}_{n,t} = S^{hp}_{n,t}\, Q^{hp,nom}_{n,t}, \tag{2.11}$$

where S^{hp}_t represents the heat pump state (i.e. ON or OFF) at time t.

In reality, most heat pumps do not have a constant COP. The electric power consumption and the thermal power delivered by a heat pump is a function of the source and the sink temperatures of the heat pump (see Figure 1.1). The heat pump's electric power consumption for DHW production (i.e. T>50°C) can be up to twice as high as SH production for floor heating (i.e. T<35°C). However, this function is non-linear. To reduce complexity, and avoid high solving time, this function is linearised in two steps. As a first step we set the source temperature T^{srce} as a parameter of the model. Indeed, in the case of Water-Water heat pumps the source temperature (e.g. the ground or a low temperature thermal network) can be assumed constant during the full prediction horizon. Temperature variations can happen, but more on a seasonal basis. In a second step, the non-linear functions $P^{hp} = f(T^{sink})$ and $Q^{sink} = f(T^{sink})$ are broken down into piecewise linear functions. The breakpoints are defined as the operating temperatures found in heat pump manufacturer datasheets. They correspond to the temperatures at which heat pumps are typically tested during the commissioning phase (e.g. 35°C, 45°C, 55°C and 60°C). Each of these temperatures has an associated power consumption and heat production (e.g. $P^{hp,35}$). We then add weighting variables w_l to the breakpoints.

$$T^{sink} = w_1 T^{min} + w_2 T^{35} + \ldots + w_L T^{max} \tag{2.12a}$$

$$f(T^{sink}) = w_1 P^{min} + w_2 P^{hp,35} + \ldots + w_L P^{max} \tag{2.12b}$$

$$1 = w_1 + w_2 + \ldots + w_L \tag{2.12c}$$

However, these equations are not sufficient, as they do not enforce that at most two neighbouring weighting variables can be positive if we want to respect the piecewise linear lines. For this reason, we assign a new binary variable for each interval (i.e between two breakpoints) and we constrain the weights w as described in Equations 2.13.

$$w_1 \leq \delta_1 \tag{2.13a}$$

$$w_2 \leq \delta_1 + \delta_2 \tag{2.13b}$$

$$w_3 \leq \delta_2 + \delta_3 \tag{2.13c}$$

$$\vdots$$

$$w_L \leq \delta_L \tag{2.13d}$$

$$1 = \delta_1 + \delta_2 + \delta_3 + \ldots + \delta_L \tag{2.13e}$$

Some solvers (e.g. Gurobi, CPLEX) provide a way to implement the previous set of equations

as a Special Ordered Set of type 2 (SOS2). Using the solver's own implementation has two advantages. First, we do not have to encode the set of equations ourselves, as they are automatically generated for us by the solver. Second, the solver's own implementation keeps the order in memory which accelerates the search for an optimal solution (when performing the branch-and-bound/branch-and-cut algorithm).

The heat pumps are modelled by adapting the equations used by (Schütz, Harb, et al., 2015) to accommodate for the additional constraints imposed by the local controllers of the heat pumps. A main contribution of this thesis is the development of a formulation that accounts for the production of both SH and DHW by a same heat pump. We also propose a solution of the problem encountered with a piecewise linear formulation of heat pump electric power consumption in the presence of negative electricity prices or when used in a MPC tracking problem. A negative price is when instead of paying for electricity you consume, you get paid to consume it. The state S of the heat pump and the breakpoint weights w are constrained by:

$$\forall n,t: \quad S^{sh}_{n,t} + S^{dhw}_{n,t} - S^{hp}_{n,t} \leq 0 \tag{2.14a}$$

$$\forall n,t,l: \quad S^{sh}_{n,t} + S^{dhw}_{n,t} = \sum_{l=1}^{L} w_{l,n,t} \tag{2.14b}$$

$$\forall n,t,l: \quad \text{SOS2}(w_{l,n,t}) \tag{2.14c}$$

The terms S^{sh}, S^{dhw} and S^{hp} are binary variables representing the state of SH, DHW and of the heat pump itself. If $S^{hp} = 0$, the heat pump is switched OFF and equation 2.14a forces S^{sh} and S^{dhw} to also be equal to zero. The weights $w_{l,n,t}$ are continuous variables [0 1], with l the index for the number of weights. For the rest of the thesis we will assume that L is equal to 4. In equation 2.14b the sum of the weights is equal to the state of the heat pump. The weights are constrained by the temperature level of the storage tanks using a big-M constraint formulation. Big-M constraints are typically used to propagate the implications of a binary, ON-OFF decision (e.g. the state of the heat pump) to a continuous variable (e.g. here the storage temperatures). They can be a source of instability for optimisation problems and have to be used carefully. In general we want the smallest value of M to be large enough to avoid cutting off a valid solution (e.g. here the maximum storage temperature $T^{sto,max}$).

$$\forall n,t: \quad T^{sto}_{n,t+1} - \left(1 - S^{sh}_{n,t}\right) T^{sto,max}_{n} \leq w_{1,n,t} + w_{2,n,t} + w_{3,n,t} + w_{4,n,t} \tag{2.15a}$$

$$\forall n,t: \quad T^{dhw}_{n,t+1} - \left(1 - S^{dhw}_{n,t}\right) T^{dhw,max}_{n} \leq w_{1,n,t} + w_{2,n,t} + w_{3,n,t} + w_{4,n,t} \tag{2.15b}$$

Equations 2.15a-2.15b ensure that the heat pump's power is correctly related to the activated storage tank and bound by its temperature level. If the binaries S are equal to zero, the left hand side of the equation becomes negative. The heat sink Q^{hp} and the electricity consumption P

are finally given by:

$$\forall n,t,l: \quad Q^{hp}_{n,t} = \sum_{l=1}^{L} Q_l w_{l,n,t} \tag{2.16a}$$

$$\forall n,t: \quad P_{n,t} = \sum_{l=1}^{L} P_l w_{l,n,t} \tag{2.16b}$$

$$\forall n,t: \quad Q^{hp}_{n,t} = Q^{sh}_{n,t} + Q^{dhw}_{n,t} \tag{2.16c}$$

$$\forall n,t: \quad Q^{hp}_{n,t} \geq 0, \tag{2.16d}$$

Equation 2.16c ensures that the heat pumps do not provide heat to both the SH buffer and the DHW storage tank.

Heat pumps are more efficient at low sink temperature. The equations 2.16b and 2.15a-2.15b bound the solution by choosing the weight w as small as possible, resulting in a fairly accurate representation of the physical operating limitation. However, when electricity prices become negative, which can sometimes happen in electricity markets due to the increased share of renewable energy sources, the equations are no longer bounded, as the optimisation will converge to solutions maximising P and as a consequence w_4 as well. Although these solutions are feasible for the solver, they wrongly assess the real consumption of the heat pumps. This is also the case when the goal of the objective function is to track a reference signal. In this particular case, deviations are often more penalised than the electricity cost which results in the same issue. In order to tackle this problem, we add a second set of constraints to the ones presented in 2.15a-2.15b:

$$\forall n,t: \quad T^{sto}_{n,t+1} - \left(1 - S^{sh}_{n,t}\right) T^{sto,max}_{n} \geq w_{1,n,t} + w_{2,n,t} + w_{3,n,t} + w_{4,n,t} \tag{2.17a}$$

$$\forall n,t: \quad T^{dhw}_{n,t+1} - \left(1 - S^{dhw}_{n,t}\right) * T^{dhw,max}_{n} \geq w_{1,n,t} + w_{2,n,t} + w_{3,n,t} + w_{4,n,t} \tag{2.17b}$$

These two equations ensure that the electricity consumption of the heat pump is linked to the storage tank temperature and is not overestimated.

2.3.5 Heat distribution system constraints

Many studies define the heat delivered to the house as an input variable and discard the heat distribution system from the formulation. This is valid for electric resistances which are often used in building control studies due to their simple modeling and control (Gorecki et al., 2017). Others focusing on the HVAC system consider the building energy consumption as a model disturbance (Schütz, Streblow, et al., 2015). They use a predefined forecast profile. Thus, storing thermal energy in the building is not possible, as the energy flux to the building is already known. This type of formulation is not interesting if we want to evaluate the thermal flexibility of the residential buildings to provide DR services. In this case we want to control the amount of heat we provide to the building. For radiator or floor heating based systems, the SH thermal flux depends on the product of the water mass flow rate and the temperature

delta given in Equation 2.18a.

$$\forall n,t: \quad Q_{n,t}^{house} = \dot{m}_{n,t}^{house} c_p \left(T_{n,t}^{flow} - T_{n,t}^{ret} \right) \tag{2.18a}$$

$$\forall n,t: \quad T_{n,t}^{flow} \geq T_{n,t}^{ret}, \tag{2.18b}$$

where $\dot{m}_t^{house}$ is the mass flow rate of the building's heat distribution system, T_t^{flow} and T_t^{ret} are departure and return temperatures, respectively. However this equation contains a non-linear product between the mass flow rate and the temperatures. Here, we assume that the circulation can be either ON or OFF. When a heat pump is switched OFF (i.e. is not supplied by electricity), the heat distribution is also switched OFF and the mass flow is equal to zero. Otherwise, the mass flow is equal to the nominal mass flow. We introduce new variables θ^{flow} and θ^{ret} that represent the product of the heat pump's current state S^{hp} and the supply and return temperature. According to Schütz, Harb, et al. (2015), these variables are constrained by Equations 2.19a-2.20d.

$$\forall n,t: \quad \theta_{n,t}^{flow} \leq S_{n,t}^{hp} \, T_{n,t}^{flow,max} \tag{2.19a}$$

$$\forall n,t: \quad \theta_{n,t}^{flow} \geq S_{n,t}^{hp} \, T_{n}^{flow,min} \tag{2.19b}$$

$$\forall n,t: \quad T_{n,t}^{flow} - \theta_{n,t}^{flow} \leq \left(1 - S_{n,t}^{hp} \right) T_{n,t}^{flow,max} \tag{2.19c}$$

$$\forall n,t: \quad T_{n,t}^{flow} - \theta_{n,t}^{flow} \geq \left(1 - S_{n,t}^{hp} \right) T_{n}^{flow,min} \tag{2.19d}$$

Here, $T_{n,t}^{flow,max}$ is a function of the minimal setpoint settings and the average external temperature bounded by $T^{flow,min}$ and $T^{flow,max}$.

$$\forall n,t: \quad \theta_{n,t}^{ret} \leq S_{n,t}^{hp} T^{ret,max} \tag{2.20a}$$

$$\forall n,t: \quad \theta_{n,t}^{ret} \geq S_{n,t}^{hp} T^{ret,min} \tag{2.20b}$$

$$\forall n,t: \quad T_{n,t}^{ret} - \theta_{n,t}^{ret} \leq \left(1 - S_{n,t}^{hp}\right) T^{ret,max} \tag{2.20c}$$

$$\forall n,t: \quad T_{n,t}^{ret} - \theta_{n,t}^{ret} \geq \left(1 - S_{n,t}^{hp}\right) T^{ret,min} \tag{2.20d}$$

$$\tag{2.20e}$$

The return temperature T^{ret} is defined as:

$$T_{n,t}^{ret} = \beta 1_n T_{n,t-1}^{flow} + \beta 2_n T_{n,t}^{flow} + c_n, \tag{2.21}$$

where $\beta 1$, $\beta 2$, c_n are parameters fitted based on the RMSE between the modeled heat supply to the house Q_n^{house} and the heat supplied by the heat pump Q_n^{sh}. Finally, Equation 2.18a is reformulated with the new variables θ as:

$$Q_{n,t}^{house} = \dot{m}_n^{house} c_p \left(\theta_{n,t}^{flow} - \theta_{n,t}^{ret} \right), \tag{2.22}$$

2.3.6 Local controller constraints

MPC is considered as a higher order control algorithm and is usually implemented on top of traditional local controllers (e.g. PID or valve actuators). MPC controllers can be easily plugged into a local control by supplying it with set-points and they can also be disabled in case of a misbehaviour without impacting the system. Indeed, in most cases, the local building controllers with previous default setpoints are stable. A majority of papers include both the building and the HVAC system (Serale et al., 2018) but it remains unclear if local control for buildings should or should not be integrated into the MPC formulations. For a simple HVAC system (e.g. electric radiators), this over-constrains the problem and removes the solver's freedom to explore new solutions. However, for more advanced HVAC systems (e.g. heat pumps) it might be necessary to ensure that the optimised strategy is applicable on the real system. Many studies evaluating the flexibility offered by buildings discard the control aspect based on the assumption that "the system providing the flexibility is smart in a manner that it is able to respond to an external penalty signal" (Junker et al., 2018). In this thesis, we include the local control aspect, since in the case of heat pumps, most systems cannot be forced but need to be encouraged to react to penalty signals.

Conventional heat pump controllers typically use heat curves associated with dead-band controls on the system temperatures (e.g. the storage tanks and the distribution) to maintain the energy level within comfort bounds. A heat curve is a program that determines the SH supply temperature setpoint $T^{flow,set}$ as a function of the external temperature. For more details, refer to section 3.1.2. The supply temperature T^{flow} is bound by $T^{flow,set}$ and T^{sto} as formulated in the following constraints.

$$\forall n,t: \qquad T^{flow}_{n,t} = \min\left[T^{flow,set}_{n,t}, T^{sto}_{n,t}\right], \tag{2.23}$$

where min is a predefined function of the gurobi solver. Equation 2.23 ensures that T^{flow} cannot be larger than T^{sto} even when $T^{flow,set}$ is larger than T^{sto}.

The main control variable in our formulation is the SH departure temperature setpoint $T^{flow,set}$. We allow it to deviate from its default value by introducing two new positive variables δ^{up} and δ^{dn}. These variables are associated in the objective function with a linear penalty cost to avoid too many changes. $T^{flow,set}$ is constrained by the following equations:

$$\forall n,t: \qquad T^{flow,set}_{n,t} + \delta^{up}_{n,t} - \delta^{dn}_{n,t} = T^{flow,set,def}_{n,t} \tag{2.24a}$$

$$\forall n,t: \qquad T^{flow,set}_{n,t} \geq T^{flow,set,min}_{n,t} \tag{2.24b}$$

$$\forall n,t: \qquad T^{flow,set}_{n,t} \leq T^{flow,set,max}_{n,t}, \tag{2.24c}$$

where, $T^{flow,set,def}$, $T^{flow,set,min}$, $T^{flow,set,max}$ are time series of the heat curve precomputed according to Equation 2.25. Equation 2.25 shows the formulation for $T^{flow,set,def}$ but we

follow the same logic for the min and max heat curves.

$$\forall n,t: \qquad T_{n,t}^{flow,set,def} = s_n * \max\left(\Delta_{lim}, T_{n,t}^{in,set,def} - \overline{T}_t^{ext}\right) + T_{n,t}^{in,set,def} + L_n, \tag{2.25}$$

where s is the slope of the curve and L is the level. Δ_{lim} is the minimal temperature difference, between the default indoor setpoint temperature $T^{in,set,def}$ and the external temperature. Below this temperature difference, the heat pump is set to summer mode (i.e. shut down for SH). $\overline{T}_t^{ext}$ is the predicted average external temperature over a window of 3 to 12 hours, depending on the heat pump default control logic. By adapting the heat curve in Equation 2.25, we can constrain the problem to force the heat pump state S^{hp} to be equal to 0 (OFF) when the difference between the indoor temperature setpoint and the 12 hours average external temperature is smaller than Δ_{lim}.

$$\forall n,t: \qquad \frac{T_{n,t}^{flow,set} + s_n \overline{T}_t^{ext,3h} - L_n}{s_n - 1} - \overline{T}_t^{ext,12h} + \left(S_{n,t}^{hp} - 1\right)M \geq \Delta_{lim} \tag{2.26}$$

Heat pump local controllers also rely on dead-band control. A dead-band limits the start-up and shut down of a heat pump outside specific temperature bounds. These bound limits are usually not fixed but vary with weather conditions or simply over time (e.g night setback). We integrate these dead-band limits in the MPC formulation as constraints, as they reflect the actual heat pump's heating ability versus what is theoretically possible. Equations 2.27a-2.27b represent the constraints for the lower bound, while 2.28a-2.28b represent the constraints for the upper bound. We use a Big-M constraints formulation to ensure that the equation holds for a specific temperature only if the sum of the binary variables takes on a specific value.

$$\forall n,t: \qquad \left(T_{n,t}^{sto} - T_{n,t}^{flow,set}\right) - \left(S_{n,t-1}^{sh} - S_{n,t}^{sh} + S_{n,t}^{dhw} + 1\right)M \leq -\Delta_{dwn} \tag{2.27a}$$

$$\forall n,t: \qquad \left(T_{n,t+1}^{sto} - T_{n,t}^{flow,set}\right) + S_{n,t}^{sh} M \geq -\Delta_{dwn} - 1 \tag{2.27b}$$

Here M represents a big number ($M = 500$) for the big-M formulation. Equation 2.27a ensures that the heat pump can change its state from OFF ($S_{n,t-1}^{sh} = 0$) to ON ($S_{n,t}^{sh} = 1$) only if the storage temperature is smaller than the setpoint temperature by Δ_{dwn}. In any other combination of binaries, the big-M makes the right-hand term value very negative and the equation is non-bounding. Note that DHW has priority over SH, thus the binary S^{dhw} ensures that the equation holds when the heat pump runs a DHW cycle although the storage temperature is below the lower bound. Changing the state is not the only option for the solver to satisfy these inequalities: the value of $T^{flow,set}$ can also be changed. Decreasing its value will allow the equation to hold even when T^{sto} is low.

$$\forall n,t: \qquad \left(T_{n,t}^{sto} - T_{n,t}^{flow,set}\right) + \left(-S_{n,t-1}^{sh} + S_{n,t-1}^{sh} + 1\right)M \geq \Delta_{up} \tag{2.28a}$$

$$\forall n,t: \qquad \left(T_{n,t+1}^{sto} - T_{n,t}^{flow,set}\right) - \left(1 - S_{n,t}^{sh}\right)M \leq \Delta_{up} \tag{2.28b}$$

Equation 2.28a ensures that the heat pump can change its state from ON ($S_{n,t-1}^{sh} = 1$) to OFF ($S_{n,t}^{sh} = 0$) only if the storage temperature is larger by a Δ_{up} than the setpoint temperature.

For any other combination of binaries, the big-M turns the right-hand term into a very large positive number and the equation is non-bounding. Again, changing $T^{flow,set}$ can also make the equation hold. Equation 2.28b avoids that the storage temperature rises too high above the dead-band control limits. If the storage temperature at $t+1$ is larger than the setpoint at t by Δ_{up} the variable S_t^{sh} is forced to be zero.

However, adding these constraints has a direct impact on the tractability, convergence rates and resolution time required for solving the problems. The addition of these constraints has increased the convergence times from a few seconds for each time step to anything from a few minutes to the maximum allowed 10 minutes (as the electricity balance markets use 15 minutes as the actuation time limit). The 10-minute time limit and the additional constraints make a problem that includes multiple heat pumps difficult to solve for optimality. As a result, this formulation has a low scalability for more buildings or longer prediction horizons. A possible solution to this problem is to use a split formulation, for which the satisfaction of the local constraints is only enforced for the first part of the prediction horizon. For the rest of the prediction horizon those constraints are relaxed.

2.4 Objective function

The objective function, also called the cost function, is the mathematical translation of the goal that the MPC tries to achieve. Different types of objective functions can be used in the MPC control algorithm from the minimisation of a single term, to a trade-off between objectives (e.g. cost and comfort). Tuning the specific parameter settings for each type of objective function has proven to be challenging. In this section we present the two types of objective function used for MPC providing DR services.

2.4.1 Economic MPC

The MPC objective function in Equation 2.29 is a representation of all operational costs, including the electricity price, the discomfort and a penalty for control action regulation. The optimisation problem includes N buildings over a prediction horizon p is formulated as:

$$\min \sum_{n=1}^{N} \sum_{t=t_0}^{p-1} \left[\lambda^{el} P_{n,t} + \lambda^{tc} \Delta T_{n,t}^{tc} + \lambda^{act} \Delta u_{n,t} \right] \tag{2.29}$$

Subject to:	$x_{t+1} = A_d x_t + B_{u,d} u_t + B_{z,d} z_t$	(2.29a)
$\forall n, t$:	$\Delta T_{n,t}^{tc} \geq T_{n,t}^{min} - T_{n,t}^{in}$	(2.29b)
$\forall n, t$:	$\Delta T_{n,t}^{tc} \geq T_{n,t}^{in} - T_{n,t}^{max}$	(2.29c)
$\forall n, t$:	$\Delta T_{n,t}^{tc} \geq 0$	(2.29d)

The problem is subjected to the constraints presented in the previous subsections. The terms λ are weighting parameters that drive the solution towards different objectives, thus allowing to choose between conservative or more aggressive control strategies. In this formulation, the electricity price λ^{el} is considered constant (0.18 cts/kWh) over the optimisation period. The cost function also includes the level of discomfort $\Delta T^{tc}_{n,t}$, described in Equations (2.29b-2.29d), with an associated penalty cost λ^{tc}. It ensures that the state variables describing the air and mass temperature of the building (see Equation 2.4) as well as the DHW storage and buffer tanks (see Equations 2.5-2.6) stay within the comfort limits given by $T^{min}_{n,t}$ and $T^{max}_{n,t}$. Finally, $\Delta u_{n,t}$ quantifies the change of the manipulated variables (control variables). In order to penalize frequent control actuation that could alter the system performances, a cost λ^{act} is associated with $\Delta u_{n,t}$. The higher the cost, the less actuation is permitted. The control actuation can be either the binary variables S describing the state or the setpoint variables if the local control is included in the MPC formulation. Cigler, Siroky, et al. (2013) highlighted on a real building system that an economic MPC can result in a poor operation with high oscillations of control setpoints. They compared different cost function formulations that allow to smooth out the system reaction by introducing quadratic penalties. However, these formulations do not allow for sharp changes in behaviour that DR services would benefit from.

2.4.2 Tracking MPC

When offering DR services, a group of devices can be asked to follow a reference power load. To do so, the MPC formulation has to be slightly updated into a so-called "tracking MPC". In tracking MPC, the objective is to minimise the unconstrained error between a given reference R and the output, in this case the clustered power (i.e. $\sum_{n=1}^{N} P_n$). The difference can be added to the objective function (Equation 2.29) as a linear or a quadratic least squares term, which results in equation 2.30

$$\min \sum_{n=1}^{N} \sum_{t=t_0}^{p-1} \left[\lambda^{el} P_{n,t} + \lambda^{tc} \Delta T^{tc}_{n,t} + \lambda^{act} \Delta u_{n,t} \right] + \sum_{k=t_0}^{t_0+d_{DR}} \lambda^{tr} \left(\sum_{n=1}^{N} P_{n,k} - R_k \right)^2 \tag{2.30}$$

The penalty for deviations, $\lambda^{tr} >> \lambda^{el}$, drives the objective function towards minimising the power deviation from the target profile R for each time step over the DR service horizon. Often, the inconsistency in the formulation of the model dynamics and limits induced by some constraints can make it impossible to fully track the target profile.

2.5 Modeling disturbances

2.5.1 Domestic Hot Water consumption

As buildings have better and better thermal envelopes, more energy is consumed in new residential buildings for DHW than for SH. Achieving a more robust characterisation of DHW consumption profiles can benefit algorithms providing DR services. A detailed review of DHW

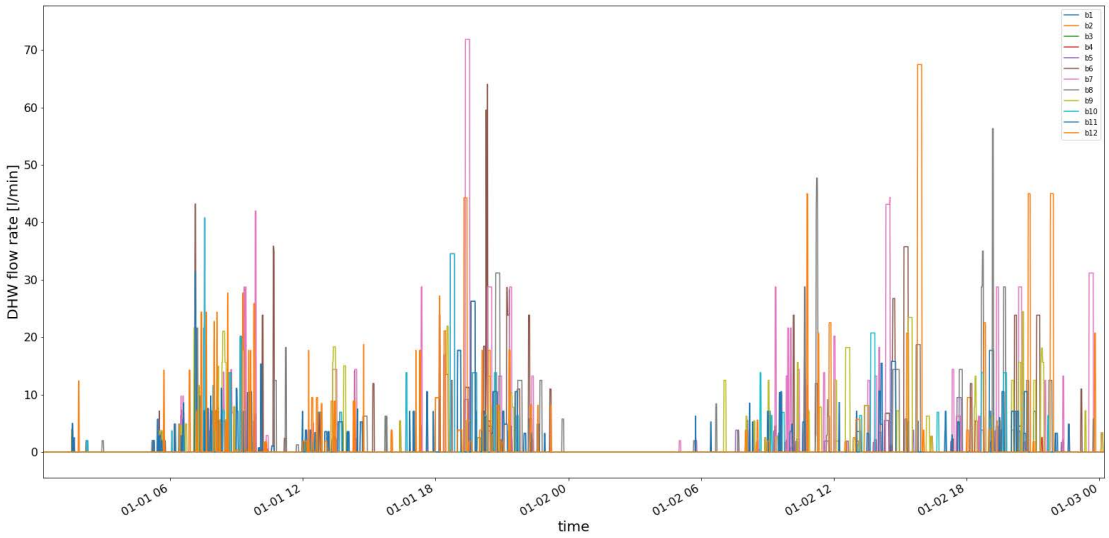


Figure 2.3 – Two days of DHW load profiles for 12 buildings (randomly generated based on guidelines).

profile modelling and its application in the building field can be found in Fuentes et al. (2018). Contrary to most studies, we do not have access to the DHW draw-off volume. de Santiago et al. (2017) developed a DHW profile generator based on a field study. They highlight the high variability of draw-off consumption per day in Switzerland (20 to 40 l per person) but show that Day of the Week (DoW) or seasons have only a small influence on DHW profiles. We generated pseudo random synthetic profiles in l/min for ten weeks based on an assumption of the number of residents per building, an average draw-off of 40 l/(d.pers) and typical shape patterns found in the literature (e.g. morning peak and higher consumption in the evening). An example of two days 1-minute sampled DHW profiles for 12 buildings of different size is given in Figure 2.3.

2.5.2 Modelling dynamic internal heat gain

Internal heat gains (IHG) in building simulations often include the effect of three main factors: (1) the heat released by residents (e.g. sensible and latent), (2) the heat induced by lighting and (3) the heat resulting from the use of appliances (Elsland et al., 2014; Jenkins et al., 2008). IHG can be more or less detailed, depending on the purpose of the simulation, but in most cases an average value is used. Here we model the IHG at each time step t based on equation 2.31:

$$IHG_t = Q_t^{people} + Q_t^{appliance}, \tag{2.31}$$

where IHG is the total internal heat gain, Q^{people} is the heat released by residents and $Q^{appliance}$ is the heat gain induced by appliances and lighting. The amount of heat produced by residents depends on the occupancy profile and the type of activity they perform as

Table 2.2 – Internal Heat Gains parameters as a function of the period of the day

		Night	Day	Evening
Time period		10pm-{6,7,8}am	{6,7,8}am-{4,5,6}pm	{4,5,6}pm-10pm
MHP	[W/pers]	80	120	100
Occupancy	[%]	80-90	20-40	50-80
Appliances W	[W/m^2]	2.1	2.1	2.1
Usage Factor Fu	[%]	20-50	30-60	50-70

presented in equation 2.32:

$$Q_t^{people} = N_p * MHP_t * Occ_t \tag{2.32}$$

where N_p is the number of residents per building, MHP_t is the metabolic heat production per person [W/pers] based on guidelines describing their activity, and Occ_t is the occupancy profile in [%].

The internal heat gains from appliances and lighting are related to the Energy Reference Area (ERA) of a building and are calculated by:

$$Q_t^{appliance} = ERA * W_t * Fu_t, \tag{2.33}$$

where W is the average power density of appliances per square meter [W/m^2] and Fu is the usage factor in [%]. If all appliances are used, Fu is equal to 1. The average power density of appliances is assumed to be constant and equal to 2.1 W/m^2. This assumption is not valid for small ERA, for which this average power density follows an exponential function. The buildings we test in this project are multi-family houses with an ERA higher than 400 m^2.

IHG profiles are generated based on three period profiles: Night, Day and Evening. The IHG parameters for these periods are summarised in Table 2.2. The transition time between Night and Day, and Day and Evening are randomly generated. For each period, our algorithm randomly generates quarterly based resident occupancy profiles and usage factors for appliances within the boundary values given in Table 2.2. The profiles generated this way are then smoothed using a rolling mean average over the last hour.

2.6 Results

2.6.1 Economic MPC in a simulated environment

Figure 2.4 illustrates the iterative steps of an economic MPC and its effect on the aggregated power load of a simulated cluster of buildings. The simulation is carried out with MATLAB ® SIMULINK ®. At the beginning the test, the buildings are controlled with a Rule Based Control included in the simulation framework. When the MPC control is activated, we use an interface

to couple the detailed simulation and the optimisation-based controls (developed in python) dynamically. The MPC algorithm is implemented in python using the gurobipy library (Gurobi Optimization, 2018) to generate and solve the optimisation problem. Simulation values, DR parameters and control setpoints are exchanged between SIMULINK ® and python every 15 minutes in simulation time. Since the problem is formulated as a MILP, the optimisation is stopped when a solution falls below a certain value of optimallity (e.g. 5% of the MIPgap) or after 40 seconds in order to limit the computational time. If the problem is infeasible, the default control setpoints are sent to the simulation controller to avoid failure of the run.

The top panel in Figure 2.4 corresponds to the first iteration of the model predictive control. Note that the MPC formulation is not exactly the one presented in the previous section, as the formulation has evolved during the project. The yellow line corresponds to the targeted power of the cluster for the next control horizon (e.g. 15 minutes), while the grey is the sequences of optimised power for the prediction horizon. The middle panel presents the result after the first iteration, where a new sequence of power is computed and the first control actions are applied to the simulation. The bottom panel presents the results after 19 iterations. We can observe that the MPC drives the system to a lower aggregated power level. The power follows the target fixed by the MPC (similar trend) but is on average always below it. This can be explained by the limitation of local control rules not taken into account in the MPC formulation (e.g minimum down time of the heat pump). In general, heat pumps take more time to start than to stop. To benefit from a better COP, thus reducing the cost, the solver will select a solution that shuts down the running heat pumps (at high temperature) and starts up the heat pumps that were previously OFF (at lower temperature).

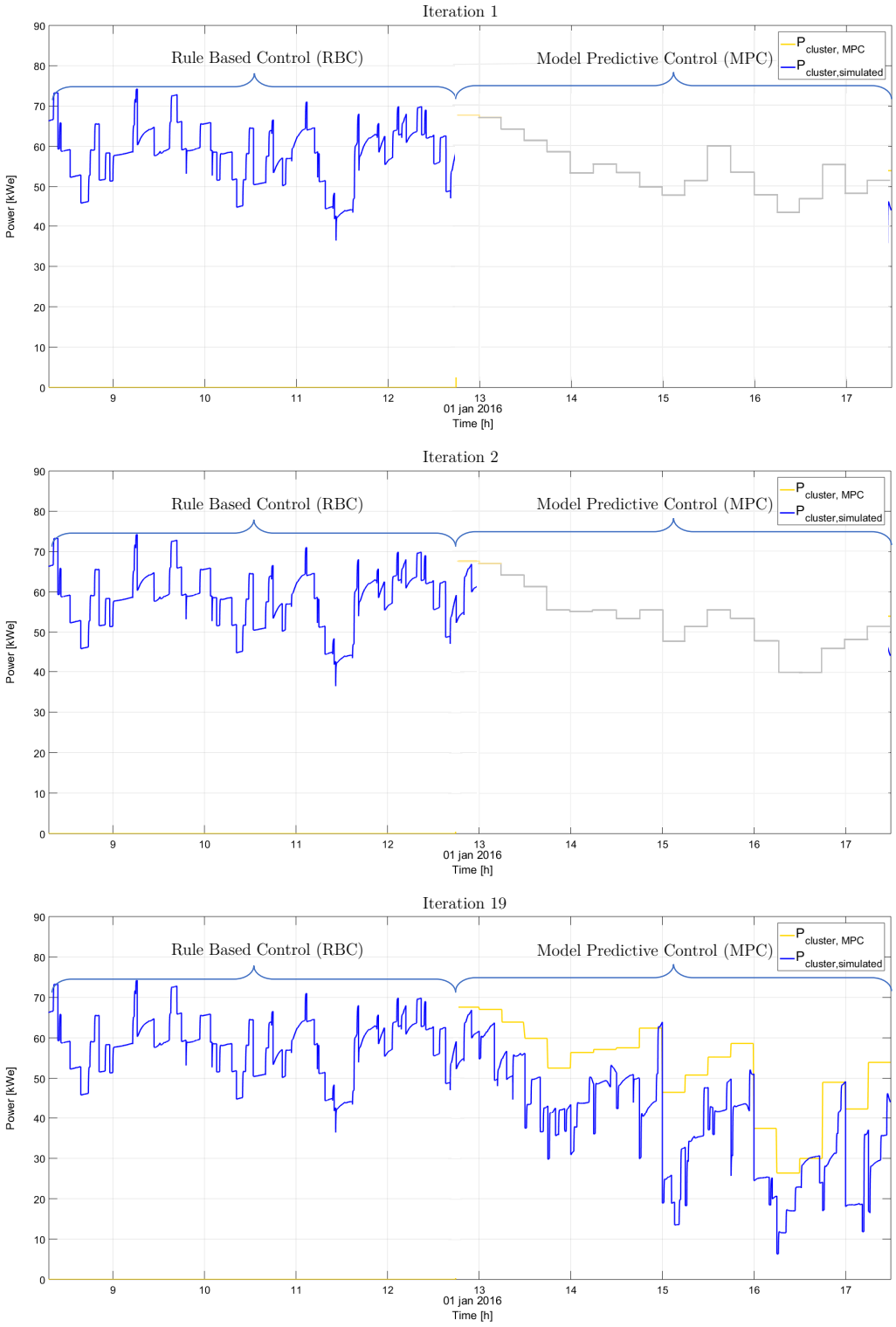


Figure 2.4 – Iterations of economic model predictive control for a cluster of 12 residential buildings in a simulation environment. In blue the simulated power, in grey the optimal power computed over the prediction horizon, in yellow the power as a solution of the optimisation.

2.6.2 Use case: Characterisation of the flexibility of residential buildings

This subsection summarises the work presented at the IBPSA bs2019 conference (Amblard et al., 2020). We characterise the flexibility offered by heat pumps in residential buildings for DR services at sub-hourly time steps. We compare our estimates of flexibility obtained with two control methods: the MPC formulation presented in this chapter and a Rule Based Control (RBC). To reduce the computational burden of extensive simulations, the potential is assessed with a simulated cluster of five buildings during typical days. We account for the effect of local control constraints and the fact that the system can fail to comply with external penalty signals. Our framework allows us to test several DR configurations, especially in relation to maximum flexibility potential fluctuations over a day or between seasons. Here, we define the flexibility as "any feasible change in power (increase or decrease) for a particular system, over a period of time, based on signals from the market/grid". Here, power refers only to electricity and not thermal power.

Indicators

The first indicator accounts for the maximum power deviation of a cluster. The amount of aggregated power consumption load P that can be increased (UP) to a maximum or reduced (DN) to a minimum operation level is computed with the following two approaches: the first computes the maximum power difference relative to the current power level at t_0 as in equations (2.34) and (2.35), while the second, equation (2.36), is based on the difference with the baseline power P^b. The terms UP and DN have to be differentiated from the existing upward and downward reserve services provided by conventional generators. In this work, "Upward" (UP) refers to an increase in power consumption, which, looking at the market side, is equivalent to a decrease in power production (downward reserve). Given the index of the heat pump n with $1 \leq n \leq N$, the call duration (d_{DR}) and the time index t with $t_0 \leq t \leq (t_0 + d_{DR})$, we can define the indicators for each *Flexhour* index t_0 $0 \leq t_0 \leq 23$, corresponding to the hour of the call.

$$F1_{t_0}^{UP} = \sum_{n=1}^{N} \frac{\max(P_{n,t} - P_{n,t_0})}{P_n^{55°C}} \tag{2.34}$$

$$F1_{t_0}^{DN} = \sum_{n=1}^{N} \frac{\max(P_{n,t_0} - P_{n,t})}{P_n^{55°C}} \tag{2.35}$$

where P_{n,t_0} is the current power level and $P_n^{55°C}$ the normalisation power. It is unclear what is the maximum power of a heat pump, thus we choose the power level of the heat pump when it is operating at a high departure temperature of the condenser at 55°C.

$$F2_{t_0}^{UP} = \sum_{n=1}^{N} \frac{\max(P_{n,t} - P_{n,t}^b)}{P_n^{55°C}} \tag{2.36}$$

Both maximum deviation indicators F are normalised by the power consumed in DHW mode with an operating temperature of 55°C.

The second set of indicators (the FlexEnergy) describes the energy deviation in kWh for a cluster of buildings and a specific duration call d_{DR}. Similar to the first set of indicators, these indicators are computed with two methods. The first one, equation (2.37), assumes that the current energy consumption at t_0 is maintained over the entire call duration. The second, equation (2.38), is based on the difference with the baseline energy consumption. The baseline consumption for a residential building is the expected consumption pattern in the absence of DR programs. Here, P values are discrete with a sampling time ts in seconds.

$$E1_{t_0}^{UP} = \frac{ts}{3600} \sum_{n=1}^{N} \sum_{t=t_0}^{t_0+d_{DR}} (P_{n,t} - P_{n,t_0}) \tag{2.37}$$

$$E2_{t_0}^{UP} = \frac{ts}{3600} \sum_{n=1}^{N} \sum_{t=t_0}^{t_0+d_{DR}} (P_{n,t} - P_{n,t}^{b}) \tag{2.38}$$

The third indicator tries to assess the rebound effect. It is computed as the energy deviation after the end of a DR call with the baseline as reference. For computational reasons, the rebound is computed for a period $d_{rebound}$ of four hours following the DR call.

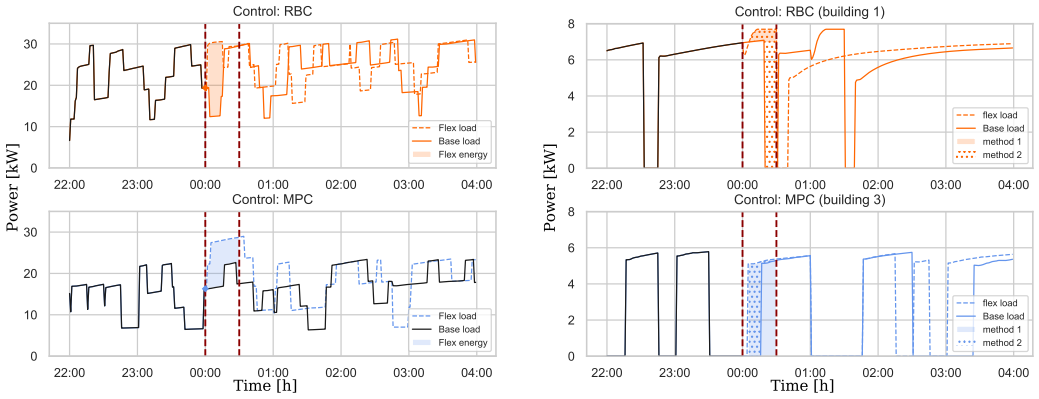
$$R_{t_0} = \frac{ts}{3600} \sum_{n=1}^{N} \sum_{t=t_0+d_{DR}}^{t_0+d_{rebound}} (P_{n,t} - P_{n,t}^{b}) \tag{2.39}$$

The time taken to reach the target power and the time spent at the target power are also important metrics of the quality of the flexibility supplied by the system. The changes in consumption patterns also have to respect the technical constraints of the systems and not compromise the comfort of users. This implies the need for a control methodology that can predict power curves and adapt over time, and makes MPC a good candidate for this work.

Typical days

Calculating flexibility for a building with various devices for an entire year using an optimisation framework is a computationally challenging task. To avoid the "curse of dimensionality", other researchers have made use of typical days that represent the entire year. Studies such as those performed by Domínguez-Muñoz et al. (2011) and Fazlollahi et al. (2014) highlight the interest of using typical days for system design optimisations. Here we use typical days to represent the year's weather, including solar irradiance and external temperature. Days are clustered into groups of three in order to be used for initialization (day 1), simulation (day 2) and prediction (day 3). When characterising the flexibility potential of residential buildings, one has to consider both SH and DHW at the same time, as most systems use a single heat pump to provide both services. When assessing the flexibility potential, one has to take into account that DHW consumption can have significant effects both on the heat pump and on the SH, as in certain cases not enough heat is supplied to the building. In order to take this issue into account, as well as the impact of the initial state of the storage tanks, the typical day simulations are run with four different samples of DHW consumption profiles per building.

Each combination of typical day (*Extreme winter*, *Winter*, *Mid-season*, *Summer*) and DHW



(a) Comparison of control method

(b) Comparison of flexibility metrics

Figure 2.5 – Example of the reaction of a cluster of buildings to a DR upward flexibility call in function of the control method.

profiles (four different ones) are simulated with one minute time steps. There are 24 runs for each combination, corresponding to each hour of a day. The flexibility is evaluated at different hours of the day (*Flexhour*) and for a duration d_{DR} of 30 minutes. Each run has an initialisation period of a day, and stops four hours after receiving a flexibility call at the beginning of the corresponding *Flexhour*.

For the sake of simplicity, we consider only one flexibility call per day, and only the maximum upward flexibility is requested per call. The time of a flexibility request is assumed to be previously unknown to the system and to the MPC controller.

Results

Figure 2.5a displays a DR call for providing upward flexibility for a typical mid-seasonal day sent to a cluster of buildings. In the top panel (orange), each building operates based on its own RBC controller. In the lower panel (blue), the cluster is operated with a centralised MPC that tries to minimise the energy consumption. It can be seen that the building with the MPC controller for cost minimisation presents a baseline that is clearly lower than the one with a RBC, which makes it difficult to compare the flexibility they offer. The DR call starts at midnight and lasts for 30 minutes (period between the two dashed vertical lines). The baseline and the power deviation of the two evaluated control strategies (RBC and MPC) are compared not only during the call but also during the following four hours.

Figure 2.5b displays the same upward flexibility call, zoomed in to the building level. It illustrates two different cases of successful activation. In the top subplot, the heat pump is already activated and providing space heating when the DR call starts. After an initial decrease during the first six minutes, the power consumption increases by 10% compared to the value at the beginning of the call. This behaviour corresponds to a switch from the SH mode to the DHW mode of the heat pump. When the call is received, the bottom temperature of the

Figure 2.6 – Variation of FlexEnergy potential with MPC, computed with method 2, for call duration $d_{DR} = 30$ min, over four typical days, with deviations.

storage tank is lower than the current heat pump operation temperature, which explains the decrease in electrical power. One can observe that the DHW cycle was started one hour earlier. Computing flexibility with reference to the initial power (method 1) results in only a small energy increase. Looking at the baseline, the heat pump would have normally stopped during this period but was kept ON by changing its setpoints. Therefore, computing flexibility with reference to the baseline (method 2) gives a much bigger energy increase. The bottom subplot with MPC illustrates the opposite behaviour. At the start of the call the heat pump is off, but it switches on after 2 minutes. However, looking at the baseline, one can observe that the heat pump would have started on its own anyway. In this case, computing the flexible energy with method 2 results in half the amount obtained with method 1. Surprisingly, however, we did not observe a significant difference overall between the two methods to assess the power deviation. This might be due to the limited number of simulated experiments. Figure 2.6 displays the aggregated flexible electrical energy in kWh of the five residential buildings considered for each *FlexHour*. Note that the lines between two *FlexHour* have no real meaning and that in this study, only one DR call per day is allowed. The duration of the call considered here is $d_{DR} = 30$ min. For this duration, the theoretical maximum flexible energy of the system considered is less than 20 kWh, assuming that all heat pumps turn on at the beginning of the

call to provide DHW throughout the entire duration. The *FlexEnergy* is low for winter days but the deviations are wider. The lines represent the mean value and the colored areas, or bands, represent the results obtained for different DHW profiles (i.e. the minimum and maximum value at each *FlexHour*. In winter, DHW cycles have a significant impact on the overall system as they often interrupt heating cycles. In summer, upward flexibility results only from an increase of the DHW temperature setpoint $T^{DHW,set}$. The thermal energy that can be stored in a DHW storage tank is an order of magnitude lower than the one that can be stored in the building mass. Once the storage tank is full, the heat pump cannot provide any flexibility for several hours. DHW charging cycles typically range from 15 minutes to a maximum of one hour if the storage is empty. This is why the value of the *FlexEnergy* would not increase much even for a longer duration call (d_{DR}. The same remark can be made for winter periods, where off times are usually very short. Depending on the method used to compute the flexibility, results could be very different. For *Mid-season* days, however, increasing the duration call up to one hour per day would result in higher *FlexEnergy* upward potential because heat pumps are more often turned OFF during those days and more thermal energy can be stored in the mass of the building.

In this chapter we used a simulation-based setting to test our models and MPC control algorithm and characterise the flexibility potential of a cluster of residential buildings. This is an important first step. However, it is necessary to validate the control algorithms on a real use case, as it is practically impossible to account for all relevant aspects of the real system with simulations. This is what we do in Chapters 3-6.

3 Demand Response implementation within existing buildings

In the previous chapter we developed a Model Predictive Control (MPC) algorithm to provide Demand Response services in residential buildings. We validated our model on data simulated to resemble a cluster of residential buildings. However, it is practically impossible to address all possible behaviours of a real system solely with simulations. Therefore, we want to implement and test our MPC algorithm on a real system.

How can Demand Response be implemented within existing residential buildings?

Promising attempts can be found in the literature. Vrettos et al. (2013) developed a state-of-the-art benchmark of a residential building including all components prone to offer DR services, namely a Heat Pump for Space Heating (SH), slab cooling, an electric resistance for Domestic Hot Water (DHW), Photovoltaics (PV) and an electrical battery. With this theoretical set-up, they reveal a significant potential for DR using an MPC controller with day-ahead and real-time prices. However, they consider a highly equipped and fully flexible system which does not exist in most building and they discard the local control aspect (translating control actions to devices). Others try to approach the problem more practically by enabling tests on real systems. Cui et al. (2017) present "OpenADR", their automation technology and communication protocol for the development and deployment of DR services. But this tool is still in development.

This chapter presents the specifications and methodology we used to equip a pilot site for the remote control of its thermo-electrical devices. The chapter is structured as follows:

Section 3.1 presents the pilot site used as experimental test facility, its actual control system and the new control framework tested. Section 3.2 details the hardware and software installed for communication. Section 3.3 presents the methods and safety protocols to ensure data integrity, and MPC implementation. Finally, Section 3.4 offers an experimental validation of the framework by presenting successful remote heat pump activations.

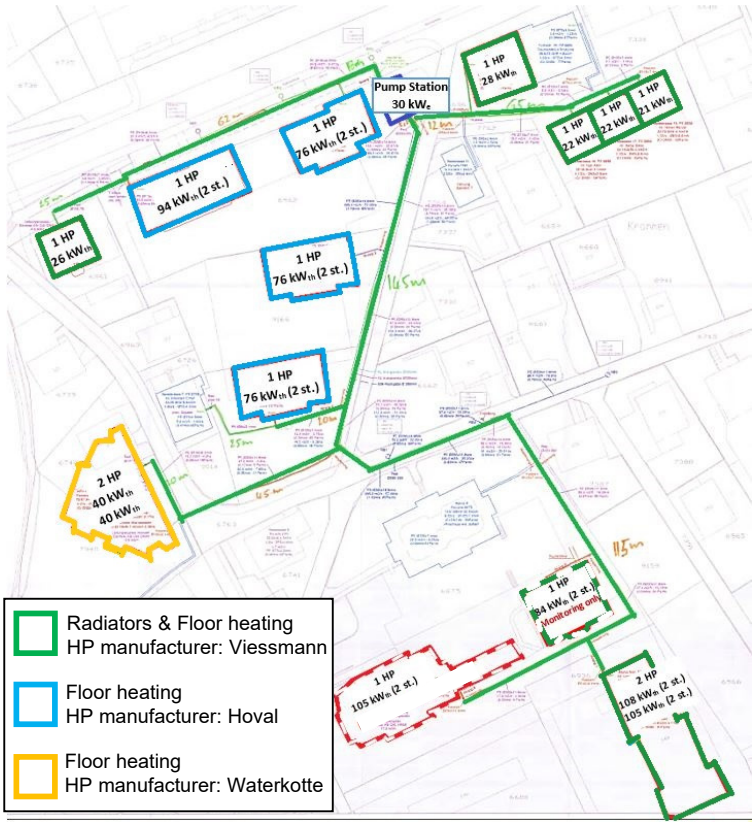


Figure 3.1 – The pilot site and Heat pump manufacturer

3.1 The Pilot site description

3.1.1 Energy systems and infrastructures

The pilot site in Naters is composed of 12 residential multi-family buildings connected to a centralized heating network. It represents 166 residential units and around 400 inhabitants. The size of the buildings ranges from 4 to 36 residential units per building with an estimated average of 2.5 inhabitant per residential unit. The buildings' construction years range from 1919 to 2015, thus their envelopes have different thermal efficiencies that can be clustered into four groups. Aggregated nominal heating/DHW demands are 838 kW for the pilot site, which corresponds to a yearly heat demand of 2 GWh. Table 3.1 summarizes the production/consumption power as well as the storage capacity of the different buildings.

The thermal demand of the buildings within the pilot site are satisfied by 15 decentralized heat pumps. The cold source for the heat pumps is a Low Temperature (LT) heating network composed of 2 low depth (20 m) geothermal extractions probes, 3 re-injection geothermal probes, powered by two circulation pumps (redundant) of 30 kW. The geothermal probes provide energy to the network through a heat exchanger of 955 kW (at full load). This system can also be used for free-cooling in some buildings in summer. The extraction and injection

Table 3.1 – Buildings' consumption and storage capacity of the pilot site

Const. year	Heat System	Q_{heat} [kW]	El. consumption [kW]			Storage [m^3]		Surface [m^2]
			$HP_{35°}$	$HP_{55°}$	El. Res.	SH	DHW	
2015	Floor	40	0	12.1	10	0	2	4'200
		40	7.8	0	2 x 10	2	0	
2011	Floor	94*	17.4	25.2	2 x 9	1.5	3†	3'040
2011	Floor	78*	14.2	21.1	2 x 9	1.5	2.5‡	2'280
2011	Floor	78*	14.2	21.1	2 x 9	1.5	2.5†	1'840
2011	Floor	78*	14.2	21.1	2 x 9	1.5	2.5†	1'840
1980	Floor	108*	18.7	26.8	2 x 15	1.75	3†	4'000
		105*	18.6	24.3			0	
1990	Floor	100*	13	21	2 x 10	1	2†	3'300
1960	Radiator	28	4.3	6.8	7	0.48	0.8	1'000
2000‡	Floor & Radiator	26	4.3	6.8	5	0.48	0.6	520
1960	Radiator	21	3.1	5.7	4	0.48	0.4	400
1960	Radiator	21	3.1	5.7	4	0.48	0.4	400
1960	Radiator	21	3.1	5.7	4	0.48	0.4	400
		838	137.6	206.4	196	13.15	20.1	23'220

* Two stage compressor heat pumps
† Two buffers in series
‡ Renovation year

temperatures (run-out) of the aquifer is around 11-7 °C in winter and 13-9 °C in summer respectively, with very low intra-seasonal variations. The temperatures within the LT network are the same but shifted by a ΔT of about 1 °C. The schematic of the pilot site energy system is presented in Figure 3.2.

Each building is equipped with one or two heat pumps, a buffer tank for SH and one or two storage tanks (in series) for DHW. Each heat pump is equipped with a single or two-stage compressor. For anti-microbial reasons, electric heaters are also used to bring the DHW storage temperature up to 60 °C at least once per week.

Three-port diverter valves allow for the separation of DHW and SH. The heat pump controller prioritises DHW production over SH. The heat is supplied to the flats by floor heating systems for the most modern buildings (by construction/renovation year) and by thermal radiators for the oldest ones. For the most modern buildings, heat pumps were a priori selected during the design phase and a custom distribution system was designed accordingly. However, for the older buildings, the heat pumps were replacements for pre-existing boilers. Thus, the distribution systems remained unchanged and were ill-suited for heat pumps.

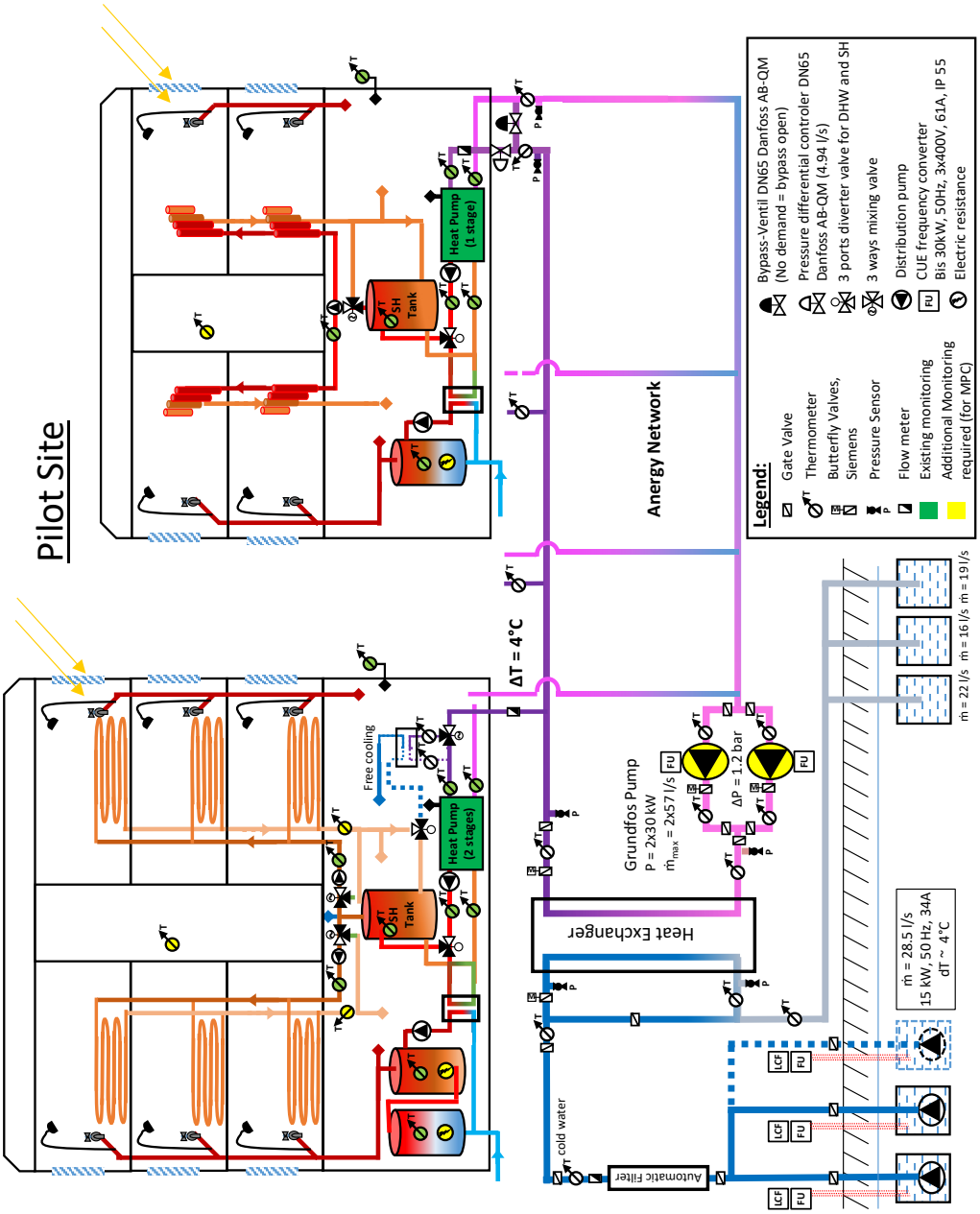


Figure 3.2 – Schematic representation of the pilot site main infrastructures and devices

3.1.2 Current control system

A management system specific to each heat pump manufacturer is installed along with the heat pump. This management system also affects the distribution system by changing the setpoints of the motorised valves. The control logic uses conventional dead-band methodology. When the temperatures of the DHW storage and the buffer tank drop below their respective lower limits, the heat pump switches on until the upper limit is reached. When switched off, different controls on the internal cycle (pressure, temperature, time) block the reactivation of the heat pump for a period of time (depending on heat pump model) to avoid operating failure. The mass flow, at the evaporator side of the heat pump, is controlled by a circulation valve to maintain, when switched on, a ΔT of 4 °C in the LT network.

For SH, a three-way mixing valve between the departure and the return of the heating distribution system guarantees the respect of the setpoint supply temperature at the departure of the distribution system. The setpoint supply temperature is computed following the conventional control schema of the heating curve (Figure 3.3). A heat curve returns the supply temperature in function of the outdoor temperature, with a signature slope and a temperature threshold specific to the building thermal characteristics. Floor heatings typically have a slope smaller than 0.6 while radiators have of slope higher than 1. Different modes can also be set up: day/night, winter/summer, a constant ΔT in the buffer tank or a stratification charge.

DHW production by the heat pumps takes priority over SH. The heat pump control unit switches OFF the DHW circulation pumps during storage charging. A three-port diverter valve separates the DHW and the water for SH. SH takes place if the operating program is switched ON and an operating status is active in the time program. The heating circuit circulation pumps are constantly switched on in SH mode. When multiple heating loops are connected to the same heat pump, the local controller takes action to always ensure the minimal requirements for the highest departure temperature.

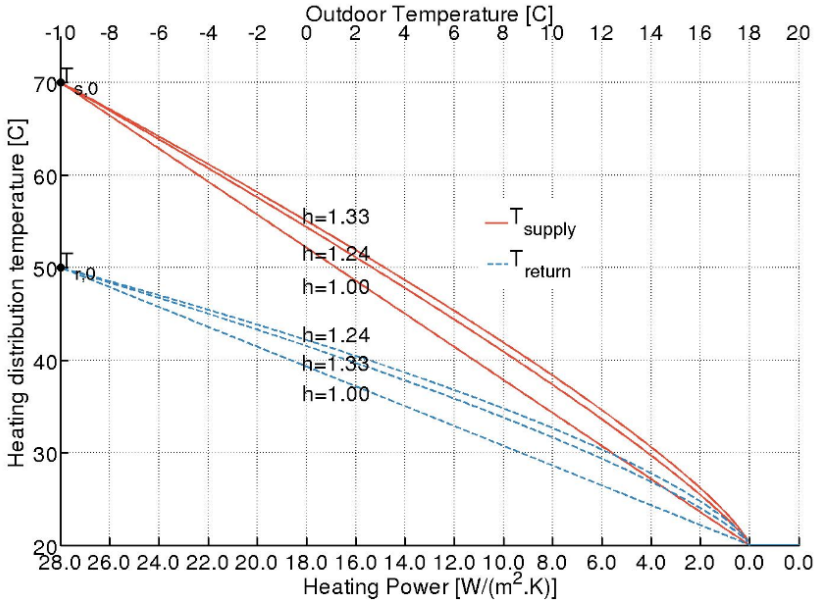


Figure 3.3 – Example of heating curves for a building with Energy Signature slope (k = $1W/(m^2°C)$) and heating threshold temperature ($T_{tr}^{hs} = 18°C$) in (Girardin, 2012)

3.1.3 New control framework

This subsection reports our control structure for DR services in residential buildings. In the scope of this thesis, the implementation of DR services is solely based on Power-to-Heat (P2H) because neither on-site electricity production nor storage is available at the pilot site. The developed strategy consists of using the thermal mass of the buildings and buffer tanks for DHW and SH as heat storage, allowing the shift of both thermal and electric power.

The buildings' owners, mostly inhabitants, also own their heating system (heat pumps, storage tank, etc.). As DR services include a modification of the control over the heat pumps, the consent of the buildings' owners is required. There are two main ways in the market to modify electricity consumption patterns of thermo-electric systems: direct load control (indirect incentives) and time variable electricity tariffs (direct incentives). Fell et al. (2015) found through an online survey experiment, that consumers prefer direct load control over time variable tariffs. Similar conclusions have been drawn from a survey within the Sim4blocks project: Inhabitants wanted to have "*control over their heating system, without doing anything*". Therefore, the inhabitants' involvement for the pilot site is planned to be as minimal as possible. Feedback from them in terms of comfort will be helpful for improving the models for the MPC controller but no automation regarding this aspect is planned. The idea is to have a minimal impact on them, in terms of comfort and implication, so that they do not see the difference to the previous system. Retribution methods to provide them with incentives are not planned. Inhabitants will not have direct control over their heating devices, except manually through their radiator's valves.

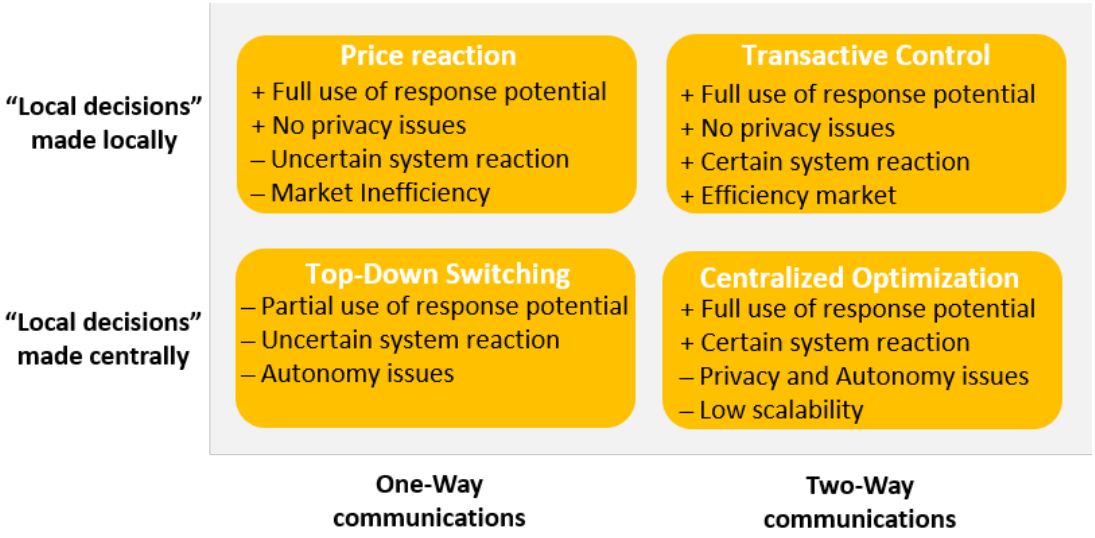


Figure 3.4 – The energy management matrix: The four categories with pros and cons from Figure 1 in (Kok & Widergren, 2016)

Kok and Widergren (2016) introduce the "smart energy management matrix" that classifies smart grid management approaches into four main categories as presented in Figure 3.4. This representation can help debating the advantages and drawbacks of different management strategies. In terms of control optimisation strategies, three main types have been proposed: centralised, decentralised and distributed control. An excellent review of the control architecture for distributed and decentralized control is found in (Scattolini, 2009). Comparison between a centralised and distributed energy management MPC is presented in (Scherer et al., 2014). With their distributed formulation, they achieved a performance comparable to a centralised MPC. They claim it is also easier to scale to larger number of HVAC units.

For this project, we chose a centralised approach because it enables full process knowledge that allows to correct for system deviations in a dependable fashion, either quickly or very consistently. It limits the need of onsite intelligence that can be costly. The relatively small size of the pilot site also limits the computation requirements that can result from centralised optimisation. Privacy issues are also not a problem for multi-family building. The system architecture of the centralised MPC deployed on the site is presented in Figure 3.5.

The heat pumps and electric heaters will provide most of the flexibility for DR services. Traditionally, the available remote-control options for heat pumps was to switch the devices OFF or ON using a centralised command. The ON state does not ensure the heat pump compressor to be running, but it powers it up. Centralised command is still widely used by DSOs for peak shaving, as it is the case at the pilot site. To extend this option, the SG-Ready interface has been developed by heat pump manufacturers and has been deployed to heat pumps over recent years. It provides a standardised, low complexity remote access to the heat pump internal controller. Four different operation states of the heat pumps can be triggered

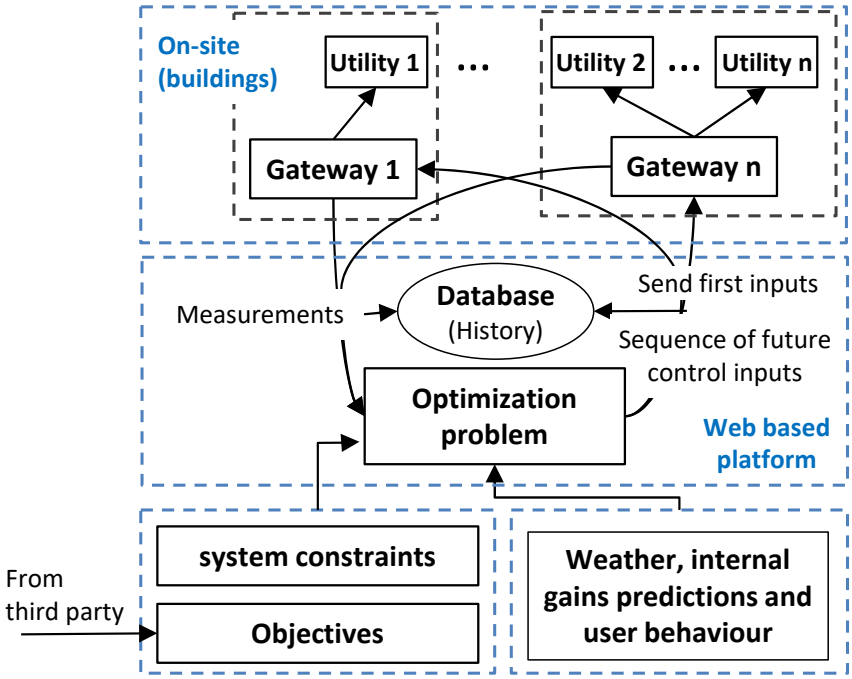


Figure 3.5 – Flow chart of the MPC framework and the system architecture

"Switch off", "Normal operation", "Recommended on", "Forced on". As is extensively addressed in (Fischer, Wolf, & Triebel, 2017), SG-Ready increases the range of options for direct load control. However, heat pumps at the pilot site do not have an SG-ready interface but can still be remotely accessed. Therefore, similar specifications as presented Table 1 in (Fischer, Wolf, & Triebel, 2017), will be implemented on the heat pump when possible.

In cybernetics and control theory, a setpoint is the desired value for an essential variable of a system. The setpoint temperature of the buffer and storage tanks as well as the measured external temperature can be modified and sent, at each time step, as input to the heat pumps' controller to change their behaviour. Without changing the internal heat pumps' controller, it is the only way to turn on the heat pumps at any desired time. Internal constraints can, however, block the start of the heat pumps for safety reasons. Turning Off the heat pumps requires only a binary signal. A few sensors (temperature, mass flow) are installed inside the buildings, but without an intrusive solution. This can be particularly challenging as highlighted in the following section. Temperature and thermal behavior of the building are estimated, at each time step, based on temperature sensor measurements of the distribution system and the storage tanks. Regardless of the DR services applied, the users' comfort will remain the main objective.

3.2 Hardware and software implementation

In order to be able to use the pilot site for testing Demand Response, several pieces of hardware and software for monitoring and communication need to be installed. These systems are essentials to integrate our theoretical control model into a real heat pump, as well as to actively assess the state of the system. One of the main challenges compared to lab field tests with hardware-in-the-loop is that the system is already up and running. Heavy modifications of its structure are nearly impossible due to space, cost or owner consent. The solution deployed has to be the least intrusive as possible in order not to lose the warranty on some piece of equipment. As mentioned in section 3.1.3, we selected a centralized approach for control, which requires two levels of communication: on site and online.

3.2.1 Instrumentation at device level

Gateways

Gateways are hardware components capable of collecting, processing and exporting data from the plants' controllers (heat pump, distribution pump) to a cloud-based platform. The gateways used in the test case don't have internal intelligence able to compute an optimal control strategy. However, thanks to small internal memory capacity, they are able to collect data for about a day if connection is lost with the central controller.

One of the main challenges is to provide off-the-shelf solutions for such a hardware component that is economical and that is applicable to a large number of energy conversion units (in this case, a heat pump). The gateway transforms the setpoints into ASCII/RTU and uses Modbus to convey this to the respective asset/device. On the Gateway is running an XML script where all datapoints are declared. This XML script manages which datapoints will be read or write of any Asset. The gateway collects all of the datapoints and sends them as a package to the cloud. The communication parameters of any assets are stored in this XML file. All calculations for the unit conversions are also carried out on the gateway, so that the values are already sent in the desired unit.

Heat pump control units

The heat pumps were manufactured by three different companies, see Figure 3.1, which implies specific communication protocols and operation modes. The Hoval and Waterkotte heat pumps can all be remotely monitored using the ModBus protocol. The Hoval heat pumps, used on the pilot site, do not support setpoints overwrite using the Modbus protocol. Thus, direct remote control is not available. The Waterkotte enable remote control using the Modbus RTU but the installed module has not proven to be useful for actuations. The Viessmann heat pumps, on the other hand are controlled through the proprietary Vitogate module as the internal control inside and between the various Viessmann components is performed through the proprietary LONworks communication protocol.

Storage tank and external temperature

The temperature sensors are connected to the heat pump by PT1000 (four wires). Their location could not be changed which forced us to develop an alternative solution when facing problems resulting from their improper installation. As opposed to detailed simulations or hardware-in-the-loop where multiple states are available, most of real thermal buffers only have one sensor. Additional sensors could not be installed which is a big limitation for accurately assessing their state of charge.

Electricity meters

To ensure and certify the traceability of the DR services offered, the power consumption of the controllable devices need to be measured. This will allow the actors in the electricity market to assess the difference between what is planned and what the energy system is doing.

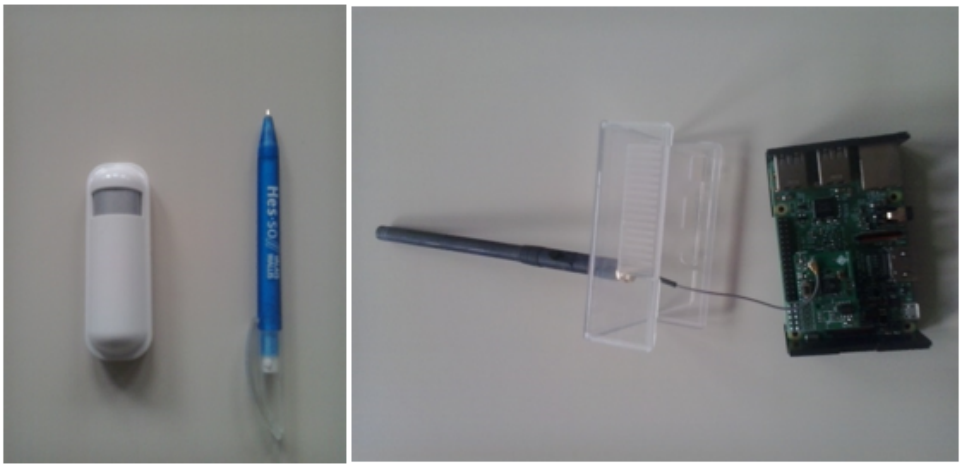
For the buildings with Hoval heat pumps, specific electricity meters dedicated to the heat pumps have been installed. The communication protocol in this case is Modbus RTU. For all the other systems the consumption/power measurements of the heat pumps are aggregated with other unknown building appliances. The communication in this case relies on the standard protocol IEC 62056-21. An Optical Probe with USB output communicates via infrared waves between the gateway and the electricity meter.

Power meters for several buildings were unexpectedly replaced during the field test by smart meters with G3-PLC communication technology. PLC is the most commonly used communication technology for smart metering. It enables faster and more reliable communications and flexible connectivity with new intelligent applications. However, these devices could not be used due to the high cost of adapting the existing communication infrastructure. Therefore, the power of four of the tested buildings are derived from a model, as presented in section 4.3.1.

Indoor sensors

Measurements of the temperatures inside the buildings (indoor, wall, floor) are important to assess the comfort of the inhabitants. Most existing buildings are not equipped due to the extra cost and also because conventional controllers are very conservative regarding comfort. Once calibrated, comfort setpoints are rarely updated. However, DR services use buildings' thermal inertia to store energy, possibly affecting the comfort of the inhabitants.

Most of theoretical work on DR within residential buildings are taking for granted the availability of measurement of temperatures. In reality, the installation of temperature sensors faces technical and consent problems. To avoid privacy issues sensors are installed in the staircases (Figure 3.7), with the consent of the inhabitants. A list of the buildings that gave us their consent is presented in Table 3.2. Sensors also need to use a wireless communication technology. For a large residential building, it can quickly become a technical challenge to equip all floors.



a) Temperature probe (z-wave) b) Raspberry Pi (with z-wave antenna)

Figure 3.6 – Indoor sensor (a) and raspberry Pi (b) installed on-site

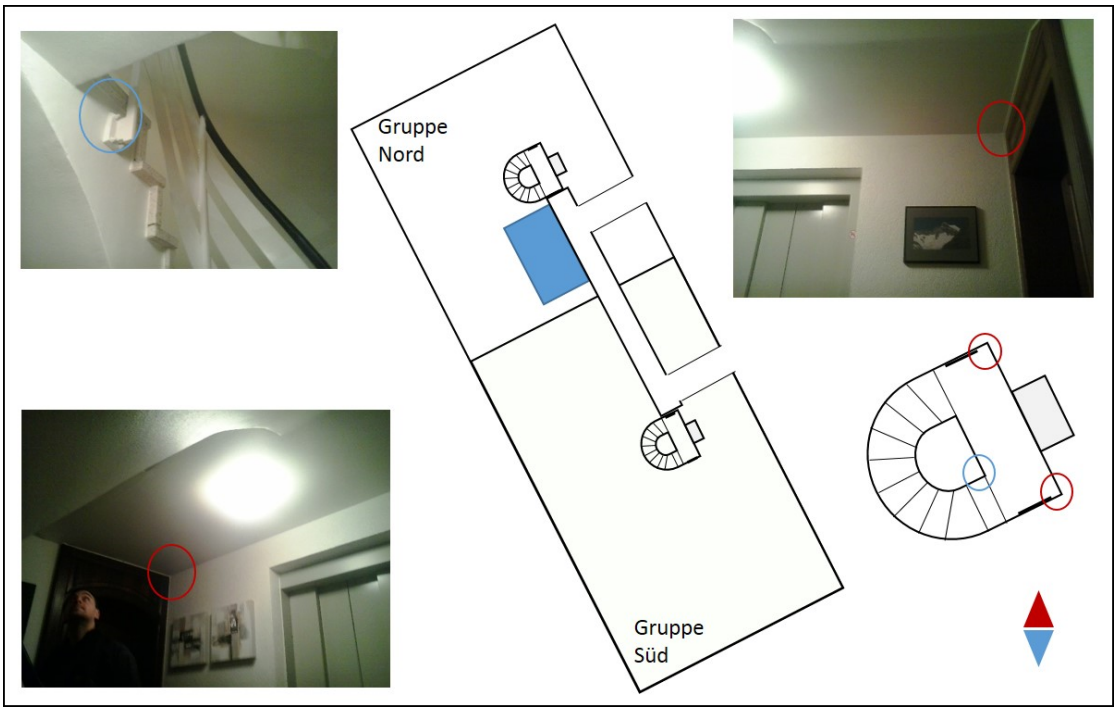


Figure 3.7 – Location for temperature sensor in one of the buildings. There are two main zones, 6 to 7 floors and two staircases (with plugs for the sensors). The blue rectangle is the room with the heat pumps and gateways. Only the north zone could be equipped

Table 3.2 – Summary of indoor sensor installation and communication quality associated. One of the sensors is always installed inside the technical room where the heat pump and buffer tanks are.

Building	Nb of sensor	Quality
Werdenberg	8	poor
Kyburg, Tarasp, Stockalper, Chillon	-	-
Gruyères	8	medium
Prangins	-	-
Oberhofen	9	good
Jegenstorf	-	-
Montebello	6	good
Grandson	6	good
Lenzburg	-	-

Wireless temperature sensors (PAT02-1B) are installed in the staircase of the buildings. The sensors use Z-wave technology (proprietary protocol) to communicate with a Raspberry Pi located in the technical room (Figure 3.6). As they run on batteries, the distance to the local room is limited. Additional sensors could be used to repeat the signal for better coverage. However, this would require the sensors to be plugged. To limit their energy consumption, they send new values only when a significant (>0.1°C) change is measured. The raspberry Pi and the router are connected via an Ethernet cable. The data is sent to CloudIO via an Advanced Message Queuing Protocol (AMQP) where data is translated to Message Queuing Telemetry Transport (MQTT) messages. CloudIO is a MQTT/AMQP broker that manages the access rights. It handles the push, pull and subscribe command to get the values from the different gateways.

Flow meters

To better understand the dynamics of the heat distribution system in the buildings and help calibrate the models, we tried to install flow meters where this was possible. In some buildings, this was impeded by a lack of proper location. When the installation was impossible, we used nominal values from the installation plan.

The distribution systems of the smallest buildings are fitted with Grundfos circulation pumps of type Alpha. They have their own internal control of the mass flow. In this case, the duty point of the circulation pump will move up or down on the highest proportional-pressure curve, depending on the heat demand in the building. The head (pressure) is reduced when heat demand declines and increased heat demand rises. For more details, please refer to their Alpha2 installation manual (Grundfos, 2013). This type of pump does not provide remote access. For this reason, we temporary installed an extra ultrasonic flow meter.

The installation of ultrasonic flow meters has proven to be more complex than expected, because it needs to be in direct contact with the pipes on a straight section of more than one

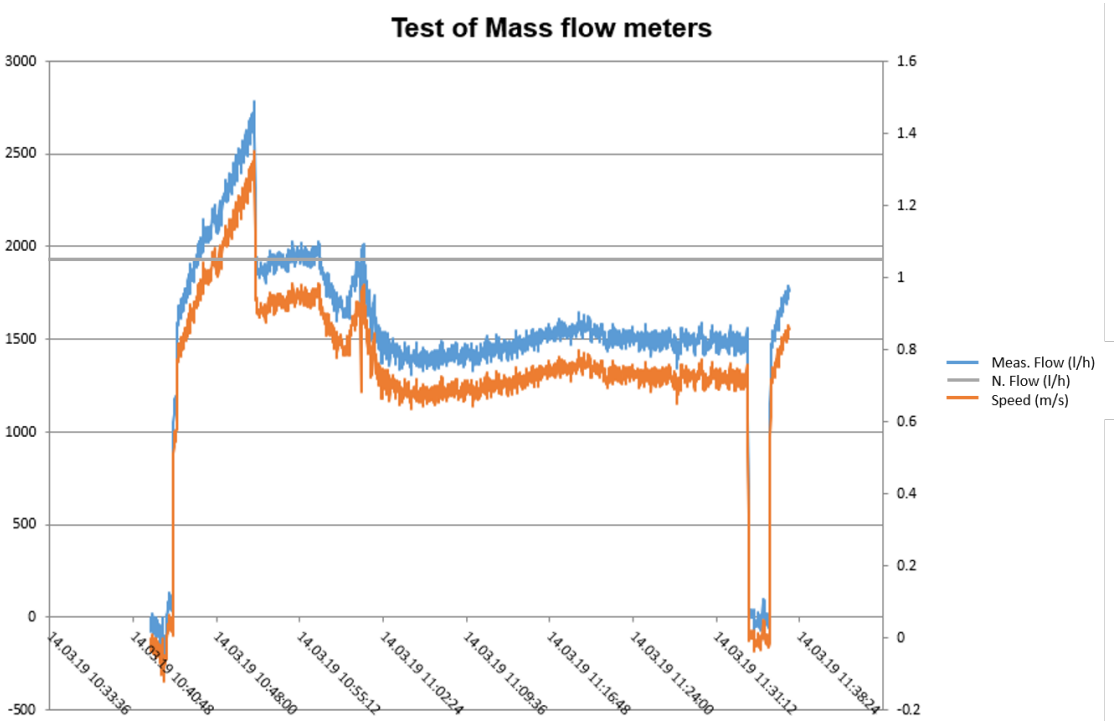


Figure 3.8 – Test for ultrasonic flow meter installation

meter. In existing buildings, this is almost impossible to comply with. We set up the flow meter after the mixing valve, as it is the only space available without insulation. The parametrisation of the flow meter is far from perfect as some parameters (thickness of pipes, materials) are not fully known. The location of the measure near a mixing valve is not ideal either because of the turbulence generated by the opening and closing of the valve. In order to calibrate the zero of the meter, the heat pump has to be shut down. When switching back on the heat pump, the mass flow would steadily increase before dropping below the design mass flow value (Figure 3.8). As the installation trials were not successful, we renounced to install the flow meters tested in lab conditions.

For larger buildings, the heat distribution systems are fitted with bigger and more complex circulation pumps, in this case the Magna series. Such devices are usually designed with remote access and IoT access in mind. Hence, data collection using Modbus RTU/Modbus TCP/IP protocol is possible for the buildings equipped with the Magna pumps. We installed a printed circuit Board (PCB) module called the CIM200 on two Magna pumps. Although remote control is also possible, we chose, for security reasons, to only monitor the system. This provides remote access to the flow rate, the head, the supply temperature, and the power consumption. Data are sent using the same IoT system as described for the indoor sensors. Contact sensors for measuring the supply and return temperature of the distribution systems are also installed.

3.2.2 Online communication protocol: the platform

The data collected locally from the gateways are transferred to the Everyware cloud, an online platform from Eurotech, over a secured MQTT SSH protocol. The cloud is connected via the datacenter of Eurotech. The platform requires an authorisation to be accessed. At each connection, a token is generated providing access to the platform for a certain duration. The structure of the platform is as follows: Each site (group of buildings) is composed of several buildings with their unique identification (Id and name). Each building can be composed by several assets with their unique identification. An asset represents a single device (e.g. heat pumps, power meters) or structure (e.g. storage tanks). Again, an asset can be further split in several datapoints, corresponding to single sensor or setpoint values. Datapoints have a name, a display name, a unit, the time of the last update and a timeseries of all recorded values.

A RESTful API in python, developed for the purposes of the project, allows to get data and post new setpoint values to the system. Different http requests allow to retrieve data from the platform. For example, to retrieve all the last values of an asset (metric = asset Id) the following http request is used: *https://eos.misurio.com/api/timeseries/metric/'+str(asset_key)+'/lastvalue*.

All monitored variables are available either as a single last value or as a time series by using the right GET command. The control variables resulting from the centralised controller are transformed and sent to the desired setpoints using the SET command. More details of the interface developed during the project and the function used are available in Appendix A.

Data are transferred at most every two minutes if a value change is recorded by the gateway. If no changes are detected for a datapoint, a value is sent only every hour. Automatically checking the time of the last update for an asset allows us to assess if a device is up-and-running. If no values were updated in the last hour, an alarm is sent by email.

3.2.3 Weather forecasting

Model predictive control used for DR service relies on the ability to accurately predict future behaviour of the system offering flexibility. For buildings this means we need to predict the thermal losses and solar gains to assess their energy requirements. Thermal losses and solar gains depend on weather conditions. Models using a weather forecast can also help assessing the comfort of the inhabitants. For real-time applications, accurate short-term predictions are mandatory, and three main forecasting methods can be adopted:

- **Online prediction**: It relies on forecasts made available by third party modeling sources. It requires an active internet connection but has the advantage of providing accurate predictions using complex models (Remund, 2008; Zhang & Hanby, 2007)
- **Offline prediction**: It does not require an internet connection as it solely uses historical data measured on-site, but it needs a model to predict future weather disturbances. The most simple model is based on a rule of thumb stating that "conditions of the next

hours, or day, would be similar to those of the previous time period (Foucquier et al., 2013). More advanced methods rely on statistical or machine learning models (Florita & Henze, 2009; Hernández Hernández et al., 2017).

- **Combined prediction**: It combines both methods mentioned above with two main advantages: (1) It can reduce the uncertainty from any discrepancy between the site location and the location of the weather station, (2) it address the risk of internet service disruption by still providing forecasts.

As there is no meteorological station accessible in the vicinity of the pilot site, we could not use historical data to model our own weather forecast. Therefore, we use the meteorological forecasts available on https://meteotest.ch/en/. Through a subscription to this website we have access to 168 hour weather forecasts by sending a URL request to their server. An App automatically sends requests, centralises and stores the history of all weather forecasts for several locations. Every minute, up to 55 times per hour, the App checks if the weather forecast of the current hour has been stored on the xMeteo server. A list of the weather forecast variables available is given in Table 3.3. The xMeteo server can then be interrogated by three HTTP requests

1. **getForeCastFor**: Return the 168 hourly forecast values computed for a specific date hour and for a site
2. **getForeCastDoneAt**: Return the 168 hourly forecast values computed from a selected date hour and for a site
3. **get_weather_data**: Return the (historical) weather data for the time interval specified

The XML format of the output is converted and reshaped into a python DataFrame that can be pre-processed and used for the simulations or as predictions of future disturbances for the MPC. Only the Air temperature `tt` and the Global radiation on the inclined plane `gk` are used to model the external predictive disturbances.

Table 3.3 – Summary of available weather forecast variables

Variable	Description
tt	Air temperature [°C]
gh	Global Radiation on the horizontal plane [W/m^2]
dh	Diffuse Radiation on the horizontal plane [W/m^2]
bh	Direct Radiation on the horizontal plane [W/m^2]
gk	Global Radiation on the inclined plane [W/m^2]
dni	Direct Normal Irradiation [W/m^2]
rr	Precipitation [mm]
ff	Wind Speed [km/h]
dd	Wind Direction [Degree]

3.3 Data integrity and compatibility with model predictive control

As presented in the previous section, the pilot site is equipped with the infrastructure to collect data on the buildings' heating systems, and to operate their controllable assets (heat pumps and valves) in order to provide DR services. The type of data provided is not consistent across buildings and heat pumps. In this section, we address the following questions:

- What are the specifications necessary for data management?
- Can existing residential buildings provide sufficient information to be controlled?

3.3.1 Data integrity

Data integrity is defined as the overall accuracy, completeness, and consistency of data over its life cycle (Boritz, 2005). Data integrity also refers to the safety of data with regard to regulatory compliance — such as GDPR compliance ("European Parliament and Council of European Union Regulation (EU) 2016/679", 2016)— and security (Giani et al., 2013). Several factors can affect the integrity of the data. A few examples include:

- **Human errors**: When someone incorrectly enters information (addresses, conversion keys) or makes mistakes during the implementation of a safeguarding procedure
- **Transfer errors**: When data cannot be transferred between the site and the platform.
- **Compromised hardware**: When a communication device or server crashes. It can limit the access to data which makes information hard to use. In worst cases, it can stop the execution of a control App or software.

Consistency of data

The consistency of data refers to the reliability or validity of a measure. The reliability check during the hardware installation is carried out manually. When a device type is connected for the first time, the delivered values are verified. In addition, there is a short verification during each commissioning.

The number of stored variables drastically changes from one heat pump manufacturer to another, from a dozen up to several hundreds of variables. A first screening of the manufacturer datasheets gives us the value ranges or the conversion keys (e.g. data points multipliers). These values, discussions with heat pump engineers and the monitoring screening of the data, can help us highlight the pertinent set of data. However, all of the selected data cannot always be useful. We conduct a semi-automatic checking of all the data to reduce the complete dataset to a more pertinent set. For example, some heat pumps have up to three possible default heat circuit configurations, each with their associated variables. The data integrity test reveals

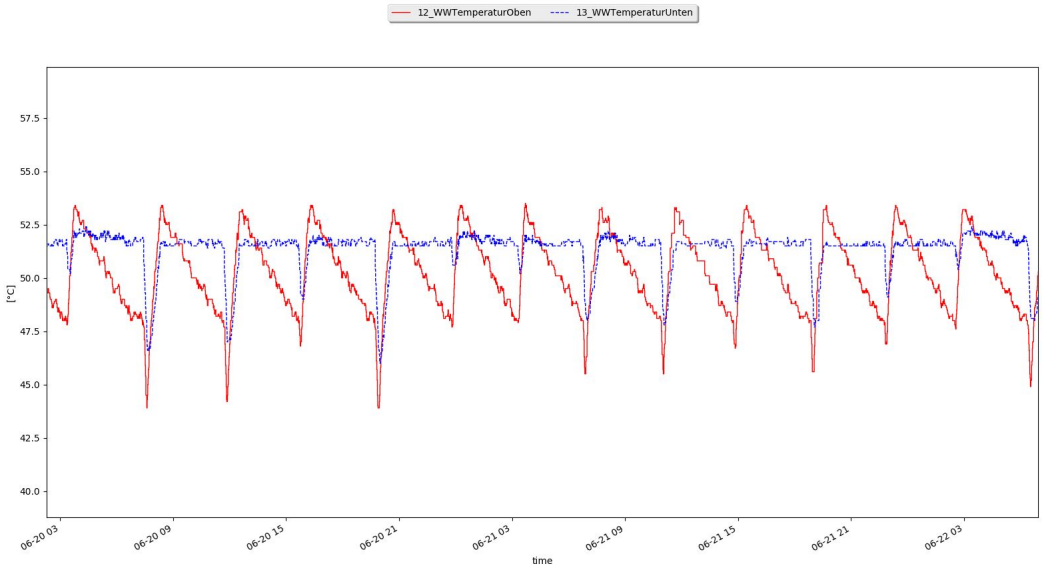


Figure 3.9 – Monitoring error detection: Inappropriate connection of a DHW buffer tank

which heating circuit configuration is activated. Moreover, measured variables with no values or constantly at minimum or maximum value are considered non-pertinent.

Several monitoring issues regarding sensors or pipe connections can only be identified by manually screening time series of sensor measurements and verifying that the values are valid. For example, in Figure 3.9, upper sensor (red line) and lower sensor (dashed blue line) temperatures of a DHW storage tank display an erroneous cycle behaviour: The top sensor monitors a drop in temperature while the lower sensor remains constant, which is the inverse of what one would expect. This could be the result of an error in the sensors' communication address or their position. However the abnormal number of charging cycles (13 in two days) led us to discover that the error is caused by a wrong connection of the inlet and outlet pipes of the DHW storage tank. Most of the sensors on the pilot site have not been re-calibrated since their installation (default manufacturing calibration).

Another example of using manual screening to detect errors is shown in Figure 3.10. Here, some external temperature values measured in winter can reach abnormally high values (up to 30°C) in the middle of the day, which clearly indicates that some sensors are directly exposed to solar radiation. This problem is addressed later in Section 4.2.

Data integrity checks can also help to reveal hidden variables (e.g. night setback schedules). Some of these variables are useful to update the state of the system or even for the formulation of the constraints. However, they are not always remotely accessible. We conduct semi-manual tests to identify night setback schedules based on the variation of the temperature inside the buffer tank for SH. We minimize the Root Mean Square Error (RMSE) between the supply temperature for SH measured and the values predicted by models using randomly defined

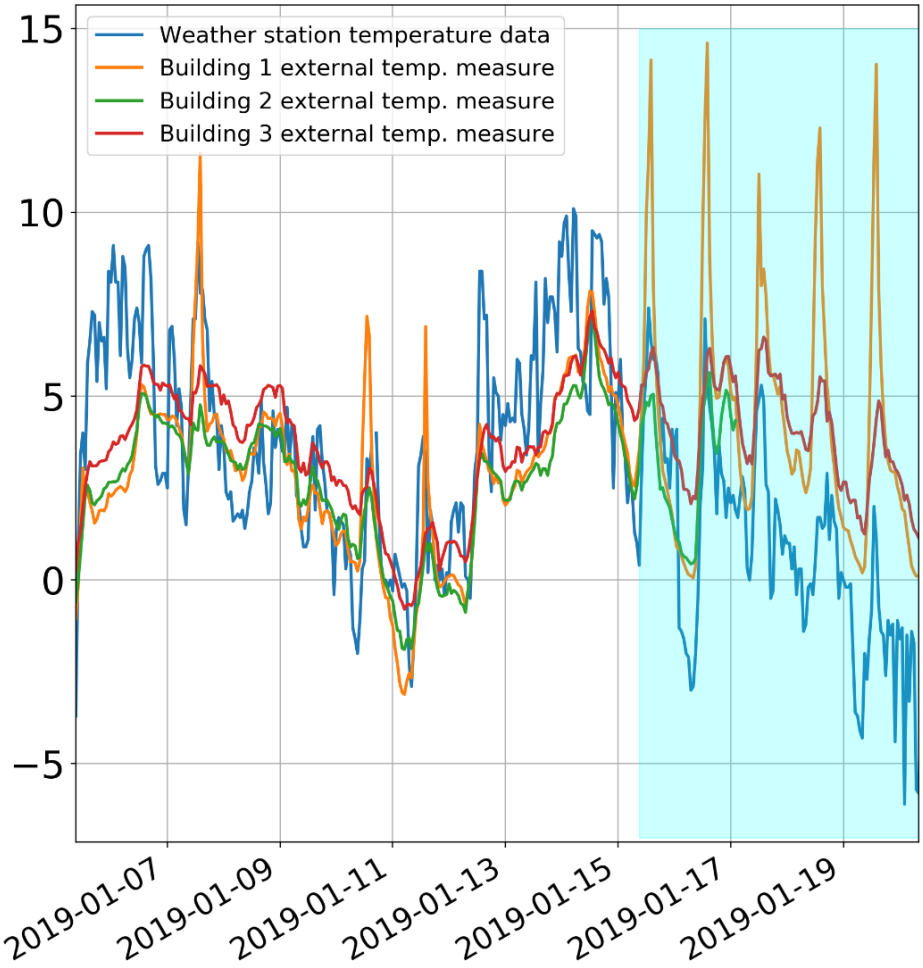


Figure 3.10 – The comparison of the external temperature readings between three buildings at the pilot site and the weather station data. In the area highlighted in cyan, one can observe the extreme spikes in external temperature registered by the heat pumps for one of the buildings, due to the exposure of the external temperature sensor to solar irradiation.

night setback schedules and different moving average window sizes. We assume the model with the smallest RMSE to be correct and retain the associated night setback schedule and window size.

Several issues with monitoring reliability are continuously noticed throughout the entire project. Monitoring systems can fail from a few hours up to several weeks, as can be seen in red in Figure 3.11. We see the temperature of the water going into and coming out of the heat pump condenser drop for long periods during a year (49 days in 2018). This is due to hardware and software problems of the monitoring framework.

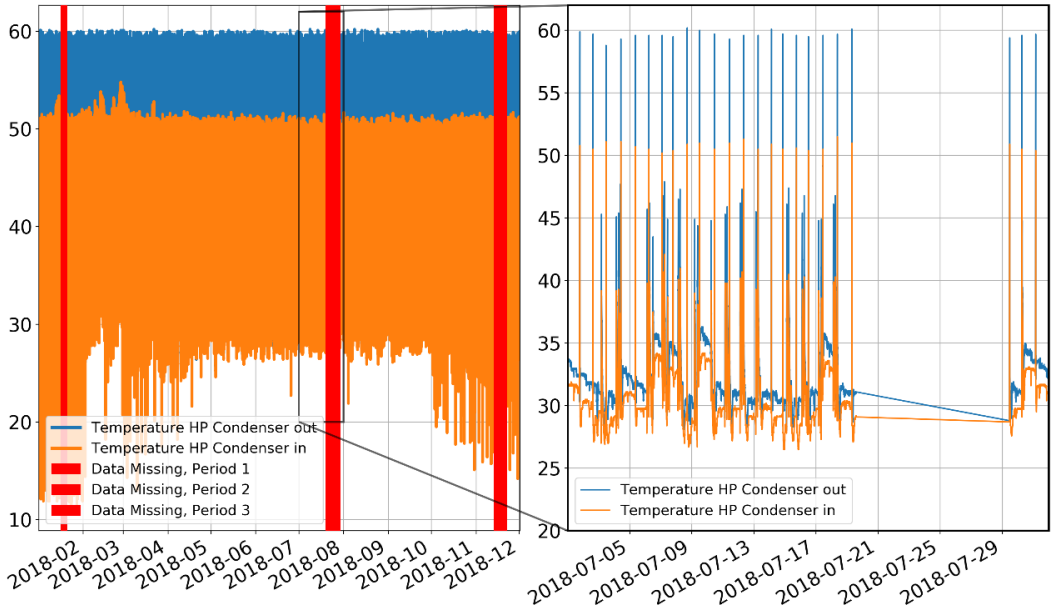


Figure 3.11 – Inlet and outlet temperatures of the heat pump condenser collected over a year for a building in the pilot site highlighting the inconsistencies in monitoring

Compatibility with theoretical Model Predictive Control

In reality, buildings and heat pumps do not provide as many variables as the ones generally used in theoretical MPC coupled with simulation. Ideally, we want to develop a model structure as generic as possible and find a minimum common ground for all buildings. Understanding the essential/critical or even available variables will later help us tune our model. Commonly used variables in building MPC, and their availability in this project, are listed in Table 3.4.

Assessing the state of charge of a building is far from simple in reality. State space models of buildings are often represented with at least two nodes to account for the two main time constants of the building: the fast dynamics of the air, used for assessing the comfort, and the slow dynamics of the internal mass, used for assessing the storage capacity. Most existing buildings are poorly equipped in terms of sensors, especially for measuring the mass temperatures of the building (e.g. T^f or T^e). Indoor temperature sensors are not always available or placed at a useful location. Comfort cannot directly be assessed if sensors are located in non-heated areas like the staircases.

Variables for the SH distribution system are also sparse. T^{flow} is often the only available. To enable T^{ret} and $\dot{m}^{sh}$ extra monitoring devices need to be installed which is not always possible (see subsection 3.2). In some cases, when heat pumps are not oversized in comparison to the building heating requirements, T^{ret} can be assumed equal to the return temperature of the heat pump condenser (often accessible).

Table 3.4 – Availability of variables from simulation to practical building MPC

Variable	unit	name	type	Availability
T^{in}	°C	air temperature	state var.	partial (staircases)
T^{f}	°C	floor temperature	state var.	-
T^{e}	°C	envelope temperature	state var.	-
T^{ret}	°C	SH return temperature	state var.	-
$T^{sto,top}$	°C	SH buffer top temperature	state var.	yes
$T^{sto,bot}$	°C	SH bottom temperature	state var.	-
$T^{dhw,top}$	°C	DHW top temperature	state var.	yes
$T^{dhw,bot}$	°C	DHW bottom temperature	state var.	-
T^{flow}	°C	supply temperature of SH	control var.	partial, f(setpoints)
$T^{dhw,set}$	°C	supply temperature of DHW	control var.	yes
$T^{src,in}$	°C	source inlet temperature	control var.	monitoring
S^{elr}	-	electric resistance switch	control var.	-
$\dot{m}^{sh}$	l/s	SH flow rate	control var.	monitoring
T^{ext}	°C	external temperature	disturbance	yes
Φ	kW/m^2	Solar irradiance	disturbance	models
G^{in}	kW	Internal gains	disturbance	theor. profiles
$\dot{m}^{dhw}$	l/h	DHW demand	disturbance	profiles/models

In theoretical MPC, setpoints are often neglected and the control variables are either the electrical power or binary inputs in the case of a non variable speed compressor. Junker et al. (2018) propose to represent the energy flexibility of buildings as a "dynamic function suitable for control", called the flexibility function. However, they make the strong assumption that the system providing flexibility is smart and able to respond to an external penalty signal, which is not yet true in most buildings.

Control variables from the MPC need to be translated into proper signals that local controllers can understand. Two strategies can be implemented in this regard: (1) integrate local setpoints into the MPC formulation or (2) transform control variables into setpoints in a post compute. These two strategies will be discussed in Chapter 4.

Finally, disturbances on-site can not be measured except for the external temperature. Internal gains rely on theoretical profiles and cannot be calibrated. For DHW, we do not have access to the consumed water $\dot{m}^{dhw}$. Back engineering of the water draw, based on temperature variations in the DHW storage tank, is practically impossible as we only have one sensor and tanks are highly stratified.

3.3.2 Specifications for data management and failure protocol

The following subsection presents the specifications we imposed on the system to safely test and control it. These specifications are the following:

- A reliable connection between the pilot site and the database to ensure data integrity
- A consistent connection between the centralised controller and the different servers
- An error prone software to handle data quality change (i.e missing data)
- A default protocol when centralised optimisation fails

Risk management and failure protocols

We imposed these specifications on the hardware and software to be able to perform automated tests with the least human supervision and correction as possible. We perform tests not only during the short working hours but also during long periods, including nights, when human supervision is not possible. In order for the tests to be automated and safe, we compiled a list of possible incidents that could lead to the failure of the test, their consequence, probability, mitigation and implemented intervention. This list can be found in Table 3.5. Whenever possible, we tried to implement an intervention that is automatically executed, but for some incidents this was not possible. In these cases an email is automatically sent to the operator, who can then execute a manual intervention.

Connection and value-up-to-dateness

All connections are checked regularly. If a connection interruption or a device error occurs, the operator is warned on the platform. He is also notified by email if desired.

Alarms

The alarms for comfort deployed on the online platform send an email:

- When the DHW boiler temperature drops below a limit value (e.g. < 30°C)
- When the SH buffer temperature drops below a limit value (e.g. < 25°C)

To avoid sending too many emails, one is sent when the value drops below the limit and one when the value rises back above. These alarms are easily set up. However, they are too sensitive, which is not ideal for an alarm. For example in the case of DHW, at almost each charging cycle, the temperature measured by the sensor drops below the limit and triggers an alarm. However, this is a normal behaviour as colder water is injected for the first minutes, leading the sensor value to drop. Reducing the limit value would not be an ideal solution as we would take the risk of discovering system failure too late. Therefore, we develop a second set of alarms that are triggered when the data extracted from the platform is pre-computed. These alarms check not only the absolute values but averages over longer periods. They can also combine different variables. They avoid having too many triggers while still ensuring comfort. However, they are only activated when the centralised controller is running during testing.

Incident	Consequence/Risk	Probability	Mitigation	Intervention
1. Lost connection with online platform	a) Unable to read nor write	small	- alarm (when platform is down)	1. Connection attempt with new token 2. Restoration of the platform connection
	b) Gateway do not reconnect to the platform			1. Check gateway reconnection 2. Reconnect gateway onsite
	c) Fixed setpoint (not to default value) - DHW too hot/cold - Overheating/Underheating - Heat pump inactive		- Reduce the range of offset value (+/-)for setpoint Temp - log file keeping track of the modified setpoints	1. Compare control setpoints to default value 2. If reconnection possible & not critical: --> resend manualy default setpoint values 3. If reconnection impossible or critical: --> Go onsite and manually reset setpoints
2. Lost connection with gateway	a) Unable to read nor write for a specific asset	small-medium		see 1.b) intervention
	b) Fixed setpoint (not to default value) - DHW too hot/cold - Overheating/Underheating - Heat pump inactive		- (+/-) offset for setpoint Temp - log file keeping track of the modified setpoints	see 1.c) intervention
3. Unable to connect to PC server	a) Unable to read nor write new values b) Fixed setpoint (not to default value)	very small	- Enable remote access of the PC server	1. Connection attempt to PC server 2. Check control setpoints values 3. Human assessment if Test can restart
5. Losing connection with weather forecast	Not being able to update MPC and to send accurate setpoint values	small	- storing longer forecast (3 days) - Warning	1. Restart weather server of HES 2. If problem remains stop the Test
6. Losing connection to indoor sensor (cloud IO)	No longer read indoor temperature	small	- storing buffer data	1. Connection attempt retry 2. Go onsite to restart raspberry Pi
7. Setpoints misstranslated	Unexpected behaviour Out of range setpoint values	very small	- Alarms on setpoints values - Feasible setpoints - Human monitoring	1. resend manualy the setpoint value 2. If persisting problem: --> Go onsite and manually reset setpoints
8. Overactivating an appliance	damage harware damage software (local controler)	small	- penalize deviation from default setpoint - penalize too frequent actuation	1. Go onsite check system 2. resend manualy setpoint default values
9. Wrong assessment of the state variables (MPC miss assessing the state of buildings)	a) Overheating building b) Overcooling the building	small-medium	- Alarms on indoor sensor - Human monitoring	1. Daily check (storage, el consumption): 2. If anomalies detected phone call 3. Stop Test for problematic buildings 4. If critical manually correct for comfort deviation with setpoint adjustment for a observation period
10. MPC gives no solution or is infeasible	a) Unable to provide new control setpoints	small	- Soft constraints to minimize the chance of infeasibility - perform status check for indentifying constraints causing infeasibility	1. Resend default control values 2. Remove Building causing problem from MPC framework and investigate the reason of the problem

Table 3.5 – Table of specifications for risk management and failure protocol

3.4 Validation of the software prototype

Once the software and hardware are installed and key monitoring data have been selected, we need to check that the the communication between the appliances on-site and the external software works as expected before the model predictive controller can be safely deployed. This validation was carried out through a series of operation field tests from June 2017 to March 2019, covering 7 of the 12 buildings in the project. These tests assess two main functions:

1. Communication check with buildings/systems
2. Heat pump activation and control

3.4.1 Communication check with systems

Communication checks are time consuming but nonetheless essential before one can offer DR services in real buildings. Each variable of a system has its own address in the communication framework. Operation field tests allow to check if the addresses link to the right systems. The standard protocol for all tests is the following:

1. Note the default value before any setpoint modification
2. Change its value for a specific period of time
3. Reset the setpoint to its *a priori* value
4. Ensure that the system is up-and-running

It might seem trivial to check that the addresses of all variables are correctly encoded throughout the framework but in reality it can be complex. For monitoring data this is often not critical but for control variables, errors can lead to system malfunction and heavy comfort violation as illustrated in the following example: In March 2019, we conducted an on-site test in two buildings. We tested if we could turn OFF the heat pumps using a setpoint related to their operation mode. This was confirmed as shown by the modification of SH cycle (lower panel of Figure 3.12). After setting it back to its *a priori* value (3 in this case), the distribution system started to run again, increasing the mass flow measured and moving the three-way valve. After a few minutes, and after increasing the room temperature setpoints, the heat pump compressor switched ON to produce SH. However, we did not realised on-site that the heat pumps were not able to produce DHW anymore. This led to a DHW shortage of a few hours in one of the buildings, as highlighted by the red oval in the top panel of Figure 3.12. One can also point out that the back-up electricity resistance did not kick in as it should have. This incident was the only time we ever received complaints from the inhabitants of the site. It threatened the possibility of conducting new tests.

Figure 3.12 – Temperature of the boilers for DHW and SH in Lenzburg during the system failure

After carefull examination, we found that the problem came from a display error on the platform, where two setpoints had the same name. The correct default value of the setpoint we tried to modify was 2 (Heizen/Kühlen/WW) and not 3. The value 3 does not exist in the manufacturer data-sheet for the controller. Surprisingly, however, the system was still able to partially run instead of throwing an error.

3.4.2 Heat pump activation and control

The last step to prove that our MPC framework can be implemented in existing buildings was to test if the heat pumps could be actuated through the identified control actions summarised in Table 3.6. In order to enable DR in buildings, we need to ensure that the thermo-electrical devices, in this case the heat pumps, can be remotely and dynamically actuated. Non-variable speed compressor heat pumps only have two operation states, ON and OFF. DR services require the heat pumps to change their default behaviours; forcing them to change their state (i.e. OFF to ON) or forcing them to keep their state constant (i.e. keep OFF).

We want to ensure that the control variable resulting from the MPC solution can be transformed into values understood by the heat pump controller. Therefore we conduct a series of tests by checking if heat pumps react to system modifications as predicted. For each test, we first decide on one of the four goals:

1. Switch a heat pump from OFF to ON
2. Switch a heat pump from ON to OFF
3. Keep a heat pump ON for a certain duration
4. Keep a heat pump OFF for a certain duration

Table 3.6 – Summary of the control actions and their value range performed on heat pumps at the Pilot Site

Goal	Setpoint	Actions	Range of X	Tested
Switch/Keep ON	Room temp.	Increase by X °C	[0 ; 30-ref]	☑
	reduced Room temp.	Increase by X °C	[0 ; 30-ref]	☑
	Return temp.	Increase by X °C	[0 ; 10]	✖
	DHW temperature	Increase by X °C	[0 ; 65-ref]	☑
	DHW preparation	OFF (=0) to ON (=1)	-	☑
	External temp.	Vary resistance value by X Ω (decrease)		-
Switch/Keep OFF	Room temp.	Decrease by X °C	[0 ; ref-15]	☑
	Room reduced temp.	Decrease by X °C	[0 ; ref-15]	☑
	Return temp.	Decrease by X °C	[0 ; 10]	✖
	DHW temperature	Decrease by X °C	[0 ; ref-45]	☑
	Operating mode	Heat/Cooling (=2) to OFF (=0)	-	☑
	External temp.	Vary resistance value by X Ω (increase)		-

To validate a test we change one of the setpoints presented in Table 3.6 following the same SG-ready principle presented in (Fischer, Wolf, & Triebel, 2017). We then visually assess on site if the heat pump compressor starts or stops accordingly. Non-monitored effects, like the time needed to start the circulation pumps and the compressors, are also measured to better help us tune in the next chapter the control framework. We consider a test as a success based on the pattern differences when setpoints values are modified versus when they are not. These differences can best be observed in the buffer temperatures and distribution temperature, combined with the power measurements.

Each combination of goal and setpoint modification is tested several times on each building. Successful control actions are marked by a tick in 3.6. For the heat pumps enabling remote control through return temperature setpoints, all tests were negative: no response from the system was obtained. The control action requiring to change the resistance value of the external sensor could not be tested. The hardware allowing the modification of the resistance value of the sensor, without losing the gage on the heat pumps, could not be developed on time.

Figure 3.13 presents the results obtained from a multi-family dwelling at the pilot site. In this case, we test if we can start a heat pump by (1) modifying DHW setpoints or (2) modifying the heat curve setpoint. For (1), we increase the DHW setpoint from 55 °C to 60 °C and we change the DHW heating status (dark green) from OFF(= 0) to ON(= 1). At the beginning of the test, the temperature measured by the sensor in the DHW tank (top yellow line) is 46.6 °C, which is close to but still above the lower bound of the dead-band.

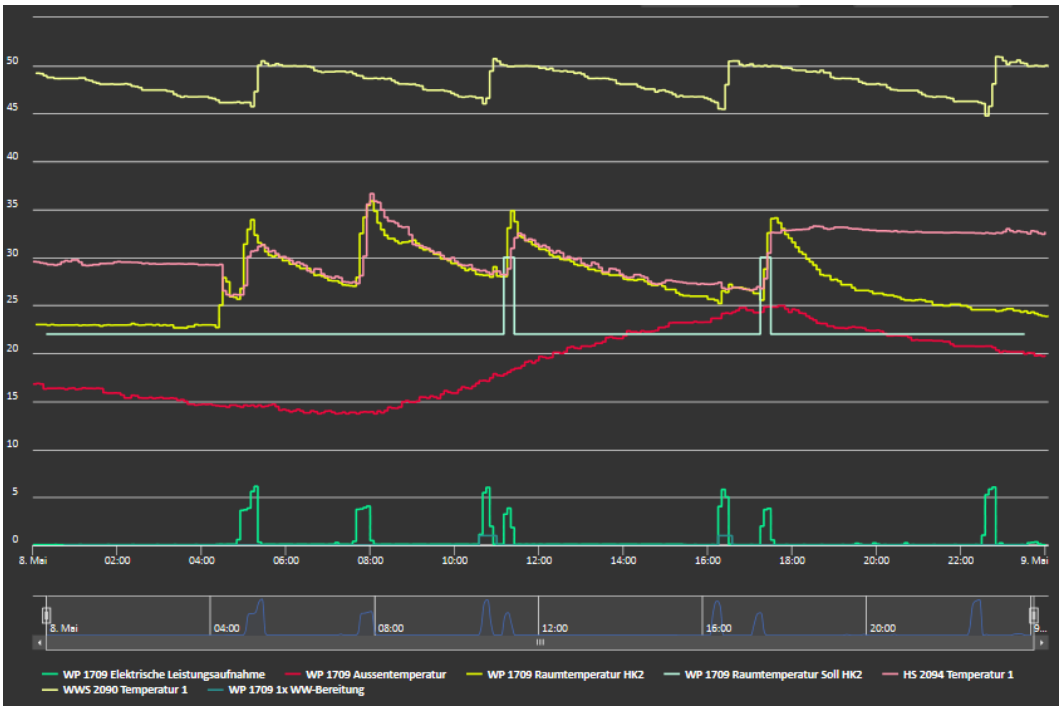


Figure 3.13 – Operation Test: Impact of changing the setpoint for the SH and DHW in a building

- After less than 2 seconds the circulation pump of the distribution system starts
- After 2 minutes 5 seconds the compressor of the heat pump starts (light green) implying a start of the heat pump based on a change in DHW heating status
- After 14 min 40 s the compressor stops by itself without any setpoint modifications. The temperature of the storage is 50.2 °C the same as the return temperature at the compressor side of the heat pump, which means that the storage tank is fully charged.

For (2), we increase the indoor temperature setpoint (white line) for the heating system from its default value (22 °C) to the maximum value accepted by the heat pump (30 °C). We perform this modification twice, and for both of them the heat pump starts as follows:

- After less than 2 seconds the circulation pump of the distribution system starts
- After 2 minutes 15 seconds the compressor starts

Setting the indoor setpoint temperature back to its default value leads to the stop of the heat pump compressor after 15 seconds. For the second time, we can see that the buffer temperature remains constant for several hours. This is a typical behaviour when the distribution system is OFF.

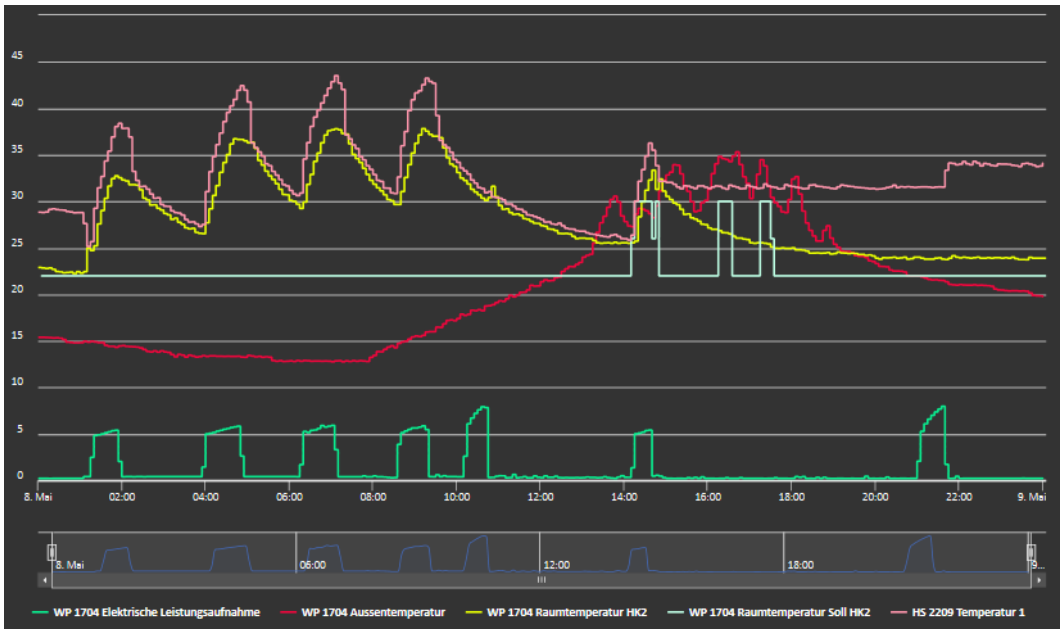


Figure 3.14 – Operation Test: Example of activation success and fail for starting a heat pump

However, increasing the indoor setpoint temperature to the maximal bound (30°C) does not always make a heat pump start as can be observed in Figure 3.14. When the external temperature is too high compared to the setpoint indoor temperature, changing the setpoint indoor temperature value has no effect on the heat pump nor the heating system.

Figure 3.15 presents the results of a test in which we tried to force a heat pump to remain shut down for several hours. In normal operation, the heat pump mode is set to heat/cooling (mode=2). Before the beginning of the test, one can observe that both cycles for DHW and SH are taking place. At 09:30, the setpoint value is changed to 0 (switch off). At 15:00, the heat pump is still OFF and did not start as expected since we explicitly set the setpoint value to 0 (=OFF). The temperature of the DHW tank is 43.7 °C, below the lower bound of DHW setpoints. The distribution pump, supplying heat to the building, has stopped during this period of time as confirmed by the non-changing, stable thermal buffer temperature of 32.6 °C. Stopping the heat pump compressor by changing the heat pump operation mode leads to a complete stop of the system both for SH and DHW. By doing so, the thermal capacity and inertia in the house is used towards providing flexibility and not the thermal capacity of the heat storage tank. Figure 3.15 also illustrates well the concept of “rebound effect”; as soon as the mode is changed back to normal, the heat pump starts to produce DHW. In the case of SH, the rebound effect can be delayed from a few minutes up to a couple of hours depending on the size and the amount of energy *a priori* stored in the buffer tank.

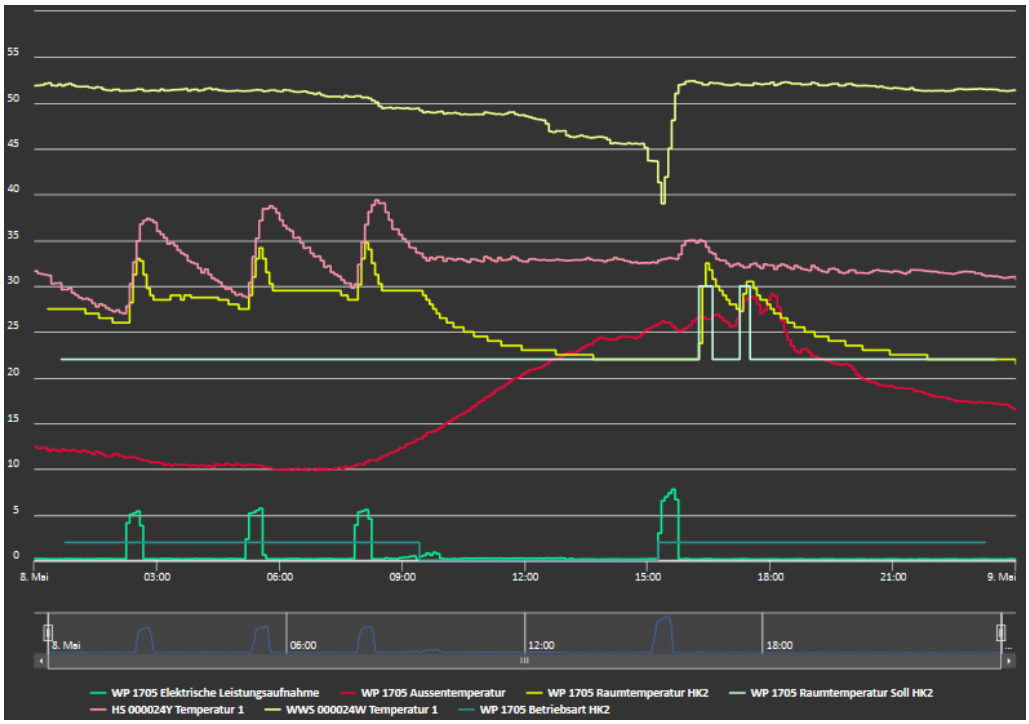


Figure 3.15 – Operation Test: Example of a long heat pump shutdown

Figure 3.16 and Figure 3.17 display the results of a test performed on the pilot site on the 2nd of May. Three buildings are actuated together via a prototype of a software developed in python. In this test the commands are generated manually every 15 minutes and processed by the software to actuate the system. Three control variables, represented as dashed lines in Figure 3.17, are changed in order to follow a predefined profile.

At the beginning of the test, the three heat pumps were running to produce SH, in the case of Montebello and Lenzburg, and to produce DHW, in the case of Grandson. In order to lower the electricity consumption to zero, the operation mode setpoints are set from “DHW & SH” to “Off” at 08:08 for each building. The three heat pumps stopped accordingly although they would have continued running without the actuation, as the temperature profiles show in Figure 28. Fifteen minutes later, the operation mode setpoints are changed back to “DHW & SH”. In order to restart the heat pumps and follow the tracking power, the temperature setpoint values have to be increased to raise the lower limit of the local deadband controllers. The ramping time takes a few minutes before the aggregated power reaches the tracking power. After another fifteen minutes, the temperature setpoints are changed back to their default values. Heat pump compressors stop after a few seconds.

A last actuation is performed at 09:23 to start two of the three heat pumps. In Figure 3.17, although the aggregated power seems to follow the tracking power, Figure 3.16 reveals that this is a coincidence. Indeed, the raise of the DHW temperature setpoints in Montebello and Lenzburg had no effect on the system and the heat pump started for other reasons.

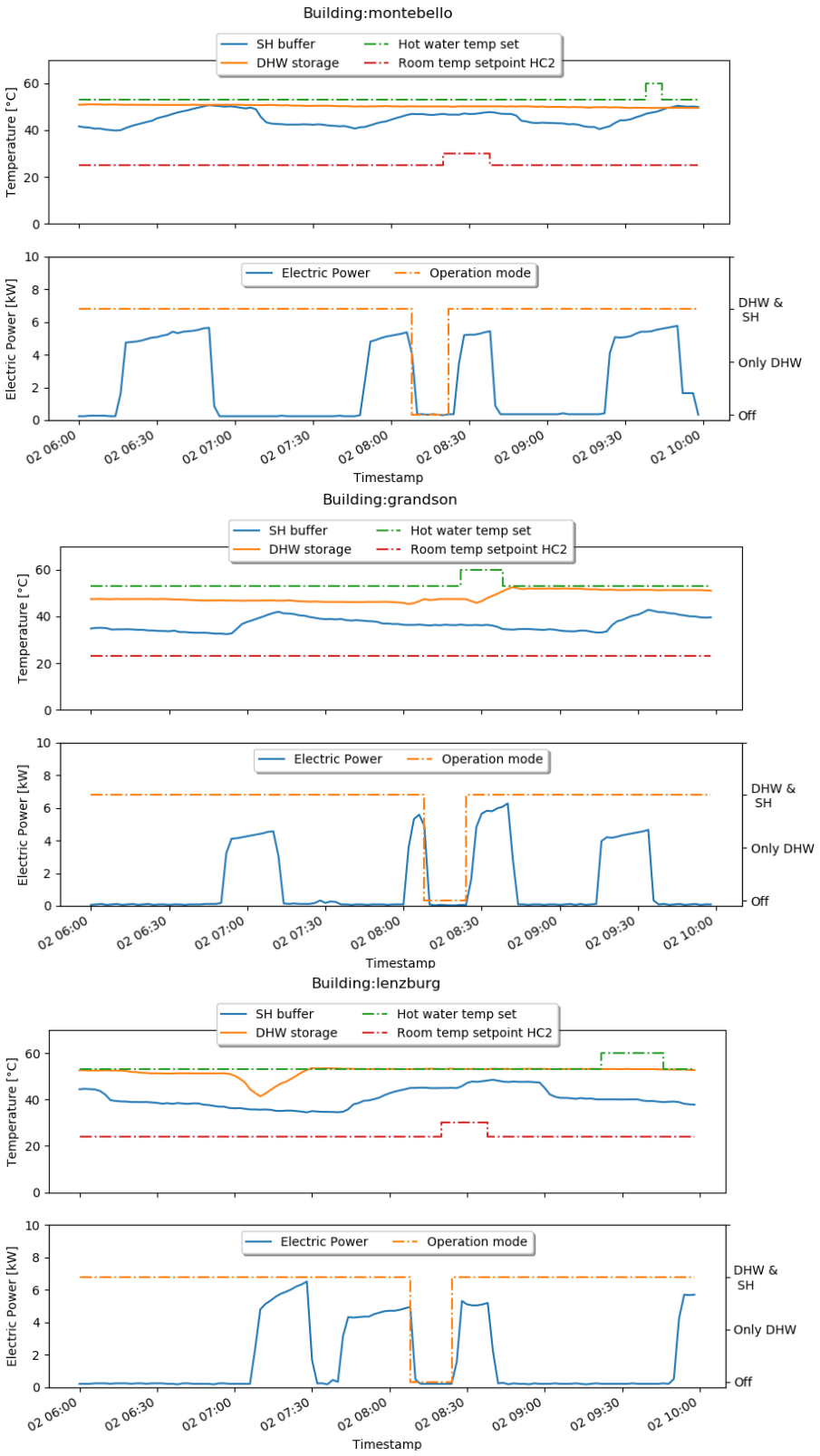


Figure 3.16 – Reaction to modification of control setpoints for three buildings. The doted lines are setpoints values.

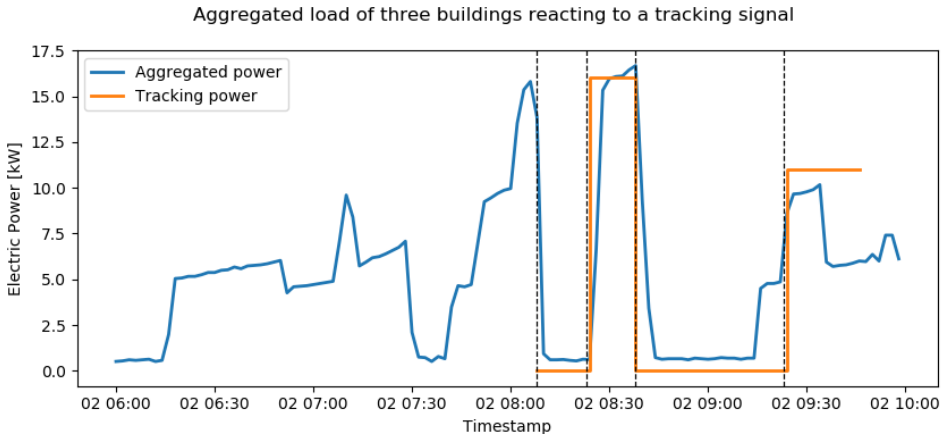


Figure 3.17 – Aggregated Load of three buildings reacting to a tracking signal

The main achievements and outputs of this chapter are the following:

- We proved that our system works in a real environment. However, only seven out of the twelve buildings have heat pumps that can be controlled and further tested.
- We imposed specifications on the hardware and software to be able to perform automated tests with the least human supervision and correction as possible
- We learnt about the physical and technical limitations of the system (e.g. heat pump compressors take two minutes to start)

From this point on, we want to test and understand how the centralised control operates. In Chapter 4 we present the methods and the tests conducted to assess the DR control system. In Chapter 5 we provide an example of DR services with the previously tuned system and assess its performance in Chapter 6.

4 Assessing the Demand Response control system

In the previous chapter we showed that we equipped residential buildings for monitoring and that we successfully enabled remote control of heat pumps. Here, we assess how well the pilot site operates once connected to our Demand Response control system using monitoring data and stand-alone tests. The chapter is structured as follows:

Section 4.1 describes challenges that need to be addressed when implementing Model Predictive Control (MPC) in real buildings. Section 4.2 details how weather and Domestic Hot Water (DHW) forecast can be improved to account for local particularities. Section 4.3 presents our model to infer heat pump power consumption from incomplete power meter data as well as an algorithm to estimate the state of the heat pumps. Finally, Section 4.4 characterises the flexibility of our system based on monitoring data and stand-alone DR service tests.

4.1 General motivation

The promising results found in simulation studies for providing predictive control services in buildings should not hide the fact that the implementation of MPC on real buildings faces considerable challenges as presented by (Cigler, Gyalistras, et al., 2013). When running an optimisation strategy for real buildings, as opposed to a theoretical MPC implementation for simulations of a building/set of buildings, the following aspects need to be addressed:

1. **Limitations of the devices at each site**: Many devices have internal control systems, which have to be interfaced with when used for the purposes of DR. This implies that, to change the behaviour of the devices, instead of standard on/off signals, the signals have to be transformed into variables and values that the devices can understand. There are two ways of dealing with this complication: either the model is adapted to account for the additional constraints with the risk of the problem becoming so complex that no optimal solution is found. Or, the model is kept simple and agnostic to local limitations but the optimal solution found might not be feasible in reality.

2. **Time duration allowed for the optimisation and rate of convergence**: DR services have to be satisfied within a short period of time. As a direct consequence, the optimisation steps also have to be concluded within this narrow period while ensuring either optimal or feasible sub-optimal set-points for the devices, thereby striking a balance between speed and performance.

3. **Model simplicity and adequacy**: The amount of thermal energy that can be stored in the building fabric is orders of magnitude greater than that the building's heat storage tanks. Therein lies the interest of using electrically powered building heating systems for DR services but also the need for using models that can reliably, yet simply, portray the dynamics of the transfer of heat from the heat pump, via the storage tank to the building fabric.

4. **Providing a good trace and tracking it**: For the use cases in Chapters 5 and 6, which require interaction with the DSO/TSO/market aggregator, particular care has to be taken to follow the trace which has been agreed upon while minimising user discomfort and failure of devices and/or part of the grids. This already needs to be considered when writing the algorithm.

4.2 Updating the forecast

4.2.1 Forecasting local sensor measurement

Model predictive control relies on the ability to accurately predict future behaviour. For residential buildings this means we need to predict the thermal losses and solar gains to assess the buildings' energy requirements. Thermal losses and solar gains depend on weather conditions. For real-time applications, fast accurate predictions are needed. We use an online forecasting platform, as presented in Section 3.2.3, for solar irradiation and external temperature.

Figure 4.1 shows that there is a discrepancy between the online external temperature forecast (in light grey) and the temperature measured (in yellow) by a heat pump in the building Lenzburg. This difference is often only a few degrees but can go up to more than 20°C at certain times of the day and for certain buildings. Since the local controller of the heat pump uses the value measured by the heat pump, we need to accurately predict this value.

To tackle the problem of weather discrepancy between the site location and the online forecast, we combine both the online and the offline (historical data) prediction to reduce the error. This also reduces the risk of missing input data for the MPC in the case of an internet service disruption. We predict the heat pump sensor value using Seasonal Auto-Regressive models with an eXogenous variable (SARX) for each heat pump. Here, we use the external temperature from the online forecast as the exogenous variable. SARX models capture some of the stochastic dynamics. The SARX model of order p with a seasonal period s and an exogenous

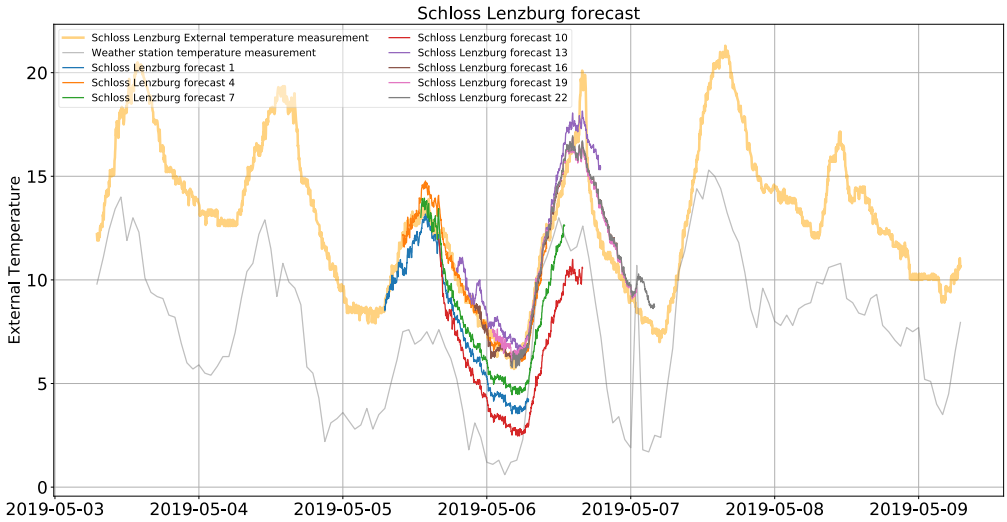


Figure 4.1 – Comparison between the outside temperature measured by the heat pump sensor, the online outside temperature forecast, and the predictions realised at different times using a seasonal SARX model (for 36 hours in May 2019)

variable x over a period n is formulated as:

$$y_t = c + \Phi_1 y_{t-1} + \ldots + \Phi_p y_{t-p} + \sum_{i=1}^{s-1} \gamma_i d_i + \sum_{j=1}^{n} \kappa_j x_t + \epsilon_t \tag{4.1}$$

Where d_i is a seasonal dummy indicator equal to 1 if $t \mod s = i$, $\Phi_1, \ldots, \Phi_P$ are the parameters of the model, c is a constant, and ϵ_t is a white noise process.

The parameters of the SARX model are inferred using the `AutoReg` model from the Python library `Statsmodels` (Seabold & Perktold, 2010). The method chooses the order of the model and the associated lag required. Once the parameters of the Equation 4.1 have been estimated, the SARX model can be used to forecast an arbitrary number of periods, but less than the model length (i.e. less than 36 hours). For the first time step, where the values are not available, we substitute the known preceding values y_{t-i} for i=1, …, p into the SARX equation. The output corresponds to the forecast for the first unobserved period. This time, we iterate the process using the predicted value arising from the previous forecasting step. A new model for each heat pump is fitted every hour.

Figure 4.1 shows the updated external temperature forecast of the temperature measured by the local controller of a heat pump in May 2019. Here we plot 8 of the predictions resulting from the SARX models updated every hour over a day. Forecast 1 corresponds to the forecast done at 9:00 on May 5$^{\text{th}}$ for the next 24 hours. The forecast 22 is updated at 6:00 on May 6$^{\text{th}}$. We see that the updated forecasts are closer to the measured values compared to the online prediction, especially for the first hours of the prediction. We can also see that consecutive days are not necessarily similar. A SARX model based only on historical value could not foresee

high changes as can be seen on the 5th and 6th of May. This justifies introducing the forecast of the online platform as an exogenous variable. We can observe that the first forecasts (e.g. 1, 4, 7 and 10) reproduce the trend of the previous days before the forecast readjusts itself and predicts a lower drop of temperature during the night before a high increase during the morning.

The forecasts for the other buildings are presented in Figure 4.2. Despite the fact that all the buildings are located within a vicinity of less than 200 meters, the external temperature measurements are very different for some buildings. For example, the temperature difference between the buildings Oberhofen and Grandson is more than 10°C at certain periods of the day. This can result from the location of the sensor (e.g. in the sun). Here, we can see that the local models are able to account for each system's discrepancy. However, as we do not use a moving average for the auto-regressive models, we sometimes end up with divergent results and noisy behaviour that can be seen as the oscillations in predictions for Jegenstorf or Oberhofen.

Figure 4.3 displays the external temperature forecast error as a function of the prediction window (x-axis) and the time of the day (colors) for the buildings at the pilot site. The y-axis represents the error between the predicted and the measured temperatures. The x-label represents the window of prediction in 2-min timestamps. Each color represents a forecast that starts at a given hour of the day (as shown in Figure 4.2). All forecasts last for 36 hours but we only plot the first 24 hours. As opposed to Figure 4.2, in Figure 4.3 the lines are shifted to start at the same point in the plot. This shift allows us to see that the error as expected increases with the prediction horizon. The black line is the mean average error of 36 forecast series and is slightly negative along the entire prediction horizon. This means that the temperature measured by the heat pumps is on average higher than the one predicted with our model. As the sensors are close to the buildings, the temperature of the air is warmer than the outside temperature predicted by the online forecast. Since we include the online forecast in our prediction, it could be the reason why we observe this negative error.

A comparison of the average errors of the external temperature prediction of the online forecast with the the average error of the local updated forecast using SARX models can be found in Appendix B Figure B.1. The forecast error over 24 hours is drastically reduced with the SARX models staying on average within a two degrees difference.

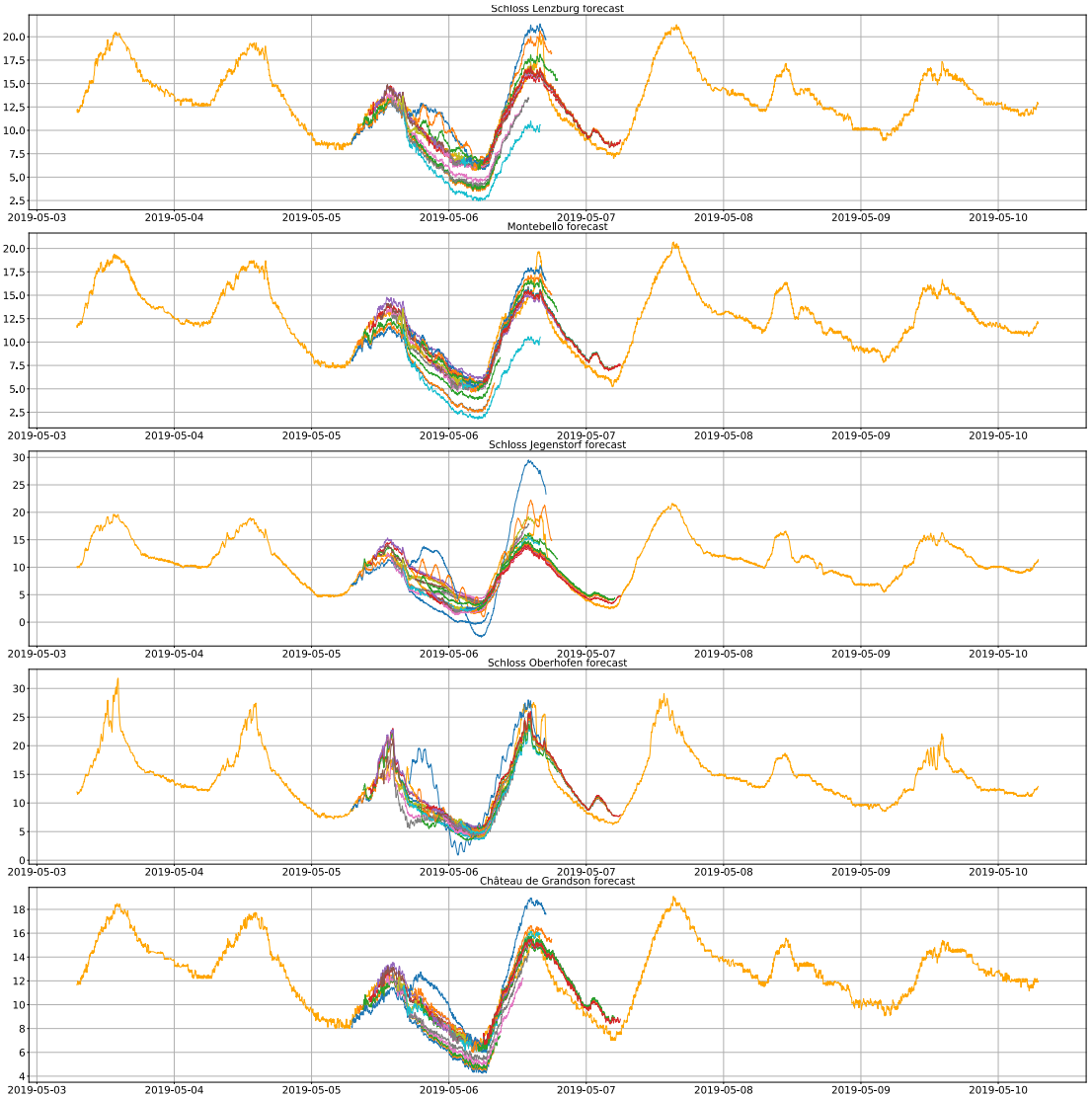


Figure 4.2 – Results of the seasonal ARX forecast model for predicting the outside temperature measured by the heat pump sensor (24 hours forecast in May 2019)

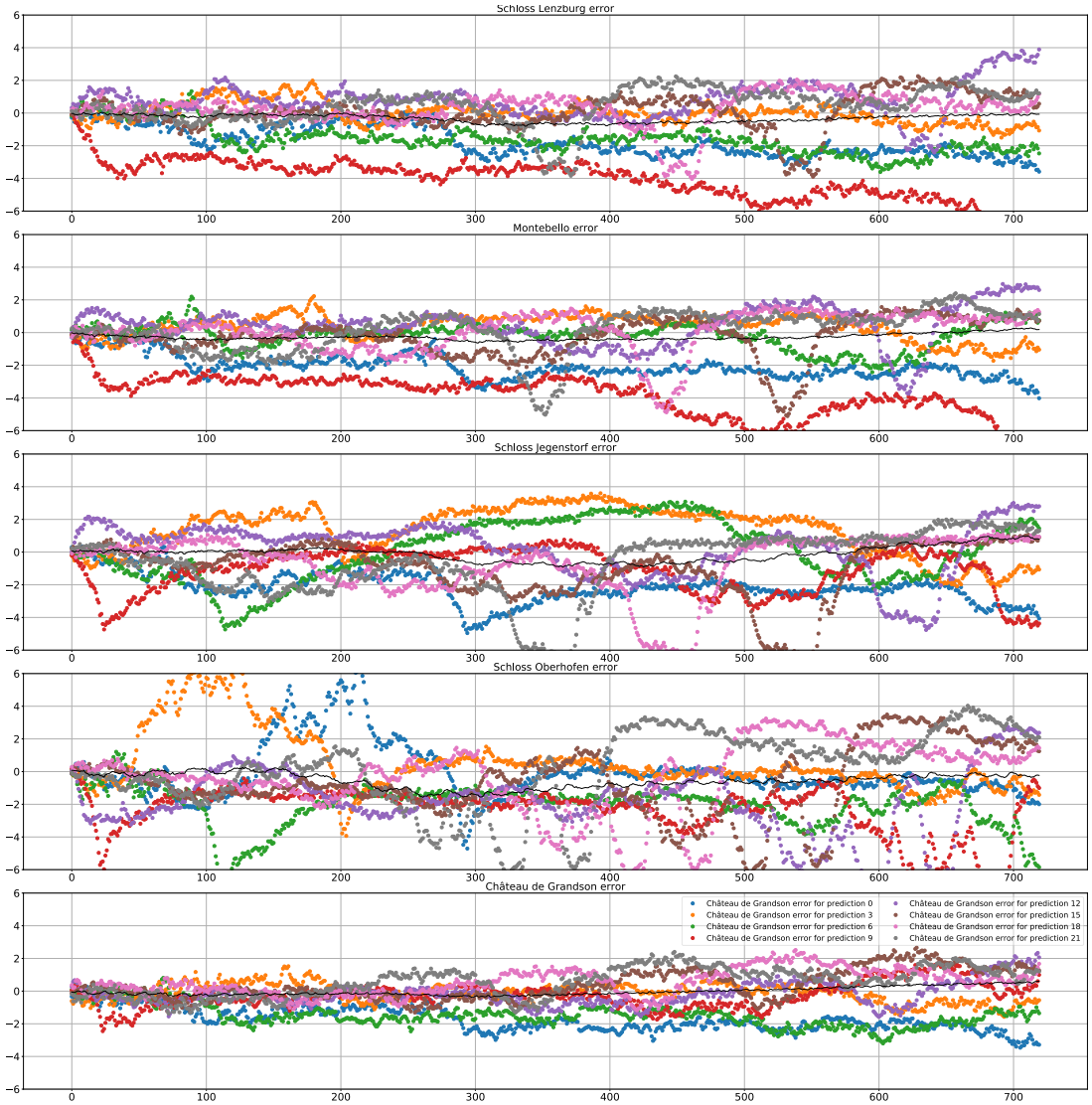


Figure 4.3 – External temperature forecast error as a function of the prediction window and the time of the day for the buildings at the pilot site. The y-axis represents the error between the predicted and the measured temperatures. The x-label represents the window of prediction in 2-min timestamps. The black line is the mean average error of 36 forecast series. The x-label represents 2-min timestamps.

4.2.2 DHW profiles

DHW consumption is by nature highly stochastic, but some patterns can be observed on a daily scale which can help us tune DHW consumption profiles using the monitored data. However, we do not have measurements of the draw volume of DHW so we cannot calibrate the DHW model in terms of l/min as it is usually done in literature. Instead, we use the thermal energy provided by the heat pump to the DHW storage. The thermal energy can give us three parameters to help us tune the DHW consumption profiles. The first one is the total energy consumed per day, including the draw volume and the heat losses of the installation. The second one is the probability of a DHW production cycle to happen at a specific time of the day. Charging cycle results most of the time from a high draw volume. Finally, it gives us the time between two cycles.

Figure 4.4 shows the activation state of the heat pump in Oberhofen as a discrete heat map (0: OFF, 1: ON) . The x-axis represents the day of the year (doy) and the y-axis the minutes of the day (mod). We can observe clear daily patterns but no seasonal trend. The heat map for two other buildings can be found in Appendix B Figures B.3 and B.5.

Figure 4.5 shows the mean activation state of the heat pump for DHW production over a day in Oberhofen. The mean of the state of the heat pump is computed over a year of measurements with 2 minutes sampling. A value of zero at a specific mod means that the heat pump was never producing DHW at that time of the day for the entire year. A value of 0.5 means that half of the time the heat pump was ON to produce DHW. The top panel is the average over the weekdays. The bottom panel is the average over the weekends. Note that the top panel has more than twice as many data points as the bottom one. Oberhofen has very specific features. As seen in Figure 4.4, there are three periods of 100 minutes (starting at 7:00, 13:00 and 19:00) with no DHW cycle. Right after these periods, there are peaks both during weekdays and weekends. This typically illustrates a peak shaving DR program with its rebound effects. This program can either have been set up directly in the local controller of the heat pump, or be the result of the use of the centralised switch by the local DSO. Weekdays and weekends are very similar, except for the peak at lunch time that does not exist for the weekdays. The results for two other buildings presenting particular features are displayed in Appendix B Figures B.4 and B.6. Figure B.4 displays a building with a single DHW charge per day while Figure B.6 shows a building with multiple random charging cycles happening throughout the day.

We have shown that estimating the DHW consumption purely based on theoretical profiles and number of residents does not account for local technical constraints (e.g. pipes preheating). Coupling the probability of DHW production with the time between two charging cycles could provide a method to generate DHW profiles. Assuming the same temperatures (e.g. energy state) at the beginning and at the end of a DHW cycle we can estimate an average DHW consumption over the cycle. With a sufficiently large set of data, we could compute density probability functions for different time intervals of a day. However, this method is could not be used on days with setpoint changes resulting from DR service activation.

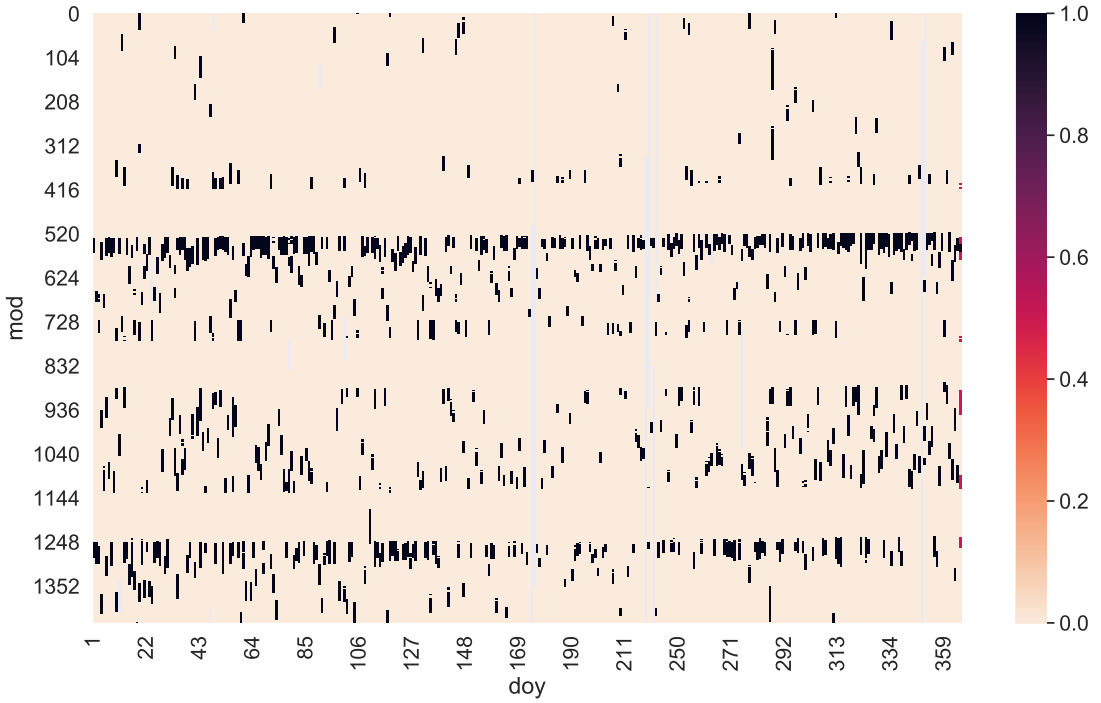


Figure 4.4 – Heat map of the DHW charging cycles for Oberhofen during a year in function of the Minute Of the Day (mod).

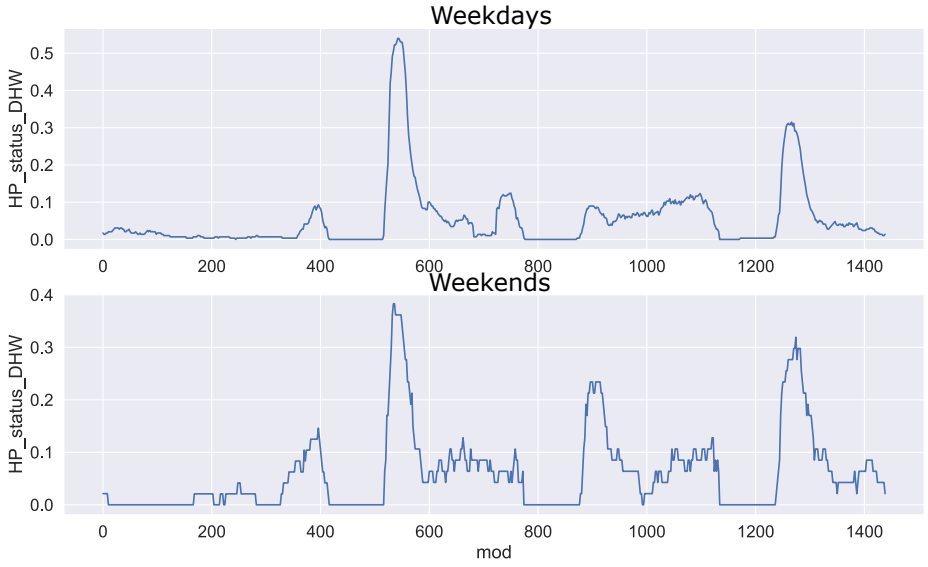


Figure 4.5 – Mean activation state of a heat pump for DHW production over a year of measurement (Oberhofen).

4.3 Updating the heat pump state

4.3.1 Heat pump power estimation

The power consumption of the heat pump is a critical variable when providing DR services. Its variation resulting from setpoint changes must be quantified to assess the performance and the reliability of the service. The power consumption of a heat pump can be modelled with a set of equations describing the physical cycles including fluid properties and heat pump components. This method does not require data. The consumption can also be assumed to follow a curve, and the parameters of the curve can be estimated based on observed data. Another way to infer the consumption is to rely on look-up tables of the power consumption at different temperatures, given by the manufacturer. In this project we use this last method as it is the easiest one to set up, does not require long history of monitoring data and is computationally fast. Once monitoring data are available, the data-sheet from the manufacturer can be extended or corrected with the onsite values.

All buildings have power meters, therefore monitoring the power should be as simple as connecting the meters to a communication device (e.g. a gateway). However, this project has encountered several issues in this regard. The first one is that it is common for heat pumps in residential buildings to not have their own electric meter. Therefore the power measurements include other devices that need to be filtered out, as they are mostly not controllable. The second issue encountered is that some power meters do not give power values but energy values on a quarter-hourly basis. The third issue, which has been the most penalising for the project, is that most of the power meters were unexpectedly replaced during the test phase as explained in Section 3.2.1. These smart meters could not be used due to the high cost of adapting the existing communication infrastructure.

To assess in "real time" the missing power measurements, we use the monitoring data of the heat pump itself and more particularly its operating temperatures. As detailed in Algorithm 1, we need the temperature at the evaporator side (i.e the source) and at the condenser side (i.e. the sink). For most water-to-water heat pumps, these measurements are available and are very sensitive to variations. We use the fact that when the heat pump is running, there is a difference in temperature between the inlet and the outlet at the evaporator side. In the first condition of the "**if**" in Algorithm 1, we check if there is a difference in temperature at the evaporator. If the heat pump is running, the outlet temperature should be lower than the inlet temperature. However, we also need to verify another condition: When the heat pump is not running, the temperatures measured in the pipes at the evaporator side quickly rise, as the environment (e.g. technical room) is at a higher temperature. Here, we set this temperature limit to 11°C. When the conditions of the "**if**" are met at time t, the power is set to 0. Otherwise, we use a 2d interpolation of the power from heat pump manufacturer data-sheets. As we cannot interpolate outside the temperature bounds provided in the data-sheet, we limit the source and sink temperature (lines 5 and 6 in Algorithm 1). The function `model` is presented in Appendix A.2.

Algorithm 1 Estimate the electric power of a heat pump

Input: $T^{src,in}$, $T^{src,out}$, $T^{sink,out}$, Heat pump parameters
Output: $\hat{P}_t$

for t = 1 to T **do**
  if $T^{src,in}_{t-1} < T^{src,out}_t$ OR $T^{src,out} > 11$ **then**
    $\hat{P}_t \leftarrow 0$
  else
    $T^{src} \leftarrow \max\left(10, \min(15, T^{src,in}_t)\right)$
    $T^{sink} \leftarrow \min\left(60, \max(35, T^{sink,out}_t)\right)$
    $\hat{P}_t \leftarrow model(T^{src}, T^{sink}, `type', `P')$ {lookup table and a 2d interpolation}

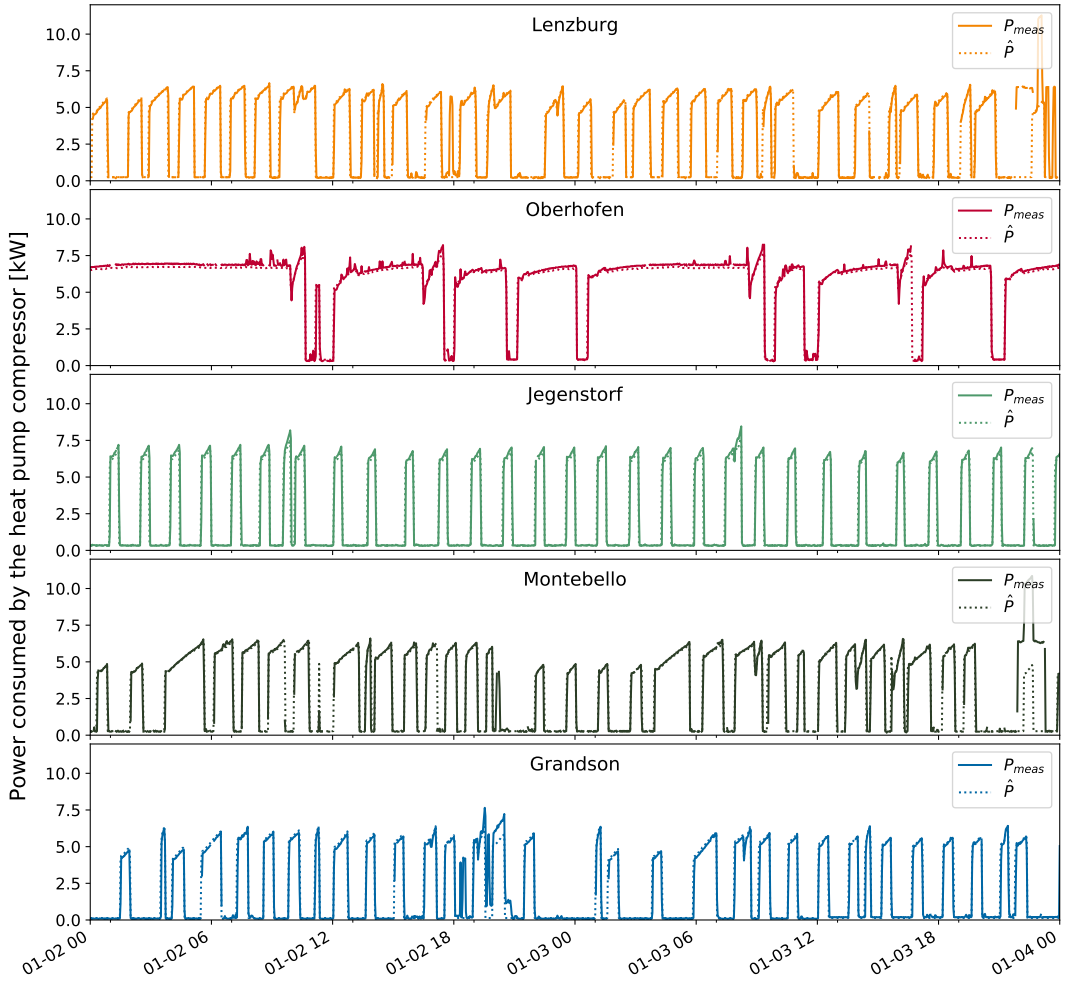


Figure 4.6 – Comparison between power meter measurements and model estimations for the power consumption of 5 heat pumps (based on Algorithm 1)

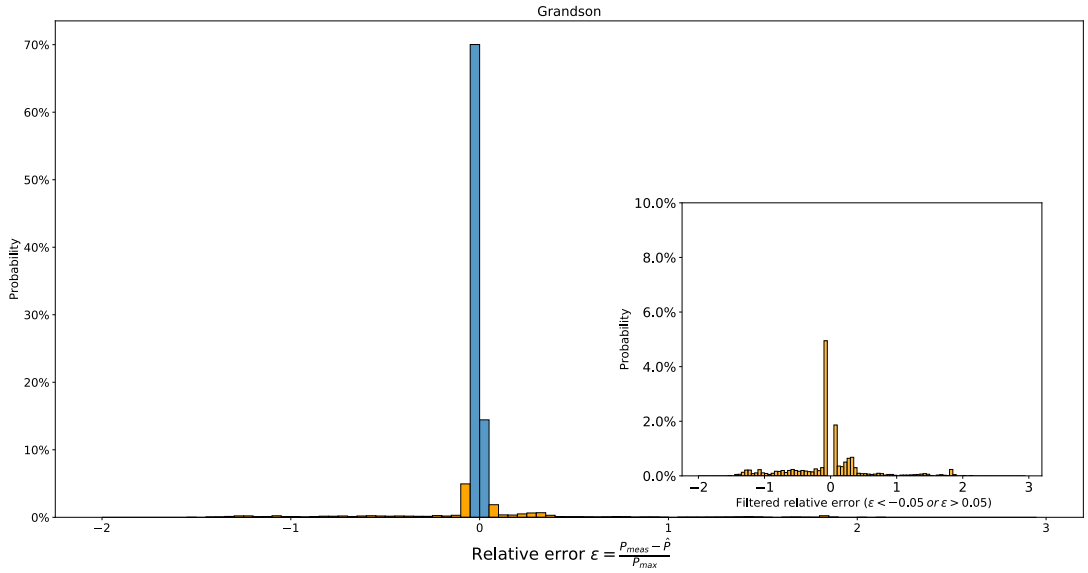


Figure 4.7 – Relative error distribution of the heat pump power model (based on Algorithm 1) for one heat pump over 20 days. The inside plot is a zoom of the relative error when its value is outside the 5% bounds.

Figure 4.6 presents the power consumed by the heat pump compressor measured (full line) and estimated (doted line) for 5 buildings over two days in February. We compute the mean of the measured baseline (residual power when the heat pump is off) and add it to the power estimation $\hat{P}$. Here, the power measured are raw data sampled every 2 minutes without interpolation, thus the non-continuous lines. The estimated power $\hat{P}$ is not affected by noise and can be used to complete missing power values.

In order to assess the quality of the model, we compute the relative error ε over 20 days as the difference between the measured power P_{meas} and the estimated power $\hat{P}$, divided by the maximum power of the heat pump (i.e. the power at 60°C). Here, we interpolate the measured value. Ideally, the relative error should be in the range of -1 to 1. However, due to monitoring errors (e.g. monitoring of unknown device power), this value can be outside this range. The relative error probability distribution for the heat pump in Grandson is given in Figure 4.7. The blue bars represent the ± 5% error bounds. 85% of the errors are within the 5% error bounds. The inside plot is a zoom of the relative error when $\varepsilon < -0.05$ *or* $\varepsilon > 0.05$. We can see a small shift towards negative errors, but the negative errors are very small. Based on this result, we can confidently accept the model. The relative error for the other buildings can be found in Appendix B Figure B.2.

4.3.2 How to identify when a heat pump is running?

Knowing if a heat pump is running to produce SH or DHW in a simulation is trivial, as it can be modeled and defined by a variable. This variable can then directly be used as input of a

MPC. However, such a variable does not always exist in the list of values monitored by heat pumps. Therefore, we develop an automated identification of heat pump cycles for SH and DHW based on the source temperature variations and storage temperature variations (see Algorithm 2). This algorithm is designed to have a short computation time such that it can be used in real time signal processing or when processing the power load for model calibration.

The algorithm is fairly simple but has proven to be very accurate. It can be decomposed in two main parts: (1) assess if the heat pumps are running (steps 4-8 of algorithm 2) and (2) assess if DHW storage tanks are charging (steps 9-12). In the first step, we compute the ΔT between the input and output of the source, in this case the Low Temperature (LT) thermal network. If this difference is greater than 2°C and the output temperature of the source is lower than the max seasonal temperature of the LT network (e.g. 13°C), it means that there is a heat exchange occurring at the evaporator and therefore that the heat pump is running. We therefore set X_n to 1. When the heat pump is not running, the monitored value of the source can increase at different rates up to 20°C due to different losses within the pipes, thus creating a $\Delta T > 2$. This is why we need the second condition to ensure that heat is extracted by the evaporator.

In the second step, we compute the moving average temperature with a window length w = 4 corresponding to 8 minutes. We then compute the gradient, i.e. the first derivative, of this moving average. If its last value is greater than 0.2 and the heat pump is running ($X_n = 1$) then we assume the heat pump is running for DHW and we set $X_n = 2$. Based on manufacturer data, a heat pump cannot run while the electric back-up resistance is running. This is why we assume that if the heat pump is running and T_{DHW} is steadily increasing, the heat pump is in a DHW charging cycle. The drawback of the method is that it depends on having a stable source temperature and it requires a sufficiently short sampling time.

Algorithm 2 Estimate the state of the heat pumps

Input: $T^{src,in}$, $T^{src,out}$, $\{T^{DHW}\}$
Output: X_n State of the heat pumps (0 = OFF, 1 = SH, 2 = DHW)

1: N ← Number of buildings available for demand response
2: X ← 0 {Set the heat pump to OFF (=0)}
3: **for** n = 1 to N **do**
4: $\Delta T^{src} \leftarrow (T^{src,in} - T^{src,out})$
5: **if** $\Delta T^{src} > 2$ AND $T^{src,out} < 13$ **then**
6: $X_n \leftarrow 1$
7: **else**
8: $X_n \leftarrow 0$
9: $g(x) \leftarrow \{T^{DHW}\}_n$
10: $\overline{g}_{\mathrm{w}}(x) = \frac{1}{w}\sum_{i=0}^{w-1} g(x-i)$ {Compute the moving average with w = 4 (8 min)}
11: **if** $\nabla g(x)[-1] > 0.2$ AND $X_n = 1$ **then**
12: $X_n \leftarrow 2$

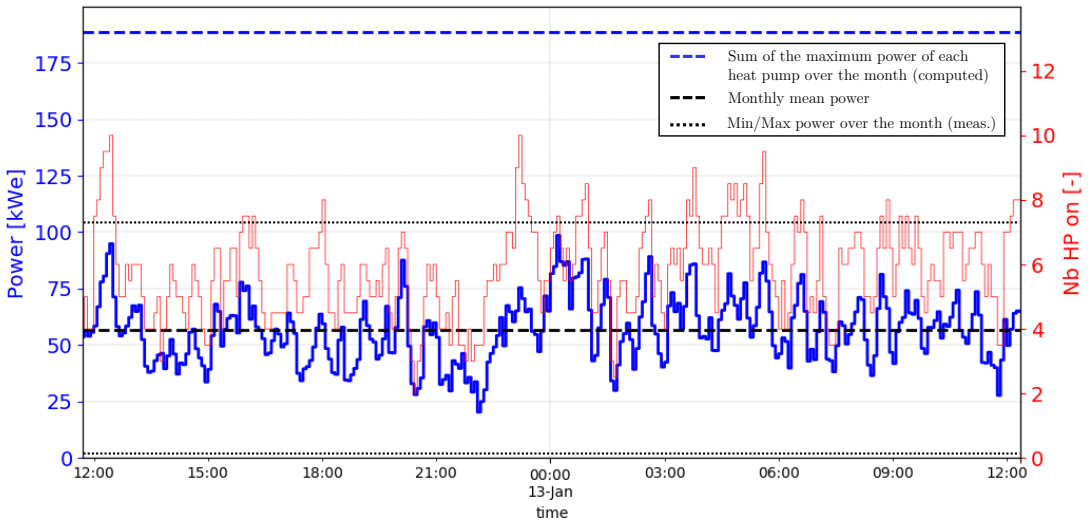


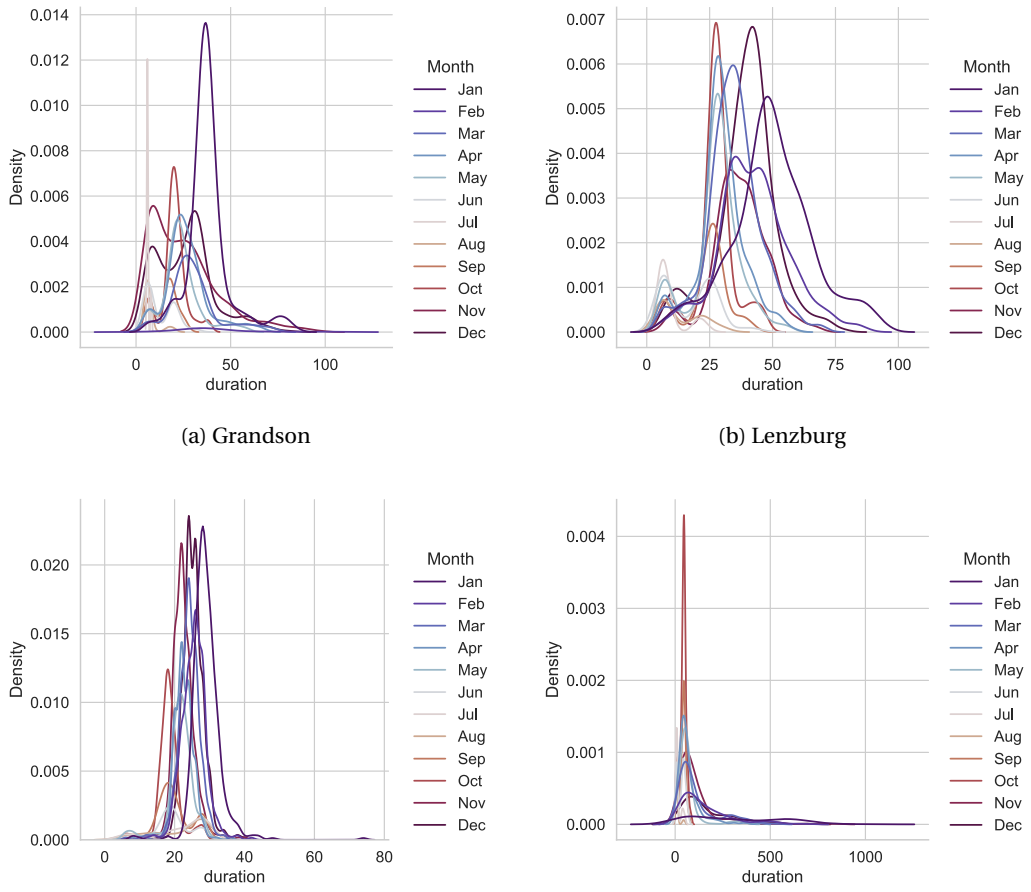
Figure 4.8 – Theoretical flexibility potential based on measurement

4.4 Assessing the flexibility

4.4.1 From simulation to reality

Figure 4.8 displays the aggregated power of the pilot site (in blue) and the number of associated heat pumps ON over a winter day. The black dashed line represents the monthly mean power while the black doted lines shows the minimum and the maximum aggregated power measured over the entire month. The dashed blue line represents the theoretical maximum limit (188 kWe), which is as if all heat pumps had their maximal consumption at the same time. It is computed as the sum of the maximum power measured over the month for each individual heat pump. The order of magnitude of the measured aggregated energy consumption has a similar order of magnitude as the simulation used in Chapter 2 (Figure 2.4). The intra-day variations are higher in reality than for the simulated buildings.

One can observe that the power consumption and the number of activated heat pumps in Figure 4.8 is not completely correlated. There are two possible explanations for this: first, all the heat pumps do not have the same nominal power and second, a heat pump's power consumption depends on its operating temperature that varies from 30°C to 60°C. One of the main findings here is that the theoretical power (dashed blue line) is almost twice the maximum measured power over this month. This indicates that in terms of power flexibility, the site has a large potential. However, as shown in Chapter 3, only 7 out of the 12 buildings can be equipped for DR control and 5 can be tested for DR services. This represents a maximum power of 32 kWe which is less than 20% of the identified potential of the site. Compared to the simulation of the pilot site, we already lost a considerable amount of flexibility.



(a) Grandson

(b) Lenzburg

(c) Jegenstorf

(d) Oberhofen

Figure 4.9 – Density distribution of the activity duration (in minutes) for space heating of the heat pumps over a year of monitoring

4.4.2 The cycle duration as a flexibility parameter

A key indicator to characterise the flexibility of a residential heat pump is the duration during which it can sustain a deviation. This information is particularly useful for estimating the energy that can be offered as DR service. From the duration of the heat pump cycle monitored, we can get a first estimate of the possible range of duration and narrow down the type of service and the typical duration bid we can offer to energy markets.

Figure 4.9 represents the density distribution of the heat pump cycle duration when they produce SH. The results are presented for the year 2019 and four different buildings. Each month is plotted separately with its own density distribution. Summer and winter months are represented with warmer and colder colors, respectively. The duration sampling is 2 minutes and we consider the cycles with a duration less or equal to 4 minutes as monitoring errors and we remove them. Due to monitoring issues, all months do not have the same number

of available data, which can sightly affect the shape of the curves. As expected, the warmer months have only few short cycles. Depending on the building, the colder months have either a higher number of cycles or longer cycle durations.

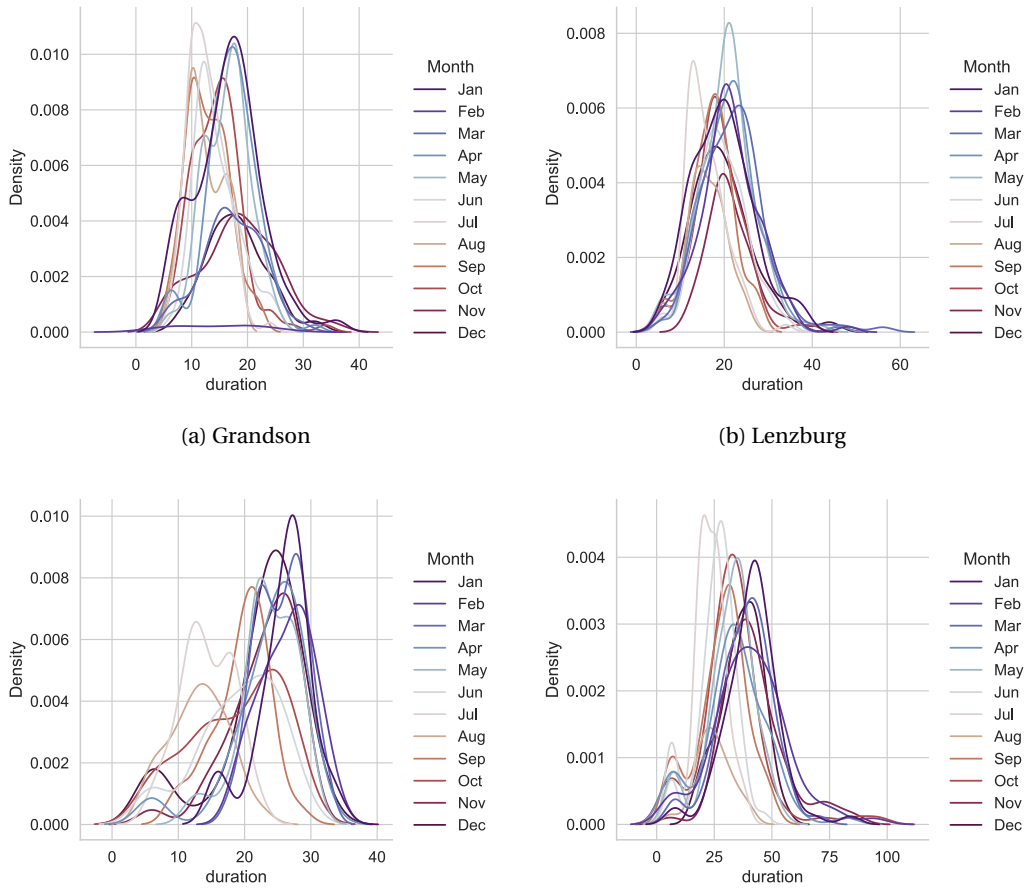
All buildings presented in Figure 4.9 have heat pumps with similar nominal power and SH buffers of similar size. Buildings (a) and (b) are identical with similar SH infrastructures. Buildings (c) and (d) are different but have the same heat pump and SH buffer installed. Although being similar on paper, buildings (a) and (b) have different cycle durations resulting from different control settings or device implementations (e.g. location of the pipes and thermal sensors). The building Grandson (a) compared to Lenzburg (b) has night setback implemented. Night setback is a widely used technique of reducing the energy consumption in commercial and residential buildings. It consists of using time-varying comfort constraints by relaxing them during night periods. This results in fewer and shorter cycles during the night. The identified night setbacks are included in the MPC formulations as system constraints.

The building Jegenstorf (c) does not display high variations over the different months and has a rather short cycle duration (less than 30 minutes) compared to the three other buildings. This is specific to floor heating systems, which have lower operating temperatures. In this case the heat pump is oversized (probably sized for DHW instead of SH), which results in a high number of short cycles. Building Oberhofen (d) presents an opposite behavior. In winter, the cycles are extremely long (up to 22 hours). This clearly indicates that the heating system is undersized and that the flexibility of the system is limited from November to March.

Figure 4.10 represents the density distribution of the heat pump cycle durations in minutes when they produce DHW. The results are presented for the year 2019 and four different buildings in the same way as in Figure 4.9. The duration of the cycle is proportional to the ratio between the capacity of the storage tank and the nominal power of the heat pump. DHW cycles present less variation over the year than SH cycles. The small shift between colder and warmer months can result from higher thermal losses happening in the technical room in winter. Although the tanks are often well insulated, this is not necessarily the case for the distribution pipes. A typical duration of a DHW cycle ranges from 15 to 45 minutes. Compared to a SH cycle, its duration can not be much extended. This is a first hint that DHW offers little flexibility.

4.4.3 Upward flexibility tests

To characterise the flexibility offered by the system, a series of stand-alone tests are carried out in the first months of 2020. The goal of these tests is to assess the main time constants of the heat pump systems (e.g. their activation time, their ramp-up and their maximum run time) and the maximal energy deviation possible. We tested different prediction horizons in order to select the most appropriate one for the MPC. An MPC with too long of a horizon took too much time to resolve and thus control actions could not be applied in time. Short horizons, although providing faster solution were less accurate. Vrettos et al. (2013) used a horizon of 16



(a) Grandson

(b) Lenzburg

(c) Jegenstorf

(d) Oberhofen

Figure 4.10 – Density distribution of the activity duration in minutes for Domestic Hot Water of the heat pumps over a year monitoring

hours since longer horizons did not lead to significant cost reduction. However, this horizon length was too long for our MPC formulation. We found 12 hours to be a good compromise.

In Figure 4.11 the external temperature average over the period is 6.3°C. At 09:00 the control variables for SH and DHW are set to their minimum value for 45 minutes before being raised at 09:45 to their maximum for 105 minutes. All heat pumps are successfully turned OFF then turned ON. Heat pump compressors take from 2 minutes to 10 minutes to start. Only two out of the five heat pumps run over almost the entire period. All heat pumps are turned OFF at 11:20. This is due to a direct control of the electricity provider switching the heat pump OFF via a centralized command. After resetting the control variables to their default values for 3 hours, the test is carried out in the afternoon as presented in Figure 4.12.

The external temperature average over the period increased to 9.4°C. At 15:00 the control variables for SH and DHW are set to their minimum value for 45 minutes before being raised

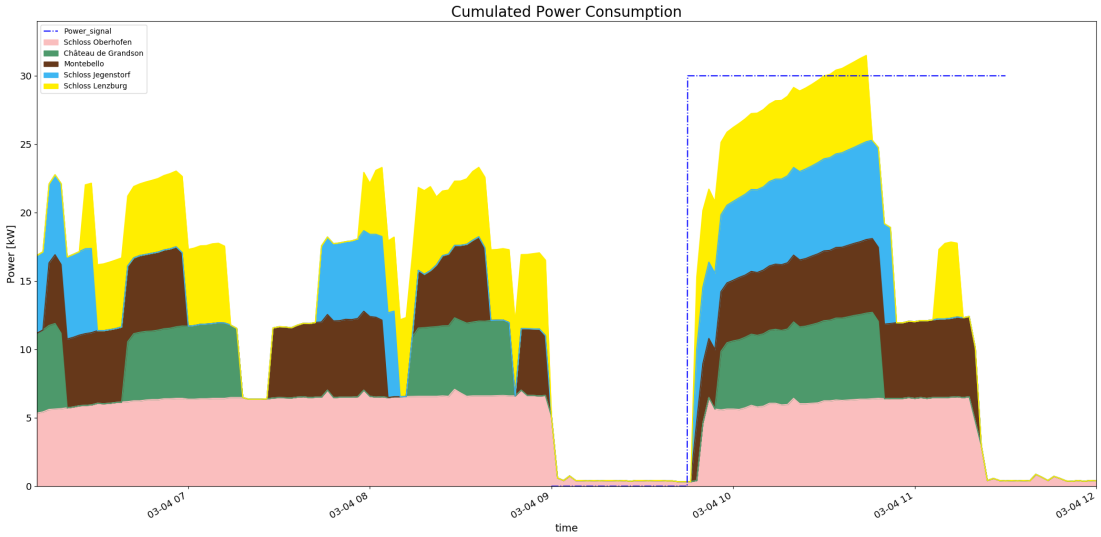


Figure 4.11 – Characterisation of upward-flexibility on the aggregation of 5 heat pumps (Morning March 4. 2020)

at 15:45 to their maximum for 120 minutes. All heat pumps except one are successfully turned OFF, then all are turned ON. This time, only one heat pump sustains the whole period. The ramp-up time to reach 80% of the power signal is fast (less than 2 min). From that point on, however, the power consumption increases only slowly before reaching the power signal value. It results from an increase of the operating temperatures. The power consumption of a heat pump with a non-variable speed compressor can hardly be maintained at a constant value.

These tests show that the flexibility parameters, namely the activation time, the ramp-up time and maximum run time are different for each building, even when the HPs installed are similar. The activation time typically ranges from 2 to 15 minutes, depending on storage temperature level or local controller logic (minimum downtime).

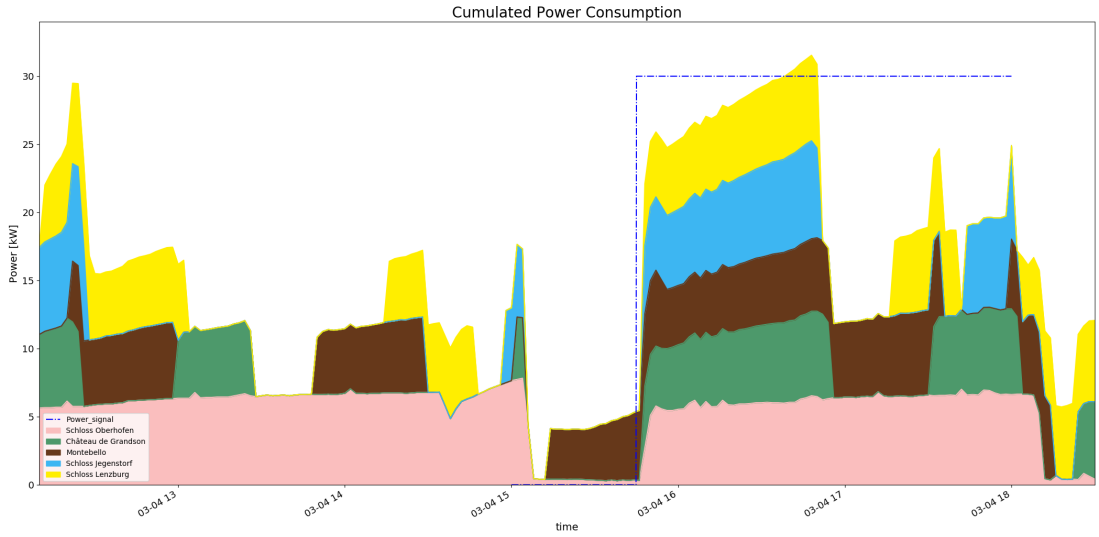


Figure 4.12 – Characterisation of upward-flexibility on the aggregation of 5 heat pumps (Afternoon March 4. 2020)

In this chapter we made use of monitoring data from a real pilot site to update our predictions, thus improving the forecasting used in the MPC. We found a way to infer the missing power measurements from the operating temperatures of the heat pumps. Confronted with the problem of not knowing the state of the heat pump, we developed an algorithm to infer it based on the heat pump operating temperatures and the variation of the DHW storage temperature. Through monitoring and stand-alone tests, we highlighted that the flexibility potential is limited by the cycle duration of the heat pumps. Now that we better understand the pilot site and assessed its flexibility potential, we will test how the DR framework we developed can be used in a relation with a third party player: the aggregator.

5 Using the demand response control system with a cluster manager

In the previous chapter we presented the tuning of the MPC framework to adapt it from simulation-based use case to a real time control use case. We have shown its ability to perform DR services at a local scale. We are now interested in using the DR control system at a larger scale that includes a third party to access energy markets. In this chapter we present the interface we created for the purpose of testing a transactive DR approach between the pilot site and a third party player in energy markets. The chapter is structured as follows:

Section 5.1 gives an overview of the energy markets in Switzerland, the future of the power system and its main actors, and introduces a new player: the cluster manager. Section 5.2 details how the aggregator and the cluster manager negotiate power loads and the technical implementation of this framework. In section 5.3, we propose a method to assess and share the availability of heat pumps. Section 5.4 describes a centralised control approach. Finally, Section 5.5 details the DR tests we performed.

5.1 Overview of the energy market and terminology

This section describes the main actors and energy markets of the power system in Europe and more specifically in Switzerland.

5.1.1 The traditional power grid and its actors

Traditionally, the power system consists of a grid, with different layers, that transport electricity from producers to consumers. Power generators compete in different energy markets. However, transmission remains managed by regulated commercial monopoly organisations called System Operators:

Transport System Operator (TSO): The TSOs are responsible for the exploitation and maintenance of the transport grid. They own the high voltage transmission lines and are responsible for securing and stabilizing the transmission network. The TSOs follow the market rules and

are regulated by the European Network of Transmission System Operators (ENTSO-E). In Switzerland the TSO is SwissGrid.

Distribution System Operator (DSO): The DSOs are responsible for the exploitation and maintenance of the distribution grid and for ensuring the quality of the supply at a local scale. There are more than 700 DSOs in Switzerland (Galus, 2017). In comparison, France and Germany have around 200 and 900 DSOs, respectively, for territories that are ten times larger.

Balance Responsible Party (BRP): The BRPs are market participants responsible for any imbalances of a group of wholesalers they represent.

In theory, most of electricity markets in Europe are liberalised and the end-user can freely make a contract with the electricity supplier (retailer) of their choice. An end-user can be a household or an industrial customer. He is connected to the low voltage distribution system and consumes electricity. An end-use customer is called "Prosumer" when (1) he possesses a decentralised production unit connected to the grid and (2) his production exceeds his consumption at certain periods.

5.1.2 The Swiss energy markets

Abrell (2016) made a detailed review of the Swiss wholesale electricity market. In this subsection we will present the main existing markets where DR services can play a role and assess their potential with regard to heat pumps.

Frequency Containment Reserves (FCR)

The FCR (also called primary control reserves) market for Switzerland has weekly bidding intervals and requires reserving a symmetric capacity band of at least ±1 MW for use on demand by the Transmission System Operator (TSO) with a full activation time of at most 30 seconds. In the case of heat pumps, their activation time is too long and with too many constraints. Therefore, only electric heaters (e.g. in DHW storage tanks) and distribution pumps could theoretically be used. The feasibility required to enter this market without batteries needs to be certified. To implement this service, simple models and fast computational times are mandatory.

Frequency Restoration Reserves (aFRR and mFRR)

The FRR (also called secondary control reserves) market for Switzerland has weekly bidding intervals and requires reserving an asymmetric capacity band of at least ±5 MW for the use on demand by the TSO (SwissGrid). During the following week, the bidder receives a control signal requiring at most 5 minutes for full activation in the case of automatic activation FRR (aFRR) and 15 minutes in the case of semi-automatic or manual activation (mFRR). This implies that to enter the secondary reserve market of Switzerland, we need to bid the power of at least 40 neighborhoods like the pilot site described in 3.1. Two levels of predictive control are

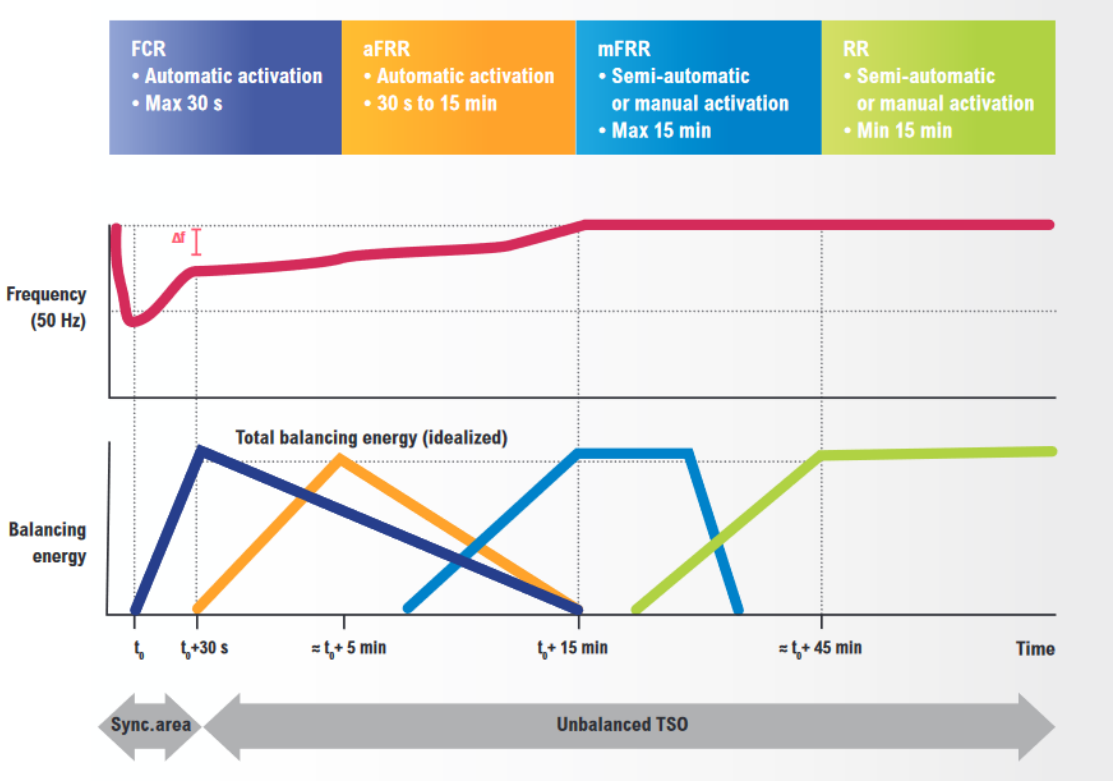


Figure 5.1 – Balancing markets process for frequency restoration (Figure 2 from ENTSOE, 2018)

needed: one weekly and one "real time". The secondary reserve market will provide incentives (for the requested increase or decrease), which will reduce the electricity price to zero or even below in some extreme cases. However, these incentives are very volatile, are made at unpredictable times and for unpredictable intervals. The benefits from such services should be carefully studied, including their potential negative effects on heat pumps.

Replacement Reserves (RR)

The RR (also called tertiary control reserves) is similar to the mFRR except that activation requests are sent by mail or phone and for longer time periods. This reserve needs to be completely activated within 15 minutes and last for at least one hour (up to a maximum of four hours). Participation in those markets, due to their long auction periods (weekly based), results in restrictions to participate in the rest of DR services.

Day-ahead and intraday markets

Alternative DR Services could be offered in energy-only markets regrouping two major market types: day-ahead auction (discrete) and intraday trade (continuous). For the first market, auctions are held the day before for a bid size of multiples of 0.1 MW and schedules need to be submitted to the TSO. For the intraday market, electricity is continuously traded (bid of multiples of 0.1 MW) to restore the balance between production and consumption resulting

from the deviation of other participants' schedule. BRP's can submit final schedules to the TSO until 45 minutes before delivery time. The bids must then be activated within 15 minutes. The price range for intraday market bids is higher than for the day-ahead market.

Balancing markets

The balancing market corresponds to the final platform through which TSOs settle the remaining deviation after the closure of the intraday market by operating the system close to real time. Figure 5.1 illustrates the main steps of the different balance markets.

5.1.3 The future of power systems

In power grids, demand and supply need to be balanced at all times to avoid frequency variations that in extreme cases could lead to black-outs (Haes Alhelou et al., 2019). The traditional approach of power suppliers to tackle this problem is top down: power generation is controlled to fulfill the energy demand curves. The growth of highly stochastic and decentralized renewable energy sources (e.g. photovoltaics, wind turbines) has increased the difficulty to achieve this power balance. An attractive option for future power systems is therefore to manage not only the energy supply but also its demand via "Demand Response" (DR) programs.

The future of Smart Grid is composed of units often referred to as Distributed Energy Resources (DER) (Raju P & Jain, 2019). DERs are defined as an indivisible set of installations possessing static and flexible loads such as heat pumps, refrigeration systems or electric vehicles. They can correspond to consumers, Prosumers, energy storage systems (e.g. batteries) or distributed power generation units. DERs require local controllers able to communicate with the rest of the system in order to enable their flexibility.

Current electricity market regulations do not allow DERs to directly participate in frequency reserve markets, as they require large sized bids that DERs cannot offer. However, if a large number of DERs are pooled together, their aggregated power can be bid in national markets. New BRP market players, referred to as commercial aggregators, are expected to manage such groupings of DERs. Certifications are required, ensuring that a certain volume of services are available (e.g. >5 MW for the secondary reserve market) and reliable. Certifications are delivered to the aggregators by the TSOs. Figure 5.2, adapted from (Halvgaard, 2014) presents the main actors in future power grids. Aggregators can play multiple roles in the market. They mostly pool the power of big consumers, such as commercial buildings and small industry, which offer larger capacity and are often well monitored, as well as distributed power generation units, such as photovoltaics and wind turbines. Recently, however, residential buildings have gained the attention of commercial aggregators, as they represent the largest cumulative resource but have little negotiation power on their own (Gkatzikis et al., 2013).

Commercial aggregators are BRPs operating a pool of DERs to sell multiple loads as a single unit in the electricity markets. They operate at different voltage levels, in different geographical regions (local up to national) and on different electricity markets (Vale et al., 2011). The DERs

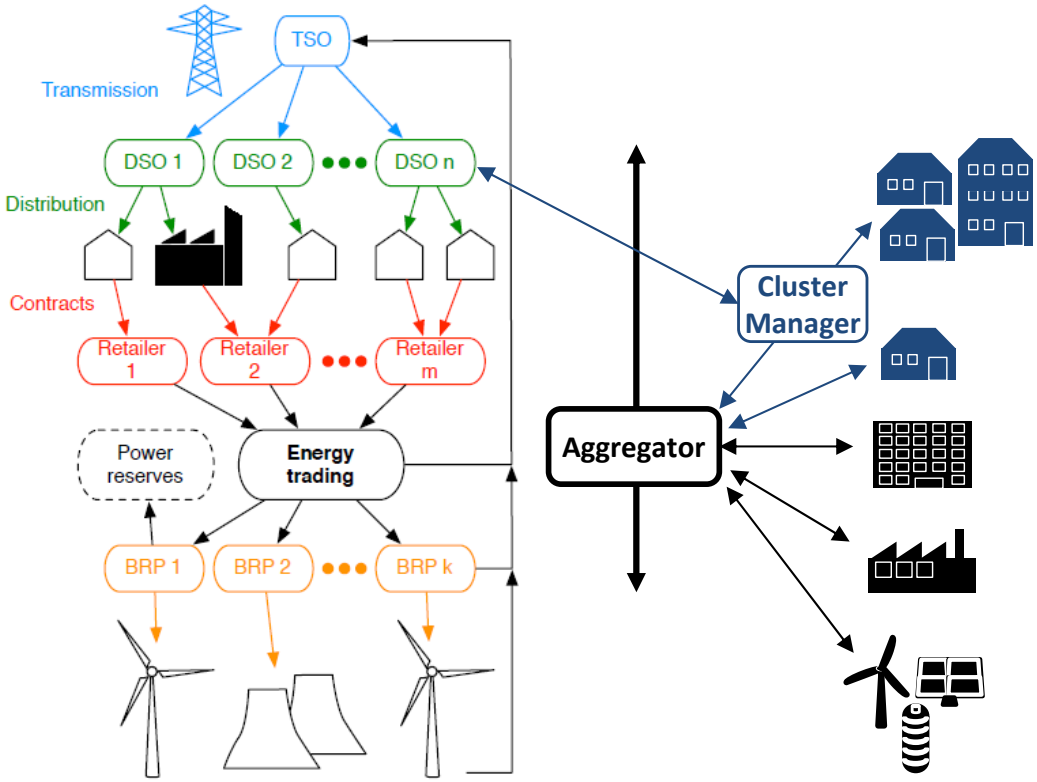


Figure 5.2 – The main actors in the future power grid (adapted from Figure 1.4 of (Halvgaard, 2014))

in an aggregator's portfolio are expected to provide their state and to respond to control signals communicated by the aggregator. Every week, the aggregator needs to assess the flexibility he will be able to provide during the following week. An aggregator will not know the detailed state of all the DERs he controls. He might only receive individual storage and power magnitude as well as the general state in which they are in (empty, full, charging, discharging).

To participate in an electricity market, an aggregator needs to make bids of the form: "`At time t, I promise to reduce/increase the power consumed P (kW) for a duration d if incentives i are greater than USD/kW`". The flexibility (up/down-reserve) is measured against the baseline consumption, and is impacted by the preceding day's consumption. Therefore, energy efficient operation results in lower DR income. In most existing markets the up-reserve and down-reserve are symmetric. Before the bidding closure, the aggregator needs to assess the reserve capacity that the pool of assets he controls can reliably provide to the TSO or DSO. He does this by solving a multi-period robust optimization problem. Because of the long horizon of prediction (day, week), which is subjected to high uncertainties for weather and consumer demand predictions, it is assumed that the system lies in a set that can be described in some quantitative way.

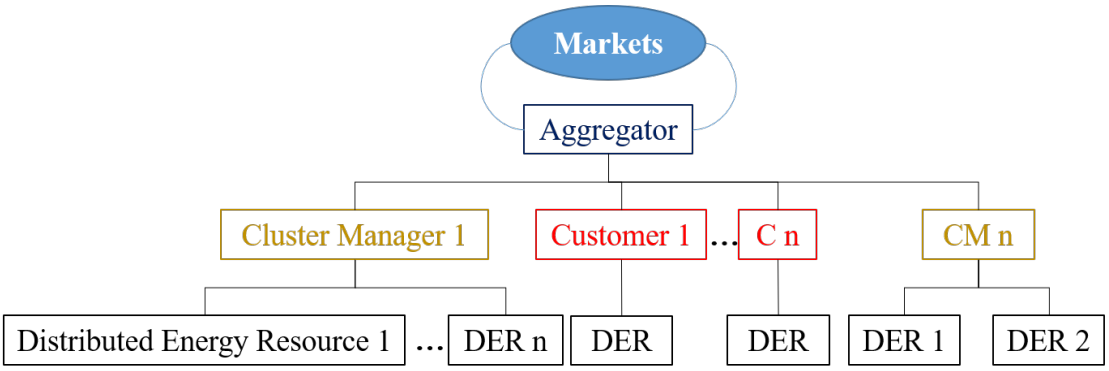


Figure 5.3 – Pool of single DER and clustered DERs in contract with an aggregator. Structure of the hierarchy control

5.1.4 A new player: The cluster manager

To use building assets as DERs, commercial aggregators need to make contracts with a large number of them. Since it is complex for a commercial aggregator to know the technical specifications of each DER in his portfolio and to assess their state in a reliable way, we introduce a new player called the Cluster Manager (CM). CMs are technical aggregators that cluster together local DERs (e.g. decentralised heat pumps in a low temperature district heating network). They have technical knowledge of the DERs, without having to deal with market specifications handled by the commercial aggregator. An example of a hierarchical structure of an aggregator's portfolio with CMs is presented in Figure 5.3.

Cluster managers can offer their flexibility to commercial aggregators. They can also offer peak shaving or local imbalance services directly to local DSOs to help them reduce penalties caused by schedule deviations. A cluster manager can deploy the communication and control framework and give the commercial aggregator direct access to its local pool of DERs. In this case, they mainly maintain the local infrastructure. However, they can also play the role of an intermediate between the DERs and the aggregator by aggregating and disaggregating load profiles and using their technical know-how to increase DR benefits while ensuring comfort.

In this work, we play the role of a cluster manager responsible for distributed heat pumps connected to a low temperature district heating network. We cluster their loads for a commercial aggregator (REstore), who is a partner of the project. We do not give the aggregator direct access to each DER, but instead perform centralised optimisation. We are not directly bound to restrictions of energy markets as they are handled by the aggregator. For example, water-water heat pumps on their own do not meet the activation time requirements for secondary reserve market. However, they can still be used in this market if they are artificially coupled with batteries by the aggregator. Requirements regarding the symmetrical aspect of up-reserve and down-reserve can hardly be satisfied by a small pool of heat pumps and are therefore not considered by the cluster manager.

5.2 The communication framework

5.2.1 Technical implementation

The interface and communication protocols presented in Figure 5.4 are adapted from the one described in Chapter 3. A minimum of intelligence installed on the gateways allows the information to transit through a platform to the cluster manager, who then centralises the data for all of the 5 buildings tested. In Figure 5.4, there is only one management platform but in reality we use a second one for indoor monitoring with its own communication specifications.

The cluster manager is the interface between the pilot site and the aggregator. The cluster manager is deployed on a Linux based computer at the HES-SO and can be remotely accessed. It handles communication between the different services, processes and analyses the data and performs the optimisations. Two processes are run in parallel. One takes care of the interface with the aggregator (see Figure 5.5) while the other one performs the data processing and the MPC (see Figure 5.6).

The local grid operator is not part of the project although he can be, as is the case for research partners working on a different pilot site in Germany. However, we represent its connection with the site in Figure 5.4, as the local grid operator can remotely switch OFF the power supply of the heat pumps to perform peak shaving on its own. This has an indirect influence on the reliability of DR services offered by the cluster manager to the aggregator.

5.2.2 Aggregator and cluster manager negotiation over power trace

An asynchronous communication between the aggregator and the cluster manager is established by using Apache Kafka. Apache Avro is used to encode the actual messages that are exchanged. The aggregator REstore has implemented this in a Python-based software running on a virtual machine. Figure 5.5 presents the flowchart of the communication and decision process between the aggregator and the cluster manager. At the beginning of each test, communication agents are started at both the aggregator's side (in light blue) and the cluster manager's side (in green). Both ends automatically check their connection and start to wait for messages to be exchanged. With this communication framework, the aggregator can listen to multiple cluster manager agents or multiple single DER agents at the same time.

For simplicity, in the case of daily and weekly tests, calls are scheduled every two hours. However, the framework can also handle calls at random times. When the system is not ready to provide flexibility, no message is sent to the aggregator and the system continues to operate normally. To simulate the effect of not knowing when a call from the aggregator will arrive, the system is not controlled before the call (i.e. the buffer is not filled or emptied in advance). Although we assume that the time of the call can be unknown to the cluster manager, it cannot be less than 15 minutes before the actuation. The communication with the aggregator is based on three types of messages:

1. **Flex Request:** sent by the aggregator to the cluster manager. It initiates the DR service. The request is always sent 15 minutes before its due time. Once sent, the cluster manager has 12 minutes to answer back the number of requested traces. An example of the structure and information contained in a Flex Request is given below:
{*"requestId": "request-2020-04-13 16:00:00.092440", "numberOfTraces": 3, "target": null*}

2. **Flex Offer:** sent by the cluster manager to the aggregator in response to a DR call. It is structured as a dictionary with keys to access its different elements. It contains the number of requested traces, each associated with a unique ID, a cost and the values of the trace itself. The values represent power consumption in kW over 15-min intervals. The first trace is the forecast baseline, as if the system were running Business as Usual (BaU). The second trace is the upper bound power consumption of the duration of the call. The third is the lower bound. The cost accounts for the electricity consumed (fixed price considered here), the comfort and the actuation cost. If the cost for all traces are similar it means that the flexibility potential of the system is considered too low. If this is the case, no Flex Offer is sent back to the aggregator.
{*"requestId": "request-2020-04-13 16:00:00.092440", "traces":*
[{*"traceId": "trace-001", "cost": 0.0, "trace": [0, 4.79, 0, 0]*},
{*"traceId": "trace-002", "cost": 5.21, "trace": [25.90, 22.80, 0.0, 0.0]*},
{*"traceId": "trace-003", "cost": 2.05, "trace": [0.0, 0.0, 0.0, 0.0]*}]}

3. **Activation:** sent by the aggregator to the cluster manager in response to a Flex Offer 10 to 12 minutes after the Flex request was sent. It contains the request Id and the “trace Id” of the Power trace selected by the aggregator. The aggregator chooses a Power trace randomly among the last three proposed by the cluster manager. The cluster manager engages in tracking the selected trace starting from the next timestamp, exactly 15 minutes after receiving the Flex Request. In the eventuality of not receiving an activation from the aggregator, the cluster manager continues to perform BaU.
{*"requestId": "request-2020-04-13 16:00:00.092440", "traceId": "trace-002"*}

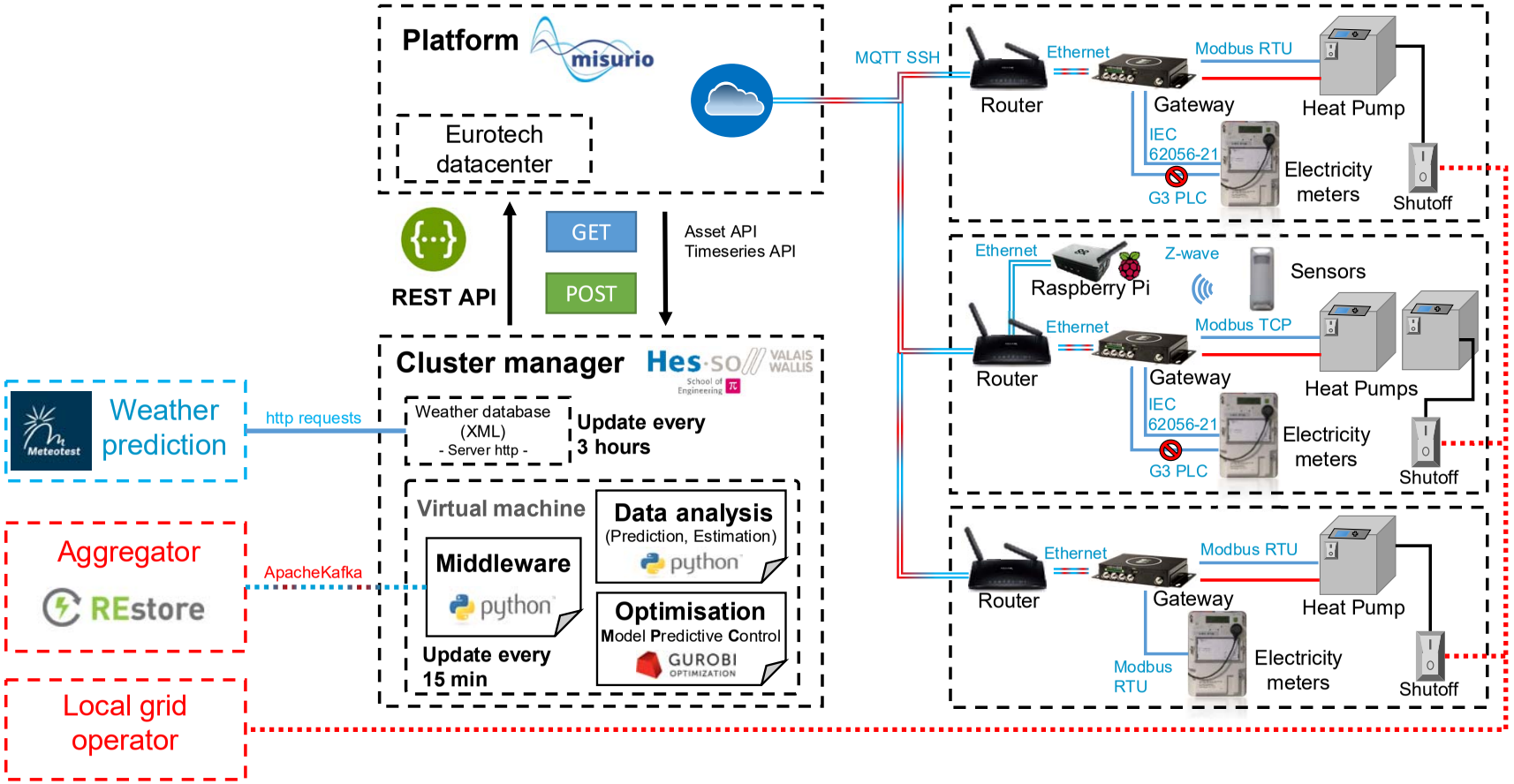


Figure 5.4 – Interfaces and communication protocols of the test case with a cluster manager

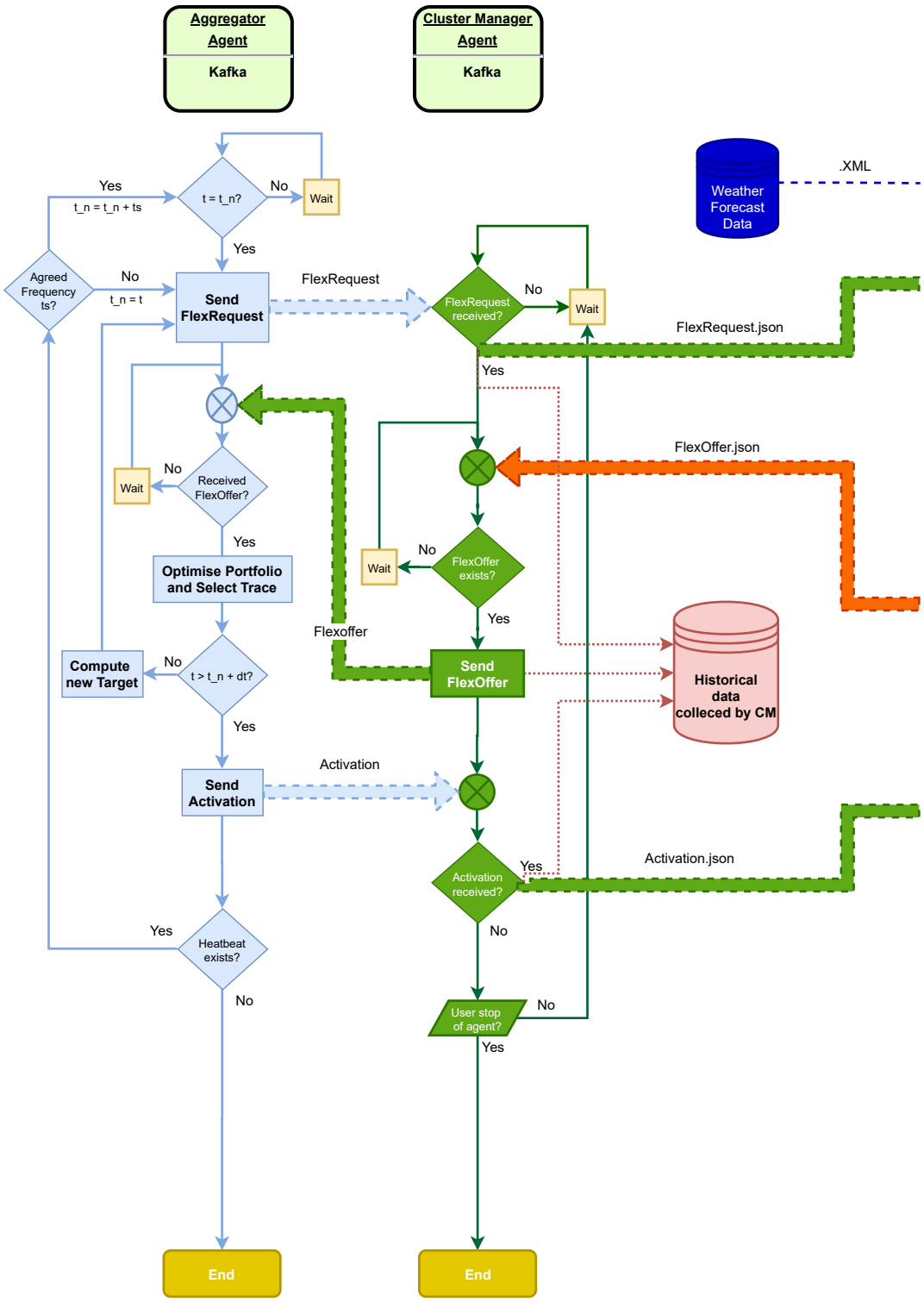


Figure 5.5 – Flowchart of a prototype DR service software (Aggregator and cluster manager interconnection)

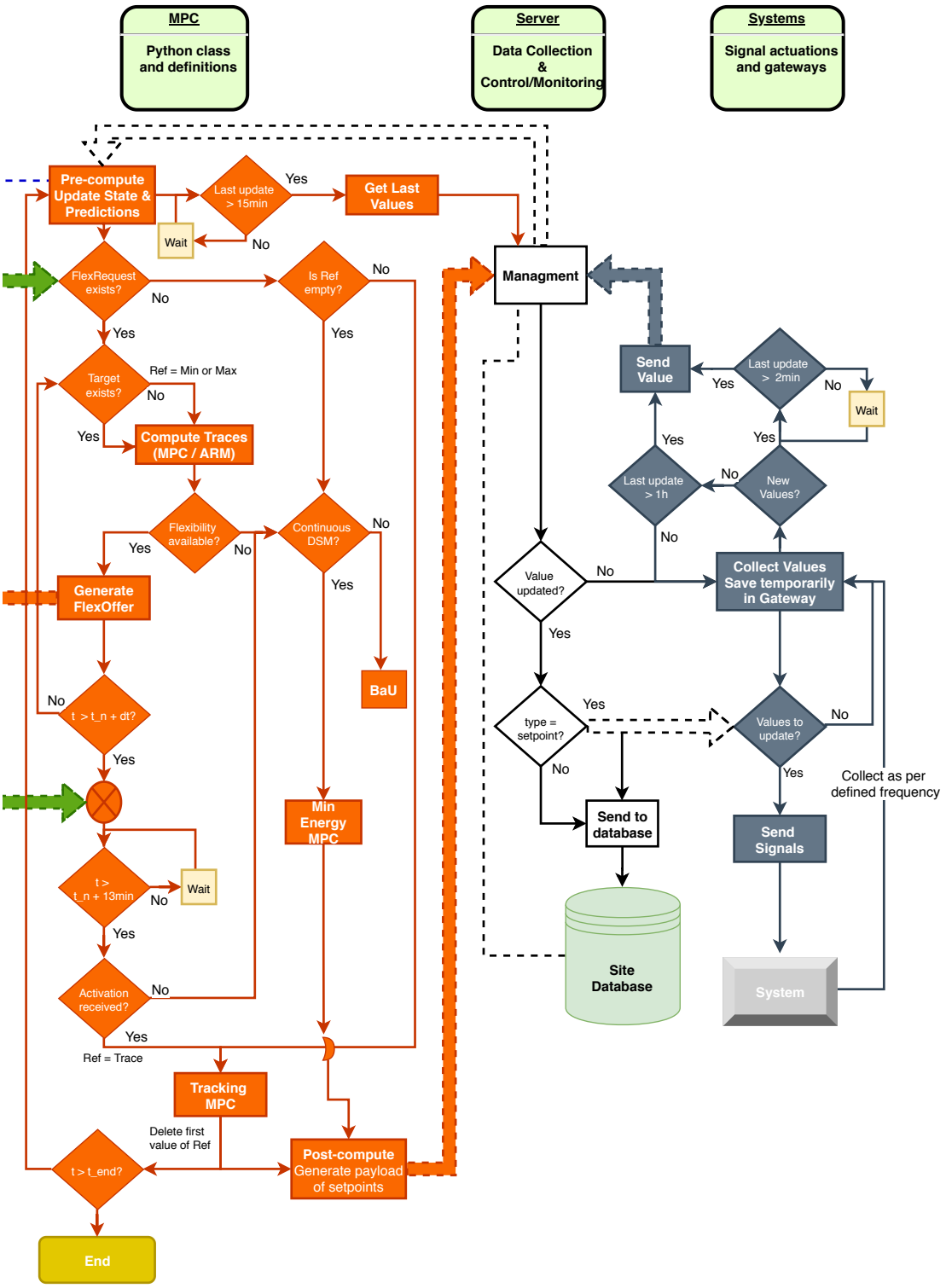


Figure 5.6 – Flowchart of a prototype DR service software (Cluster manager and Pilot Site)

5.3 How to assess and share the state and availability of a system?

5.3.1 Updating the states and the forecasts

To update the states of the system we need to collect the last values from the pilot site before each optimisation (i.e. at a 15-minutes basis) or whenever a Flex Request is received from the aggregator. We check the last time the values were updated on the platform. If it is not recent enough this means that they are unresponsive, so we temporarily remove the buildings from the test. We then update the DHW prediction of each building based on weekly profiles and the internal gains G_{in} based on semi-random daily profiles (described in Section 2.5.2). The weather forecast is updated every hour using an online platform. The new forecast is used in the MPC formulation to predict the thermal losses of the building, as well as its respective heat curve. As presented in Chapter 4, we observe a discrepancy between the online external temperature forecast and the temperature measured by the heat pump. The local control uses the value measured by the heat pump. Therefore, we use the online forecast as an exogenous variable in auto-regression models to predict the external temperature measurements from each heat pump sensor (see Section 4.2).

Sensor values cannot directly be plugged into the model of the storage tanks nor the building. We compute the initial model states and update the model constraints with the algorithm presented in Section 4.3. This algorithm infers if the HPs are currently running, and if so, if they are in DHW or SH mode. Finally, we update the building models (T^{in} and T^{m}) by simulating one step using the thermal power estimated from the sensor values.

5.3.2 The baseline consumption

Characterising the energy flexibility of buildings is challenging and there are various methods to enable and evaluate the successful implementation of a DR program. Most of them rely on the determination of the baseline load (*a priori* or *a posteriori*). The baseline load for a residential building is the expected consumption pattern in the absence of DR programs. This load can be separated into two parts; the controllable load, also called responsive load (e.g. Heat pump, DHW electrical heater or smart appliances) and the non-responsive load (e.g. oven, lights). In markets, the calculation of the baseline load of a building is often based on the average consumption in the previous days (Sharifi et al., 2016). Overall, there are three main different methods to estimate the baseline load:

1. **White or grey box simulation**
2. **Load pattern matching method**
3. **Regression analysis**

With the first method of using simulation, it is trivial to compare different control strategies, since one can simply define one of them as the baseline (Amblard et al., 2020; Fischer, Wolf, Wapler, et al., 2017). In building simulation, advanced simulation tools using white or grey box

models can be use to compute the baseline. These models, however, are complex to calibrate and take time to run, making them difficult to be applied in real time control.

The second method relies on load pattern clustering techniques to estimate the baseline. The clustering can be temporal (Day Matching Method) or spatial (class of buildings). The assumption is that the baseline can be deduced from days with similar conditions (e.g weather, day-of-week) or from the consumption of buildings presenting similar patterns also referred to as Control group method. Li et al. (2017) estimates the baseline load during DR event periods for DR participants based on the actual load of non-participants of the same cluster.

The third method relies on regression models to estimate the baseline. The parameters of the models can be estimated on both DR and non-DR days. Mathieu et al. (2011) developed a regression-based electricity baseline model with exogenous variables for commercial and industrial buildings. They used a time-of-week variable and a piece-wise linear and continuous outdoor air temperature dependence. The drawbacks of these methods in general are (a) that they require data of several preceding days to estimate the model coefficients and (b) data from DR-days falsify the inference of the pure baseline.

In practice, it is almost impossible to accurately define a baseline for single or small clusters of residential buildings at quarter hourly based timestamps. Patterns can be identified for weekly or daily timestamps, but at shorter time intervals the uncertainty is more significant, especially for DHW consumption. Wang et al. (2018) develop a data driven model to simulate a generic cluster of buildings subjected to disturbances and uncertainty. They show that the uncertainty of the energy flexibility decreases when the aggregated number of buildings increases.

New methods that do not rely on an estimation of the baseline have gained focus in recent years. Junker et al. (2018) propose a method to model the energy flexibility of buildings as a dynamic function, describing the relation between a penalty signal and the response of a system. The "Flexibility Function", as they call it, could be used to characterize the energy flexibility or directly to provide ancillary services. The main advantage of this method is that it only requires one-way communication compared to transactive energy approaches, where states and costs are iteratively negotiated in a two-way communication. However, this method is still limited to time-invariant and linear systems. Furthermore, they assume that the system providing flexibility is smart and able to respond to an external penalty signal, which is not true yet in most residential buildings.

In this project we have set-up a two way communication system allowing us to test a transactive energy approach. At each transaction the cluster manager estimates the controllable baseline load of the building (e.g. heat pump compressor and circulation pumps). This baseline estimation helps the commercial aggregator quantify the flexibility potential and to compute the retribution. The baseline is computed following two different approaches:

- Solving a scheduling optimisation problem
- Using a seasonal Auto-Regressive model (AR)

With the first approach, we estimate the baseline by solving a scheduling optimisation problem with the same models and constraints used in the MPC formulation (Chapter 2). The objective is to minimise the energy cost while ensuring comfort. The baseline is then the sum of the HP electric power loads resulting from the optimisation. We assume that the MPC controller for energy minimisation would have selected and applied this control strategy in the absence of a DR-event. In the case where BaU is run instead of energy minimisation in between DR-event, this method can still be used by tightening the constraints on the setpoints variations.

With the second approach, we compute the baseline by training AR models for each building using historical synthetic power data, typically from the previous 3 days, a daily seasonality (720 samples) and a lag of one hour (30 samples). We here use the term synthetic because the power is not directly measured (see Section 4.3) but reconstructed from operating temperature time series and manufacturer datasheets. Statistical models do not have a physical interpretation and can sometimes generate non-physical behaviour (e.g. negative power). This is even more true when the system we try to model is steep because of ON-OFF patterns. The time discretisation, here with 2-minute time steps, can also cause problems. The aggregated baseline for the site is computed by summing the estimated baseline of each building following Algorithm 3:

Algorithm 3 Baseline estimation using seasonal AutoRegressive Model

Update the historical values (2160 samples)
for n in len(Buildings) **do**
  Train AR model and fit its parameters
  Predict multi-steps out of sample values for the next X hours: $\hat{P}_n$
  for t in len($\hat{P}_n$) **do**
    if $\hat{P}_{n,t}$ < threshold **then**
      $\hat{P}_{n,t} \leftarrow 0$ {Corresponds to HP being OFF at time t}
    else
      $\hat{P}_{n,t} \leftarrow 1$ {Corresponds to HP being ON at time t}
  Resample to 15-minutes values
  $\hat{P}_{n,t} \leftarrow$ Power using HP manufacturer datasheet
$\hat{P}_{cluster,t} \leftarrow \sum_{n=1}^{N} \hat{P}_{n,t}$

The seasonal AR(p=3) with a seasonal period s is formulated as:

$$Y_t = c + \Phi_1 Y_{t-s} + \Phi_2 Y_{t-2s} + ... + \Phi_p Y_{t-ps} + \epsilon_t \tag{5.1}$$

Where $\Phi_1, \ldots, \Phi_P$ are the parameters of the model, c is a constant, and ϵ_t is white noise.

Once the parameters of Equation 5.1 have been estimated, the AR model can be used to forecast an arbitrary number of periods, but less than the model length (i.e. less than 6 hours), into the future. For the first time step, where the values are not available, we substitute the known preceding values X_{t-i} for i=1, ..., p into the AR equation. The output corresponds to the forecast for the first unobserved period. This time, We iterate the process using the predicted

value arising from the previous forecasting step.

The main advantage of the method is that it is fast and no *a priori* knowledge of the physical system is needed. Parameters for the method, such as the length for the seasonality period or the lag are fitted off-line, based on historical values. The drawback is that the forecast time series are heavily influenced by recent DR services. To mitigate this problem, exogenous variables such as setpoint values can be used. However, in this case it does not significantly improve the quality of the results. We use a threshold value to filter the sequence of estimated power and transform it into a sequence of binary variables (the value at time t is 0 if $\hat{P}$ < threshold, and 1 otherwise). We can then convert it back into power using. The quality of prediction depends on the function used to convert binary to power. Here we assume a constant COP. However, with regards to the results of the test presented in Chapter 6, this constant COP assumption has proven to be a poor choice. A combination of piece-wise linear COP models based on storage temperatures would have reduced the error.

5.3.3 Generating traces

During the negotiation phase, several power traces are exchanged between the cluster manager and the aggregator with a duration of x hours (x = 1h for most of the tests) and a frequency of 15 minutes. The frequency of 15 minutes has been chosen as it satisfies electricity market specifications and technical limitations for water-water HP. Traces are generated following two methodologies:

1. Solving scheduling optimisation problems (p: prediction horizon of 6 hours)
2. Using a seasonal auto-regression model based on data from the previous days

The first option relies on a scheduling optimisation problem, as presented in Chapter 2. It has the advantage of being close to the one used for the tracking MPC and thus traces generated that way have a higher chance of being followed than with the second, more distant, method. The drawback is that each trace generation takes up to 80 seconds, as we need to solve an optimisation problem. For the first set of tests, Min and Max consumption traces were generated by penalising the deviation between the aggregated power consumption and the theoretical maximum or minimum aggregated consumption of the site, respectively. For the last set of tests, the aggregator added a target to the Flex Request (i.e. [100, 100, 0, 0]) at the beginning of the negotiation phase. This signal is converted into power and used as a tracking reference for the scheduling MPC. The portfolio optimisation performed by the aggregator is out of the scope of this thesis. Such an optimisation can be found for example in (Gkatzikis et al., 2013). For the sake of simplicity, the aggregator here chooses to generate only two default target profiles.

With the second option, the maximum and minimum consumption traces are derived from the baseline by identifying the periods where all the HPs can be activated or where multiple HPs

can be switched OFF, respectively. As the cluster manager has no market price information, it only tries to find periods with maximal or minimal consumption potential. The simple idea behind these periods is that a HP activation can be shifted forward and backward only for a rather short duration (a few hours). For example, if two HPs are forecast to start at the beginning of the DR period and three at the end, their activation can be shifted in order to have them start all together at a certain time in the middle of the period.

5.3.4 Cost of the traces

The MPC objective function in equation 5.2, is a representation of the operational costs including the electricity price, the discomfort and the penalties for changing the behaviour from the baseline. The optimisation problem is formulated as:

$$\min \sum_{n=1}^{N} \sum_{t=t_0}^{h_p} \left[\lambda^{el} P_{n,t} + \lambda^{tc} \Delta T_{n,t}^{tc} + \lambda^{act} \Delta u_{n,t} \right] + \sum_{k=t_0}^{t_0+d_{DR}} \lambda^{tr} \left(\sum_{n=1}^{N} P_{n,k} - R_k \right)^2 \tag{5.2}$$

The λ are weighting parameters that drive the solution towards different objectives, thus allowing to choose between conservative or more aggressive control strategies. In this formulation, the electricity price λ^{el} is considered constant ($0.18 cts/kWh$) over the optimisation period. The associated penalty cost for deviation, $\lambda^{tr} >> \lambda^{el}$, drives the objective function towards minimising the power deviation from the minimum, the maximum or the target profiles, denoted here by the variable R. The cost function also includes the level of discomfort $\Delta T_{n,t}^{tc}$, described in equation (2.29b-2.29d), with an associated penalty cost. It ensures that the state variables describing the air and mass temperature of the building as well as the DHW storage and buffer tanks stay within the comfort limits given by $T_{n,t}^{min}$ and $T_{n,t}^{max}$. Finally, $\Delta u_{n,t}$ quantifies the change of the manipulated variables (setpoint variables). In order to penalize frequent control actuation that could alter the system performances, a cost λ^{act} is associated with $\Delta u_{n,t}$. The higher the cost, the less actuations are permitted.

As the cluster manager has no market prices in this case, the associated cost of a trace is computed post-optimisation against the baseline cost $Cost_0$ as follows:

$$Cost_i = \left(C_i^{comf} + C_i^{act} \right) - \left(C_0^{comf} + C_0^{act} \right) + \lambda^{el} \max \left(0, \sum_{k=0}^{d_{DR}} P_{i,k}^{clust} - \sum_{k=0}^{d_{DR}} P_{0,k}^{clust} \right) \tag{5.3}$$

C^{comf} and C^{act} are the penalties for the comfort deviation and actuation of the heat pump. They include all the penalties over the prediction horizon to account for long term disturbances. The index 0 represents the baseline. The right side of equation (5.3) accounts for the energy deviation cost. It takes into account only the deviation of the DR service duration d_{DR} and can only be positive. P^{clust} is the aggregated power of all buildings. In-between DR services, when the system is run as BaU and the baseline is estimated using an auto-regressive model, $C_0^{comf} + C_0^{act}$ is set equal to zero. We assume there is no discomfort.

Following the baseline is assumed to have a cost of zero, meaning that the CM will not be remunerated if he follows it. Due to model uncertainty and forecasting errors (i.e. weather and DHW consumption), it is very difficult in practice to predict and follow a 15-minute timestamp baseline over several timestamps. Therefore, deviation from the baseline should not be heavily penalized, otherwise the risk taken is only borne by the cluster manager and customers.

5.4 Centralised control

5.4.1 Operation in-between DR services

Business As Usual

The pilot site is left running with default settings. If the setpoints of the heat pumps are not at default values, we automatically set them back. The software prototype continues to collect the last values from the site and to update the building's models on a 15-minute basis. Weather predictions are updated every hour. We run alarm checks and send warning emails in case of identified problems. To evaluate the effect of DR calls, the systems are set back to BaU after each tracking period.

Local Energy minimisation

When this control option is chosen, the pilot site is controlled by the centralised MPC. The objective of the MPC is to minimise the energy cost while ensuring thermal comfort for all buildings. The right-side part of equation 5.2 is removed. If a building completely loses its connection to the monitoring platform (i.e. did not send new values during the past hour), the building is temporarily removed from the test and the framework is automatically updated to ensure data integrity. If the setpoint values are not at default and the connection cannot be reset, a technician is sent on-site to reset the control values and restart the communication system. For long-run tests including the night period, we decide not to run the energy minimisation MPC in between DR calls for mainly two reasons:

1. To better assess the effect of the DR services on the system
2. To avoid comfort violations

The second reason is the most critical in poorly monitored buildings where assessing comfort is complex and error-prone. Model predictive control for local energy minimisation will drive the system towards lower boundary conditions (low thermal limits). Running BaU instead of a centralised MPC reduces the risk of having the tests stopped prematurely due to comfort violation. We assume that the actual control system is sufficiently well designed and does not operate the system close to a boundary condition.

5.4.2 Tracking power profiles with Model Predictive Control

Optimisation

Once an activation is received, a trace is selected and the cluster manager will track the agreed trace starting from the next time step, exactly 15 minutes after receiving the Flex Request, as presented in Figure 5.6. The prototype DR service software calls a centralised tracking MPC every 15 minutes. To account for the solving time of the optimisation problem and the ramp-up time of the heat pumps, we call the tracking MPC two minutes before the new actuation time. We find this to be a good compromise between the time issues previously mentioned and the accuracy of the last state. Updating the state too much in advance can lead to a faulty estimation of some variables or constraints. This may cause the solution found by the optimisation to not be feasible in reality.

The tracking MPC is formulated with a prediction horizon of 6 hours and a control horizon of 15 minutes. The length of 6 hours is a good compromise between forecasting horizon and solving time. With a shorter horizon the optimal solution is computed faster but we lose the advantage to foresee constraints violations or shifting opportunities. With a longer horizon the solver struggles to find an optimal solution within the time constraints and the forecasting errors increase due to the simplified models used or error in weather predictions.

The formulation is composed of a model of the building, the heat pump, the heat circuit and the two storage tanks (SH and DHW) for each building. Deviations from the trace agreed on with the aggregator are penalised using the objective function presented in Equation 5.2. After each optimisation, we remove the first value of the reference sequence R, and increase the associated penalty λ^{tr}. When the length of R is much shorter than the prediction horizon h_p, we need to change the penalty because otherwise the solver may choose to ignore the tracking error. A set of constraints imitates the effect of the local dead band controller. This constrains the HP to start only under certain conditions that can be altered by changing the setpoint control variable. We tested different methods to avoid too many changes in the control variables and found the method of (Cigler, Siroky, et al., 2013), which penalizes quadratic deviations from the setpoints, to work the best.

Due to time constraints, each tracking MPC run is limited to 80 seconds. Either an optimal solution is found within this time limit or the best sub-optimal solution is kept and implemented.

Post-compute

The building systems for SH and DHW on the pilot site are controlled through the HP controller. Three identified control strategies were tested on some heat pumps, as presented in Section 3.4.2. The possible control actions are listed below

1. Switch **ON/OFF** the heat pump by increasing/decreasing the “Room temperature setpoint”. This setpoint is used by the local controller via a heat curve to compute the

departure temperature of the heating circuit and decide whether the HP compressor needs to start or not. When night setback is implemented on the local controller, the "Reduced room temperature setpoint" needs to be updated during night-time instead of the "Room temperature setpoint"

2. Switch **ON/OFF** the heat pump by increasing/decreasing the DHW setpoint

3. Switch **OFF** the heat pump by changing the "operating mode" from "heating/DHW" (= 2) to "Off" (= 0)

Although a heat pump can forcefully be shut down by changing its operating mode from Heat/DHW to Off (strategy 3), this option is not used on the pilot site. The main reason is that the communication framework installed onsite (i.e. the gateways) has no "reset" capability or system failure protocol. If the communication is interrupted between the monitoring platform and the gateway, or if the latter encounters failure, neither the heat pump or the gateway can be remotely restarted. Therefore, they remain switched OFF, requiring the onsite intervention of a technician. As this problem occurs regularly enough and no solution is provided by the pilot site manager (ELIMES), this control strategy is discarded for the long duration tests.

Once the optimisation finds an optimal solution or reaches the maximum solving time, the manipulated variables corresponding to different control actions are post-processed to send the setpoints to the HP. Two options are tested:

1. **MMSP**: In 'MinMax SetPoint', we increase/decrease setpoints by a fixed ΔT whenever the HP state resulting from the optimisation is ON/OFF". This method is similar to the heat pump solution proposed by SMART GRID READY option (Fischer, Wolf, & Triebel, 2017).

2. **VARSP**: In 'VARiable SetPoint', we increase/decrease setpoints based on the new values of the heat curve optimised by the MPC

The advantage of the first option compared to the second is that it requires a less detailed formulation of the MPC, as only the state of the HP is required (control variables). The ΔT in the setpoints is often higher, causing a larger shift in the dead-band constraints. Indeed, in the second option the optimization penalizes high deviations from the default setpoint. This can be insufficient to drive HPs to start or stop on time. For this reason, a small ΔT is added/deduced from the optimization results.

The biggest drawback encountered by the first method is that it only has two operation states, either ON or OFF. This has proven to be a poor formulation in the presence of a heat buffer. The second option can provide more "operation flexibility". By slightly increasing the temperature setpoints, it can shorten HP OFF time without starting the HPs. This is useful when the buffer tank should be emptied in prevision of a future activation of the HP.

5.5 Description of the tests

5.5.1 Goal of the tests

The main test objectives are to specify and study in a pilot site the technical, social and economic characteristics that will enable dynamic DR services to be offered by blocks of buildings. The tests should also help describe the reliability and the performance of MPC methodology to provide DR services. The focus is put on power-to-heat applications and the optimisation of available energy vectors in buildings. One of the main sources of flexibility in blocks of buildings are heat pumps. Hence the goal is to confirm the potential and challenges of flexible control of heat pumps by performing multiple tests as a flexibility service provider. Actions toward achieving these objectives include:

- Analyse system performance, develop optimisation procedures and improve system use;
- Evaluate the impact of the developed DR solution on real systems with respect to energy, costs and user acceptance;
- Design a DR evaluation methodology in blocks of buildings;
- Quantify the reliability of bundled flexibility of smaller buildings.

We want to test a transactive DR approach (two way communication system) and assess its reliability and performance over multiple consecutive days with multiple DR-events per day. Due to delays in the project, the seasons fall and winter could not be thoroughly tested.

5.5.2 Main milestones of the Tests

Below are the dates detailing the main implementation steps of the DR service tests:

- **Winter 2018:** Test of the Kafka interface between the aggregator & the cluster manager on a simulated version of the pilot site
- **January 2019:** Change of the communication and management platform
- **February to Mai 2019:** Test of the novel platform communication protocols
- **August-September 2019:** Installation of indoor sensors in five buildings
- **Winter 2019-2020:** Standalone tests to tune the framework
- **April 2020:** Daily experiments of DR services on five buildings with aggregator & cluster manager interaction (4 weeks from 8:00 to 20:00 except weekends)
- **Mai 2020:** Weekly experiments of DR services on five buildings with the aggregator & the cluster manager interaction (2 weeks day and night except weekends)

Table 5.1 – Main parameters of DR tests on the pilot site

	Day Test	Week Test
Start time	08:00	04:00 Monday
Stop time	20:00	20:00 Friday
Last negotiation	18:00 (Friday 16:00)	Friday 16:00
Negotiation duration	10 - 12 min	12 min
Calls duration	1-2 hours	1 hour
Time between calls	2 hours	2 hours
Max call per day	6	11
Period of Test	April 2020	1-15 May 2020

5.5.3 Summary of the Tests

The data from the pilot site are collected and processed on a server at HES-SO by the cluster manager (HES-SO in this testcase) at a 15-minute frequency rate. The framework can be divided into three steps:

1. Waiting to provide DR services by running **Business as Usual (BaU)**
2. **Negotiation**: The cluster manager and the aggregator exchange information on the asset flexibility. Several traces are generated (e.g. Baseline, High consumption, Low consumption) with a duration of x hours (x = 1h for most of the tests) by:
   (a) Solving a **scheduling optimization problem** (p: prediction horizon of 6 hours)
   (b) Using a **seasonal auto-regression model** based on the past 2 days of data
3. **Tracking of the agreed trace with MPC** (p: 6h, u: control horizon of 15min) and using one of the two control strategies:
   (a) **MMSP**: In 'MinMax SetPoint', we increase/decrease setpoints by a fixed ΔT whenever we want to Start/Stop a HP
   (b) **VARSP**: In 'VARiable SetPoint', we increase/decrease set points based on the new values of the heat curve optimized by the Model Predictive Control (MPC)"

The tests are scheduled in two phases: Day tests and Week tests. The main parameters of the tests are given in 5.1:

6 Performance and Reliability

In the previous chapter we presented the framework of the tests and described the actors and methods to perform power load control with heat pumps in a real residential building block. Based on these tests we here assess the reliability and performance of DR transactions between the pilot site and an aggregator. The chapter is structured as follows:

Section 6.1 presents selected indicators to quantify the performances and reliability of the flexibility offered by heat pumps in residential buildings providing DR services. We focus on indicators that can be assessed in real systems. In section 6.2 we evaluate the reliability of the test by assessing how consistently the framework is able to answer DR service calls from the aggregator. In section 6.3, we assess the performance of the heat pumps in terms of flexible power and energy. We evaluate the aggregated results as well as the detailed results per building. Finally, section 6.4 presents the results of a collaboration in which we model the flexibility as a dynamic function.

6.1 Indicators

The tests are performed on real systems with varying initial conditions (e.g. state of charge of the storage) and disturbances (e.g. weather). As a result, it is complex to compare the performances of the different control methodologies implemented. To help assess the performances of these methodologies, we select a series of indicators. This approach will not necessarily lead to a like-for-like comparison, as there are still too many external factors that affect the results. However, the indicators are expected to provide a strong base to make objective assessments.

6.1.1 Optimisation Indicators

Optimisation indicators provide metrics for comparing the different optimisation methodologies, by quantifying the methodologies' ability to solve the problems and to participate within the constraints imposed by the electricity markets.

Resolution time

The time taken to solve each iteration of optimisation is of vital importance, especially for certain markets. The higher the optimisation time, the lower the attractiveness of the solution proposed. A methodology achieving optimal solutions but requiring computing time longer than the activation period can not be implemented in "real time" control. Hence, the value of the indicator is bound between 0 and the maximum tolerated time before the termination of the optimisation (e.g. 70 seconds).

Optimality gap

This indicator is applicable mostly to algorithms that rely on solvers like Gurobi and CPLEX, which use linear programming or variants of linear programming. The optimality gap is the difference between the best known feasible solution and a value that bounds the best possible solution. The global optimality is achieved if this difference is equal to 0. The higher the gap, the lower the quality of the solution and the lower the attractiveness of the framework proposed. However, for complex problems such as Mixed-Integer Linear Programming (MILP), a solution with a lower gap compared to a previous solution with a higher one can have a similar value of the best feasible solution. In fact, the gap is not only reduced by finding a better feasible solution, but also when the solver changes the best bound (e.g. by cutting the optimisation space). In general, this indicator may not be the best metric, as it depends on the solver used and on the constraints imposed on the problem. Moreover, in alternate frameworks, using heuristic and other solving methods, this indicator might not be applicable.

Infeasibility count

For optimisations using linear/mixed integer programming approaches, one can count the number of times an optimisation arrives at an infeasibility. This is a good metric to quantify the ability of providing any solution for the next given duration of time. If the infeasibility count is too high, the method will not allow a consistent participation in the market as penalties will be higher than the benefit.

Number of actuations

This metric counts the number of heat pump actuations proposed by the optimisation solution. An actuation is defined as the action of control over a machine or a process. The number of actuations can be compared with the actual number of switch ON and switch OFF on a similar day without any optimisation. If the optimisation proposes too many quick changes in the functioning of devices, especially for devices that use compressors, it will be counter-productive as that would result in a reduction of the lifetime of the device and an increase in maintenance costs. So, for a good optimisation, the number of actuations proposed should not be too high.

6.1.2 System Indicators

System indicators provide metrics for assessing how reliably local control implements the solutions found by the optimisation. There will always be some divergence between the theoretical number of actuations proposed and the actual ability of the devices in the site to implement the proposals. The lower this difference is, the more reliable the system is.

Actuation success (*AS*)

This metric reflects a real device's acceptance of the change in behaviour proposed by the optimisation. A good optimisation is able to successfully propose implementable changes in system behaviour. A "non-change" in a heat pump state (e.g. the heat pump stays ON), if planned by the optimisation, is in this case also considered as a success. *AS* can be computed per building n over a call of duration d_{DR} (e.g. the number of time steps t) using equation 6.1.

$$\forall n, t: \quad \delta_{n,t} = \begin{cases} 1, & \text{if } T^{set}_{n,t} \neq T^{set}_{n,t-1} \\ 1, & \text{elif } T^{set}_{n,t} \neq T^{set}_{n,0} \\ 0, & \text{otherwise} \end{cases} \tag{6.1a}$$

$$\forall n, t: \quad \Delta_{n,t} = \left(1 - \mid \hat{X}_{n,t} - \overline{X}_{n,t} \mid\right)\delta_{n,t} \tag{6.1b}$$

$$\forall n \quad AS_n = \frac{\sum_{t=t_0}^{d_{DR}} \Delta_{n,t}}{d_{DR}} \tag{6.1c}$$

Here T^{set}_n are the respective temperature setpoints for DHW or SH, $T^{set}_{n,0}$ is the setpoint value previous to the DR event. $\hat{X}_t$ is the estimated state of the heat pump resulting from the tracking MPC at time t and $\overline{X}$ is the average state estimated using algorithm 2 (see chapter 4) over time step t. If there are multiple (k) DR events per day, the daily AS can be computed as the weighted average over all $AS_{n,k}$ with the weight given by the number of time steps (d^{DR}_k).

Percentage of failed actuations

This metric reports the percentage of failures of the DR call after a trace has been agreed upon between the cluster manager and the aggregator (see section 5.2.2). This includes all the times the DR service framework fails due to optimisation or communication issues.

6.1.3 Comfort Indicators

One of the difficulties in DR for residential and commercial users is that there is a trade-off between user comfort and financial gain. The question is often, how much discomfort is the user willing to accept for how much financial gain. User comfort, or the lack thereof, can be quantified by comfort indicators. One way to quantify thermal comfort, which is often used in simulation studies, is to monitor the interior temperature, which needs to be within certain comfort limits. The longer these limits are exceeded and the higher the exceeding margin, the lower the user comfort. In real residential buildings, however, the indoor sensors

are often nonexistent or located in non-representative places, making this method difficult to implement (Section 3.2.1). Thus, we rely on the next indicator for comfort assessment.

Number of phone calls/complaints

Phone calls expressing dissatisfaction and the need for changing the strategy imposed by the optimisation are considered for this metric. In that sense, this is a placeholder metric. It is a placeholder quantifying the severity of the discomfort/displeasure of the tenants. The more phone calls are received from the tenants in each of the buildings, the poorer the DR service is. Since in residential buildings we do not have access to representative indoor temperatures, it is impossible to know the tenants' comfort without their direct feedback.

DHW temperature

More than miss-regulated indoor temperatures (i.e. outside the comfort boundaries), miss-regulated Domestic Hot Water (DHW) is perceived as a great discomfort by building residents. The DHW temperature can either be too cold (e.g. < 30°C) or too hot (e.g. > 70°C). This indicator evaluates the number of times per day at which the DHW storage temperature is outside the comfort limits. The time spent outside the limits is a better assessment of discomfort than the number of times the limits are exceeded. The cold water circulating when the DHW buffer is charged will often make the monitored temperature drop below the limits for a short period. However, this is normal and does not cause much discomfort.

Monetary gain for customers

Another important metric for resident satisfaction is the final monetary remuneration that they can gain. The higher the monetary remuneration for the proportional discomfort (expressed through the number of phone calls/complaints), the better the framework will be considered. This indicator could not be tested during this project.

6.1.4 Energy/Power Indicators

In this definition of indicators we want to address both the performances of the method (e.g. tracking MPC), as well as the flexibility available in residential buildings via load control.

Absolute tracking error between the trace agreed upon and reality

This metric quantifies the absolute energy deviation over a period. It is helpful to assess if heat pumps fail to provide positive or negative flexibility. The absolute tracking error is given by:

$$\varepsilon^{abs,dev} = \frac{ts}{3600} \sum_{t=t_0}^{t_0+d_{DR}} \left(P_t^e - P_t^a\right) \tag{6.2}$$

Here, ts is the time (in seconds) of the window period (e.g. 15 minutes) of each actuation. If the value is positive it means that the realised power P^e exceeds the tracked power load P^a.

Cumulative tracking error between the trace agreed upon and reality

This metric corresponds to the difference between the real performance of the site and the trace agreed upon for the site by the cluster manager and the aggregator. It indirectly quantifies the performance of the scheduling optimisation. The lower the deviation between the agreed upon trace and the actual performance, the better the optimisation framework. The cumulative tracking error is given by:

$$\varepsilon^{dev} = \frac{ts}{3.6e^3} \sum_{t=t_0}^{t_0+d_{DR}} \mid P_t^e - P_t^a \mid \tag{6.3}$$

6.2 Reliability

In this section, we will discuss the reliability of the automated DR service test performed during six weeks from April 2020 to mid May 2020. We define the reliability here as the consistency with which the DR service calls are answered. Reliability is negatively affected by software issues, optimisation problems (e.g. infeasibility) and communication failure. The quality of the answer to the calls will be discussed in the following section about performance.

Figure 6.1 represents the reliability of the framework to offer DR service during the test period. Each bar represents a day of test. The height of the bar compared to the small line marker represents the number of successful negotiations compared to the number of calls from the aggregator received for the day. The color map displays the rate of failed actuations in percentage over the day.

From the 6th April 2020 to the 15th May 2020, 186 out of a total of 216 DR negotiations were successfully conducted (86%). The main cause for the 14% of unanswered calls was software failure caused by infeasibility for the scheduling optimisation problems. After each failure or at the end of the day, we updated the framework by relaxing hard constraints to fix the problems, and we added safeguards in the script to avoid termination. As displayed in Figure 6.1, the reliability was improved every week. During the last two weeks, “weekly tests” were conducted all day long including night periods but without the weekend. This corresponds to 42% of the total DR negotiation calls, of which 95% were successfully answered.

Out of the successful negotiations, the percentage of failed actuations during the DR service duration for the whole test period was 13%. It was again mostly caused by infeasibility of the load tracking optimisation problem. This percentage dropped to 5% for the weekly tests. Most of the infeasibility problems occurred for the DR call at 16:00. This corresponds to a time period of the day where the moving average of the outside temperature monitored by the heat pump is the highest. During this period, values for the storage tank or for the heat are close to hard constraint values, which can result in infeasible problems.

The upper panel of Figure 6.2 represents the daily actuation success per building for the last 3 weeks of tests as computed by equation 6.1. The buildings in the heat map are ranked by

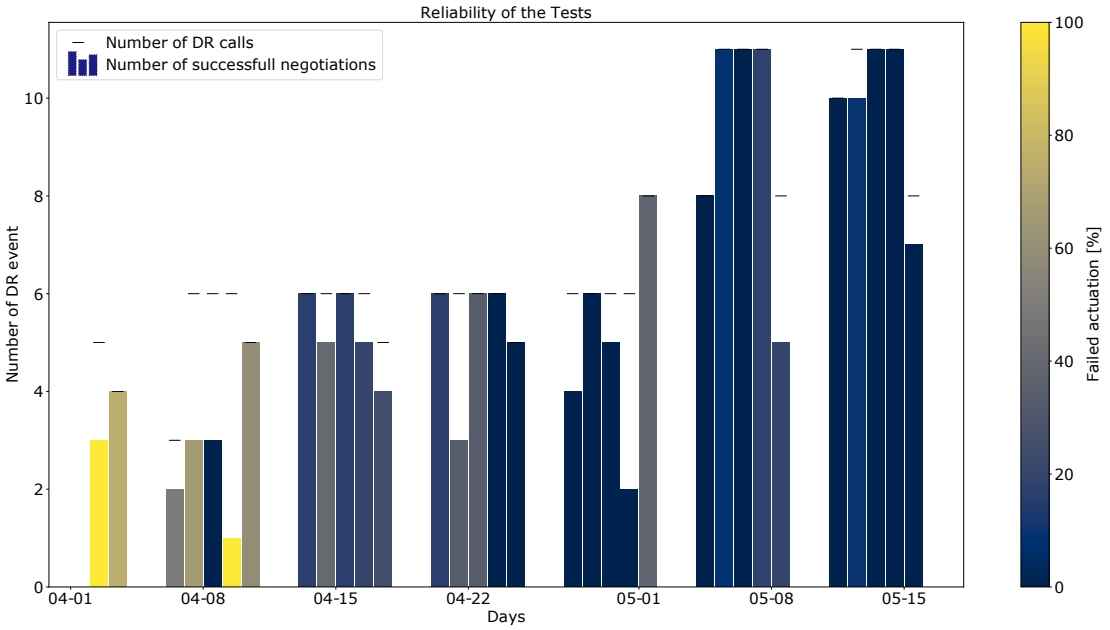


Figure 6.1 – Reliability of the DR service tests. The reliability is represented here by the successful negotiations (vertical bar) compared to the number of DR calls (black dash) and by the actuation success (colormap).

their AS score, from Lenzburg with a score of 94.6% to Jegenstorf with a score of 87%. Overall the cluster of 5 buildings has a score of 91.1%. As presented in Figure 6.1 the last two weeks of tests have a more successful negotiation. Here, one actuation failure will have less influence on the score than it does during the first week, in which there were fewer negotiations.

Jegenstorf is a renovated building with floor heating and therefore has very low SH departure temperature even in winter. At the current outside temperature level, most of the actuation flexibility comes from the DHW. During the second week, the optimisation tries to intensively run the heat pump to produce DHW. This failed most of the time in reality, which can explain the relatively low score of 75%. After changing some constraints for DHW in the scheduling and tracking MPC, the optimisation reduces its DHW actuation and the score increases.

The lower panel presents the weather conditions (e.g. outside temperature and solar radiation). Outside temperatures are monitored by the heat pumps and averaged over a period of 3 hours. It corresponds to the value used by the local controller. We see that for the period of test, monitored values range from a minimum of 10°C at night to more than 25°C during the day. The building Oberhofen, whose sensor is not in the shade, measures during 3 hours an averaged outside temperature of above 35°C. Overall, it does not seem that outside temperature and actuation success are strongly correlated, but due to the relatively small set of results, a correlation cannot be excluded solely based on those results. Other factors influence the results, such as the trace chosen by the aggregator. For these tests the aggregator picks a trace randomly (i.e. 35.3% baseline, 34.5% positive, 30.2% negative).

Figure 6.2 – Heat map of actuation success per building and per day compared to the weather conditions (lower panel)

Figure 6.3 – Heat map of actuation success per hour and day of the week.

In Figure 6.3, the actuation success for the six weeks of tests are clustered per day of the week and per hour of the day. The scores are strongly influenced by the number of tests for each hour and day. For example, the hours from 20:00 to 6:00 were only tested during 2 weeks, which can explain the low scores (i.e. 50%). The low score for the days at the beginning of the week can be explained by the fact that we implemented corrections during the weekend and we tested them at the beginning of the week. It is difficult to see a clear hourly pattern from Figure 6.3, but the actuation success is a bit worse in the afternoon for the DR calls at 16:00 and 18:00. This can be explain by the fact that they correspond to the highest 3-hours averaged external temperatures. As seen in section 4.2, the forecast are also less accurate for these hours of the day. These two factors in combination probably explain why changes in setpoint temperature do not result in heat pumps starting up. To test this hypothesis and confirm the effect of the external temperature on the actuation success, one should perform a similar test on colder days. Unfortunately, we did not have time to test this.

6.3 Performances

Table 6.1 presents a summary of the tests performed from the 2nd of April to the 15th of May 2020. We quantify their reliability with the number of successful activations and the rate of success of the heat pump actuation. The performance is quantified with the absolute and cumulative tracking error between the agreed upon trace and the reality (e.g. Abs. and Cum. Deviation in kWh). The mean power deviation allows us to compare the tracking performance. The duration of a DR event is not mentioned in this table as all tests have the same length

of 1 hour. As presented in Chapter 5, we test two ways of generating a trace (e.g. Scheduling optimisation or Autoregressive model) and two ways of translating control actions from the centralised optimisation to the local controller (e.g. Min/Max setpoints or variable stpoints).

We can see in Table 6.1 that results for the last weeks have a higher actuation success and lower power deviation than the tests at the beginning of the testing phase. This is mostly a consequence of the improvement we made on the framework along the tests. At the end of the testing phase the mean power deviation is less than 3 kW. It corresponds to half of the power of a heat pump out of the 5 tested. The absolute deviation is mostly negative for all the tests. This means that during the period tested, heat pumps had greater difficulty to follow a positive deviation (i.e. switch ON heat pumps) than a negative one.

The best results have been obtained with a combination of a scheduling optimisation for the trace generation and variable setpoint control. Drawing strong conclusions here is not reasonable due to the small set of tests performed had the limited period of the year tested. However, these tests can still provide us with some insight on the limitations of the methods and the system tested. In the following subsections we will look into the results more in detail by discussing them for a specific day (i.e. the 14^{th} of May 2020).

6.3.1 Aggregated Results

Figures 6.4-6.13 present the results of a day from a week-long test of direct load control services. To ease reading, vertical areas featuring the period of active testing are repeated in all figures. The light grey vertical areas display the 12 minute negotiation periods between the aggregator and the cluster manager. The light red vertical areas show when control actions are performed on site as solutions of the tracking MPC optimisation. They are flanked by light red vertical lines that mark the duration of a DR service. For the day presented, each Flex Request is answered with a Flex Offer and the MPC tracking is uninterrupted. In between DR calls, the system is set back to BaU (i.e. default setpoints).

Figure 6.4 displays the aggregated power (blue), at a frequency of 1 minute, of 5 participating heat pumps on the 14^{th} of May 2020. The daily average external temperature is 18°C with temperatures above 20°C from 12:00 to 20:00. Therefore, most heat pump consumption occurs during the early hours of the day when the external temperature is still cold.

The thick red lines are the power loads P^a agreed upon by the aggregator and the cluster manager. There are 11 of them lasting 1 hour each with a 15 minute sampling. Each one corresponds to a power trace proposed by the cluster manager and selected by the aggregator resulting from a 6-hour forecast scheduling optimisation problem.

The black dashed lines are the forecast baselines. At each negotiation, baseline forecasts are re-evaluated for the next 4 hours because the previous DR power service affected the system behavior. The displayed lines stretch from the end of each negotiation period until the start of the next one. They are not re-evaluated directly at the end of the DR period to highlight

Table 6.1 – Summary of the test characteristics and associated results. In the column reporting the actuation success, the value na. were not computed as their score was non-significant. The last column reports the mean power deviation in [kW].

days	day of week	type of test	trace gen.	control	Nb activ.	Actua. Succ.	Dev. [kWh] Abs.	Dev. [kWh] Cum.	Mean P dev
02.04.2020	Thu.	Daily	Sched.	MMSP	3/3	na.	-3.01	21.25	7.08
03.04.2020	Fri.		Optim.		4/4	na.	25.03	25.89	6.47
06.04.2020	**Mon.**				2/2	na.	11.61	12.82	6.41
07.04.2020	Tue.				3/5	na.	-2.78	19.32	6.44
08.04.2020	Wed.	Daily	Sched. Optim.	MMSP	3/5	na.	7.46	7.46	2.49
09.04.2020	Thu.				1/6	na.	2.49	2.49	2.49
10.04.2020	Fri.				5/6		-9.33	21.95	4.39
13.04.2020	**Mon.**		Sched.		6/6	na.	-7.69	37.8	6.30
14.04.2020	Tue.			MMSP	5/6	na.	-4.98	35.34	7.07
15.04.2020	Wed.	Daily	AR		6/6	na.	-21.72	35.2	5.87
16.04.2020	Thu.		model	MM/VARSP	5/5	na.	-20.83	39.92	7.98
17.04.2020	Fri.			VARSP	5/5	na.	0.64	19.49	3.90
20.04.2020	**Mon.**				6/6	na.	-8.02	28.04	4.67
21.04.2020	Tue.				3/3	na.	-20.6	28.2	9.40
22.04.2020	Wed.	Daily	AR model	VARSP	6/6	na.	-10.3	28.9	4.82
23.04.2020	Thu.				6/6	na.	-5.5	32.29	5.38
24.04.2020	Fri.				5/5	89%	-0.13	27.48	5.50
27.04.2020	**Mon.**				4/6	91%	-2.05	10.61	2.65
28.04.2020	Tue.		AR model		6/6	96%	9.85	22.15	3.69
29.04.2020	Wed.	Daily		VARSP	5/6	91%	-4.1	14.9	2.98
30.04.2020	Thu.		Sched.		2/6	nan	0.6	10.9	5.45
01.05.2020	Fri.		Optim.		8/8	91%	12.8	41.1	5.14
04.05.2020	**Mon.**				8/8	88%	-16.2	39.9	4.99
05.05.2020	Tue.				10/11	92%	-3.8	40.6	4.06
06.05.2020	Wed.	Weekly	Sched. Optim.	VARSP	11/11	91%	-10.3	37.3	3.39
07.05.2020	Thu.				10/11	92%	-8.06	36.7	3.67
08.05.2020	Fri.				5/8	85%	-23.4	29.5	5.90
11.05.2020	**Mon.**				10/11	90%	-5.3	32.8	3.28
12.05.2020	Tue.				9/11	88%	-7.6	32.9	3.66
13.05.2020	Wed.	Weekly	Sched. Optim.	VARSP	11/11	91%	-20.2	48.3	4.39
14.05.2020	Thu.				11/11	94%	-6.39	32.72	2.97
15.05.2020	Fri.				7/8	96%	4.23	19.36	2.77

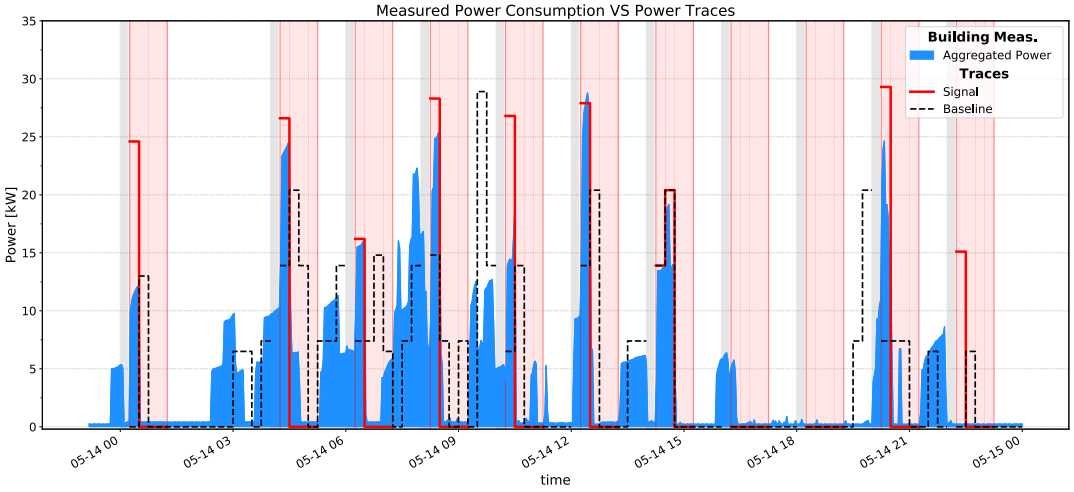


Figure 6.4 – Power consumption versus Power signal (traces) over a day of a weekly Test

possible rebound effects of the system. When the baseline P^b and the signal are identical, it means that the selected trace was the baseline.

To assess the performance of the direct load control services, the deviation between the power signal P^a (agreed upon trace) and the pilot site consumption P^e is presented in Figure 6.5.

The top panel of Figure 6.5, represents the power deviation from the agreed upon traces based on the aggregated data interpolated at a frequency of 1 minute. The selected traces are assumed to be constant over the sampling period of 15 minutes. When the values are negative it means that the on-site power was lower than the expected trace. When values are positive, it corresponds to an over-consumption in comparison to the agreed trace. The relative deviation over the day is -6.4 kWh and the cumulative, computed as the sum of all the absolute deviations, is equal to 32.7 kWh. This is as if on average, half of a heat pump was never correctly tracking the power signal. We see that for some activations, three heat pumps do not respond to direct control load.

When high power change occurs because of direct load control, high deviation spikes can be observed. The negative spikes correspond to an activation delay of the heat pumps. Even when conditions for the local controller are met, heat pump compressors are only started after a 2-minute delay by the local controller. To compensate, the tracking MPCs are launched 2 minutes before the new actuation periods. As soon as an optimal solution is found or at the latest 40 seconds before the new actuation time, the new setpoints are sent. Setpoints to switch OFF heat pumps are sent at the actuation time. Heat pump compressors directly stop when conditions are met, except when an explicit minimum running time is implemented by the local controller. The positive spikes observed are often a result of the monitoring sampling rate of 2 minutes and of the way power is measured/computed. The power consumption of 4 out of 5 heat pumps is not directly measured (not anymore) but reconstructed from

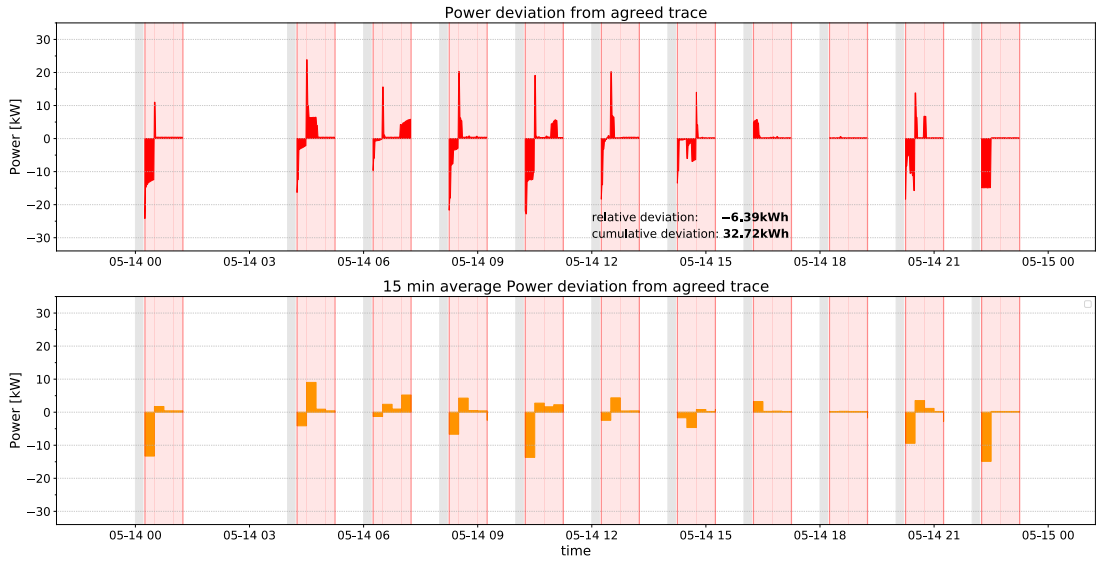


Figure 6.5 – Power deviation compared to the agreed upon traces resulting from the DR calls

operating temperature time series and manufacturer datasheets. The interpolation and the model formulation can sometimes create mismatches.

The lower panel displays the average power deviation based on 15-minute sampling. This is the value that interests electricity markets that are based on 15-minute bids. The amplitude of the deviation is in this case smaller than for the top panel.

The previous graph provides a good visualisation of the power deviation resulting from load control but less of the energy deviation (coloured areas). To better picture the flexibility in terms of energy, the cumulative energy consumption, as well as the cumulative energy of the baseline and the trace, are presented in Figure 6.6.

The blue line corresponds to the cumulative energy over the day measured on-site. The red lines are the cumulative energy if the tracking of the trace were perfect. The orange dashed lines correspond to the cumulative energy of the solution of the MPC. The black dashed lines represent the cumulative energy of the forecast baselines. The intercepts of the cumulative energy traces of all dashed lines are adjusted to match the blue line value at the beginning of each DR service period (after the negotiation). The cumulative energy pattern is only slightly affected by the DR call. This is because the first 15 minutes of most traces have a high variation in power compared to the baseline, as assessed by the small steps. Energy-wise the forecast of the baseline and the realised power are fairly similar, although we tend to slightly overestimate the baseline consumption. Finally, we can observe a plateau in the consumption during the night and late in the afternoon. This is typical for end of Spring and early summer days, where external temperatures rise above the temperature setpoints.

To better understand the performance and reliability of the site, we also must look at the

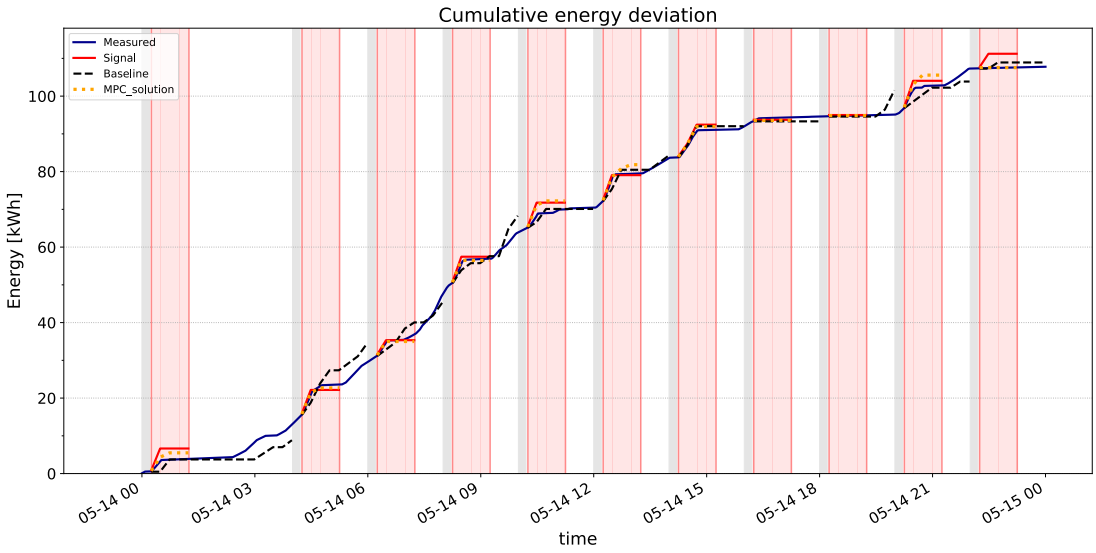


Figure 6.6 – Cumulative energy and its deviations compared to trace and baselines

results per building, as some buildings tend to perform better than others. In Figure 6.7, each building consumption is pictured with its own colour. A difference between the power signal and the measured power mainly results from either an error in the tracking optimisation or a non-implementation of the solution onsite.

The upper panel of Figure 6.7 displays the optimal power consumption found by solving the tracking MPC. The model used in the optimisation is a discrete model with 15-minute control steps. Therefore, consumption is constant across all 15-minute intervals. The values displayed here correspond to the first step optimal solution of each MPC solved, with in this case 4 rounds of tracking MPC solved per DR service duration (between light red vertical lines). At 22:15, for example, the optimal solution does not match the power signal agreed upon by the aggregator and the cluster manager. For nine of the DR calls, the error between the power signal and the MPC solution is less than 2 kW. This can be explained by different initial state values between the time when the trace was generated and when we tried to track it. As tracking MPC is only used during the DR call, the white vertical band periods are empty.

The second subplot is similar to Figure 6.4, as it represents the power measured compared to the power signal, but it details the effect of each heat pump individually. Each stacked coloured area corresponds to one heat pump. In contrast to the upper subplot, power is not only displayed for the direct load control periods but for the entire day. This allows us to see the different rebound effects as well as the initial power level before each call.

When compared with the upper panel, we can see several successful activations (at 04:15, 06:15, 08:15, 12:15 and 14:15) of the heat pumps corresponding to the ones predicted by the tracking MPC (e.g. at 06:15 when Montebello, Oberhofen and Lenzburg heat pumps are activated). But sometimes heat pumps do not react, e.g. Oberhofen at 00:15. To better understand those errors

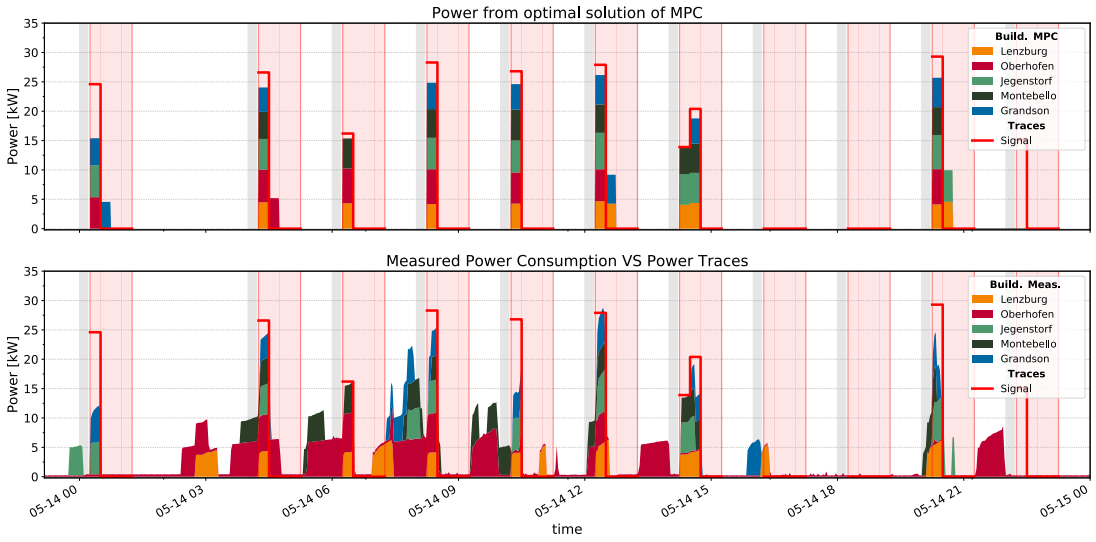


Figure 6.7 – Power consumption resulting from direct load control (MPC solution versus Measurement)

and activation failures, we need to look into the states of each building individually, especially the buffer and setpoint temperatures, as will be presented in the following subsection.

6.3.2 Results per building

In this subsection each figure represents the results of the test detailed at the building level. The colours correspond to the colour of the building as displayed in Figure 6.7.

Figures 6.8-6.12 are split into three panels. The top one shows both the temperature setpoints used for controlling the heat pump and the power measurements. The dotted lines correspond to the setpoint values of “Room temperature setpoint” and “DHW setpoints”, respectively. Outside the direct load control periods, the values of those setpoints are set back to their default values. The solid coloured line displays the measured power consumed by the compressor of the heat pump. We can clearly see the difference between when heat pumps operate to produce SH (lower value) or DHW (higher value and shorter duration). The solid coloured bars are the power consumption given as the solution of the MPC.

The middle panel represents the effect of direct load control on SH. The solid line corresponds to the only temperature measurement of the buffer tank. The dashed line corresponds to the measured departure temperature of the heating circuit after the 3-way valve. The dashed line represents the theoretical departure temperature of the circuit as given by the heat curve of the heat pump. Here, we use a linearised representation of the heat curve, since the value of the local controller is not available to us. The departure temperature $T^{flow} = f(T^{in,set}, \overline{T}^{ext})$ is a function of the “Room temperature setpoint” displayed in the top panel and of the external temperature averaged over 3 hours.

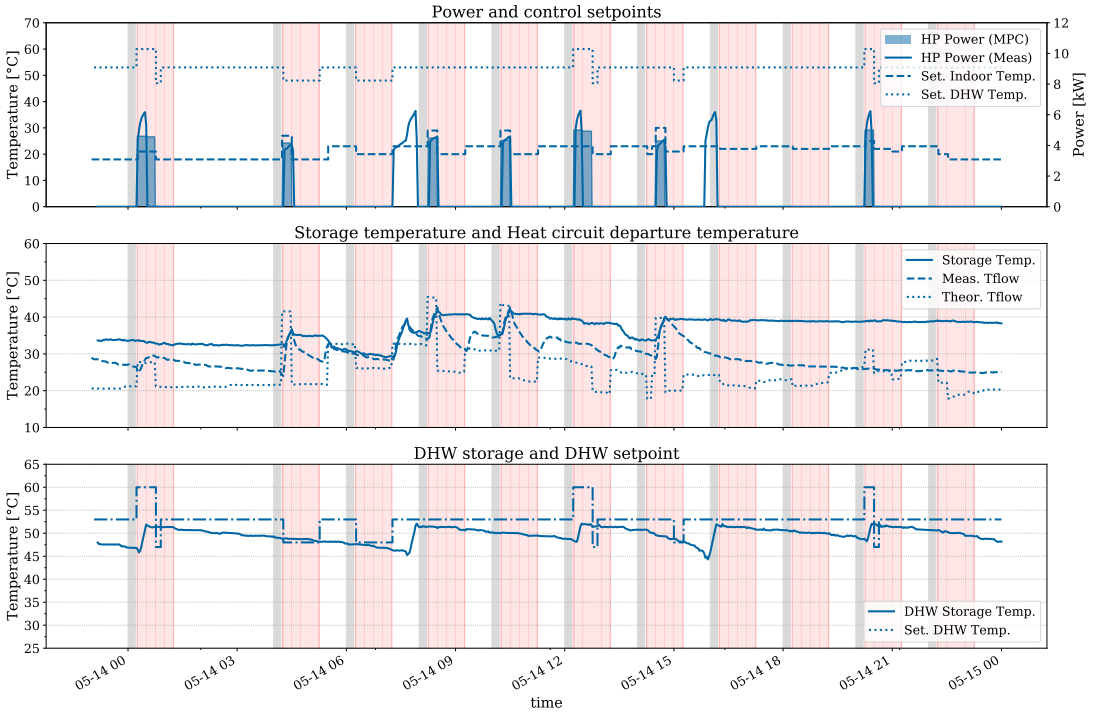


Figure 6.8 – Power and Temperature variation resulting from direct load control for Grandson

The bottom panel represents the effect of direct load control on DHW. The solid line corresponds to the DHW temperature sensor (one per storage). The dotted line represents the setpoint value used to control DHW cycles.

It is not always easy to assess what a system would have done without direct load control, but coupling setpoints, temperature and power measurements can visually help. As a reminder, the heat pump's local control works with a hysteresis on the temperature of each storage. When storage temperature drops too far below the setpoint value of the hysteresis, the compressor starts, and the heat pump runs until the upper value of the hysteresis is met. This is of course the theory, but other unforeseen events or aspects can sometimes change this behaviour.

Figure 6.8 displays the results of direct load control for the building Grandson. The building has a night-setback for SH. This means that setpoint values are reduced during the night. The building seems to perform well in terms of direct load control. Every time the MPC solution proposes to run the heat pump for SH, setpoint values for SH are increased (dashed line of the top subplot) and the compressor starts. When heat pumps need to be stopped, setpoints are decreased, which accurately stops the heat pump. We can easily link each power consumption to SH or DHW by looking at which storage tank temperature increased at the given time. Between 06:15 and 07:15 we can see a perfect example of a successfully delayed cycle of a heat pump. The value of the theoretical departure temperature is reduced over all the period therefore keeping the hysteresis lower bound below the SH buffer temperature. At the end of the DR call, as soon as the setpoints are set back to the default value, the heat pump starts.

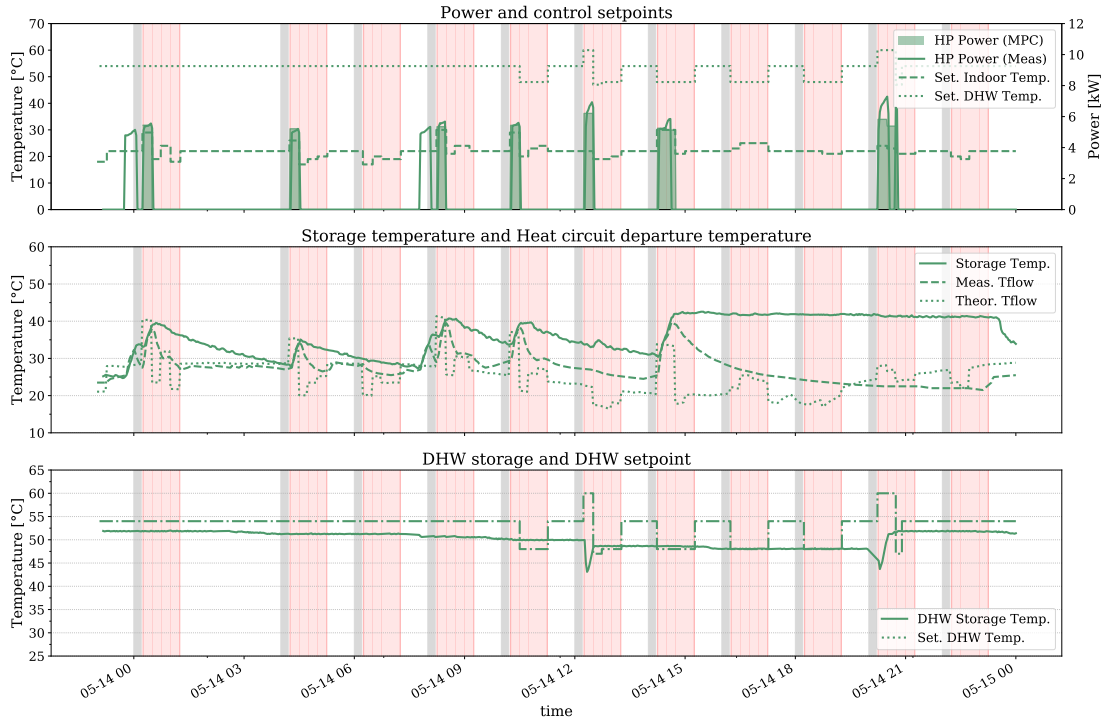


Figure 6.9 – Power and Temperature variation resulting from direct load control for Jegenstorf

This is shown as an increase in temperature inside the buffer tank.

The main identified failure here comes from the direct load control when we increase the DHW setpoint. The model of the MPC foresees that the heat pump can run for two consecutive periods (blue area from 00:15 - 00:45 and from 12:15 - 12:45). Even with the DHW storage temperature standing 7°C below the new setpoint value, the compressor stops, probably due to high (>60°C) departure temperature from the condenser. Figure 6.9 depicts the results of direct load control on a building with floor heating. This building performed well, with most of the tracking successfully implemented. As opposed to some other buildings, theoretical T^{flow} and measured T^{flow} match well. The system has a certain inertia, therefore changes between two values take time before stabilizing. Because the system operates with low temperatures and the heat pump is over-sized, the buffer tank quickly fills up in less than 30 minutes. As increasing the “Room temperature setpoint” directly affects the distribution system temperature, it is practically impossible to raise the buffer temperature higher without generating serious discomfort. In the afternoon after 15:00 and the last charge of the SH buffer, the temperature in the buffer tank remains constant for more than 9 hours (middle panel). The measured T^{flow} follows an exponential decay down to values close to 22°C which suggests that this decrease corresponds only to thermal loss through the pipes with no mass flow rate. Due to the high external temperatures, the heat pump seems to have activated the summer mode and deactivated the heat distribution system. This phenomenon is present in all buildings (Figure 6.8 to 6.12). Increasing the setpoints does not necessarily allow shifting back to space

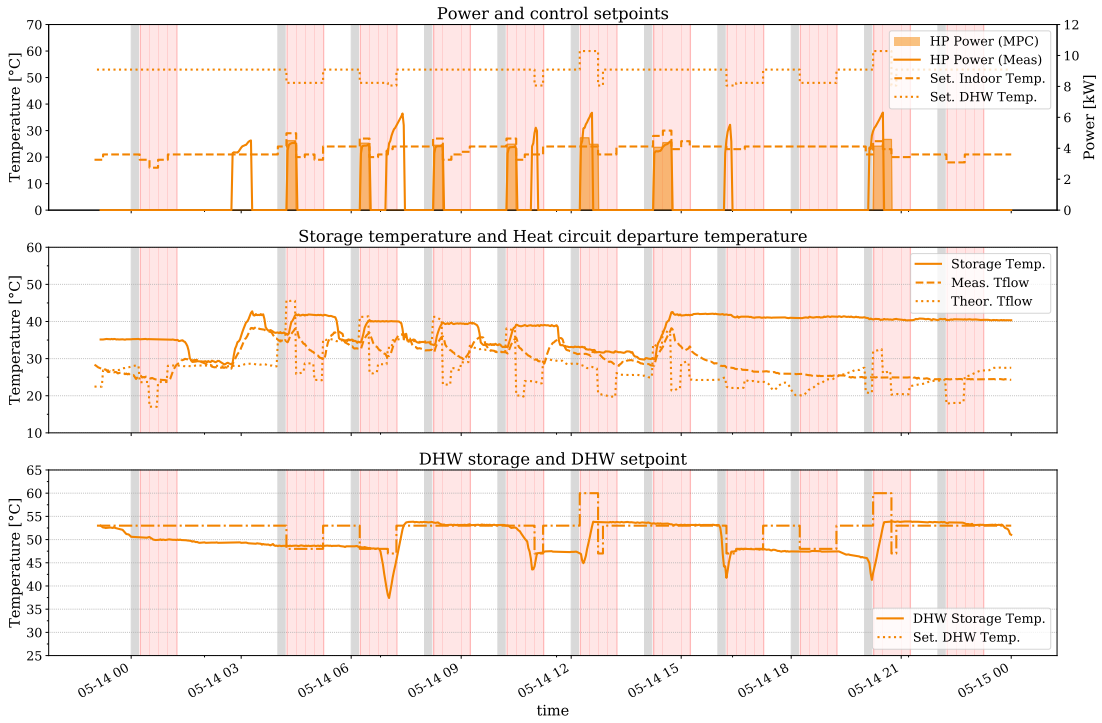


Figure 6.10 – Power and Temperature variation resulting from direct load control for Lenzburg

heating. Figure 6.10 displays the results for building Lenzburg. Direct load control of SH perfectly matches the results of the tracking MPC. However, direct load control of DHW leads to worse results than direct load control of SH and this is also the case for the other buildings. Direct load control with DHW appears to be more difficult due to the bigger uncertainty in consumption forecast. As opposed to SH, where both the demand and the production can be regulated, for DHW only the production can be controlled. Having only one sensor to assess the energy state inside the DHW storage tank makes it complex to predict when a new cycle will occur. For comfort reasons, DHW is always prioritized and setpoints are only reduced to a minimum of 47°C. Therefore, delaying a DHW cycle for more than 30 minutes is not always possible, as demonstrated for the DR call at 06:00. In the bottom panel, we can see that the storage temperature at the start of the period is low, which is why setpoints are set to the lowest value possible. At 06:40 a DHW consumption brought the storage temperature below the lower bound of the hysteresis, which starts a new DHW cycle.

The DR call at 10:00 is a perfect example of the usefulness of MPC when dealing with direct load control. When the power traces are generated, the storage tank temperature is maximal. There is only a small chance that a DHW cycle will happen in the next hour. However, within the third 15 minutes interval, a sudden high DHW consumption puts the storage temperature below the lower bound of the hysteresis and the heat pump starts a new DHW cycle. At 11:00, to avoid deviating further from the trace, the DHW setpoint is reduced, which directly stops the heat pump. The same happened a second time at 16:00.

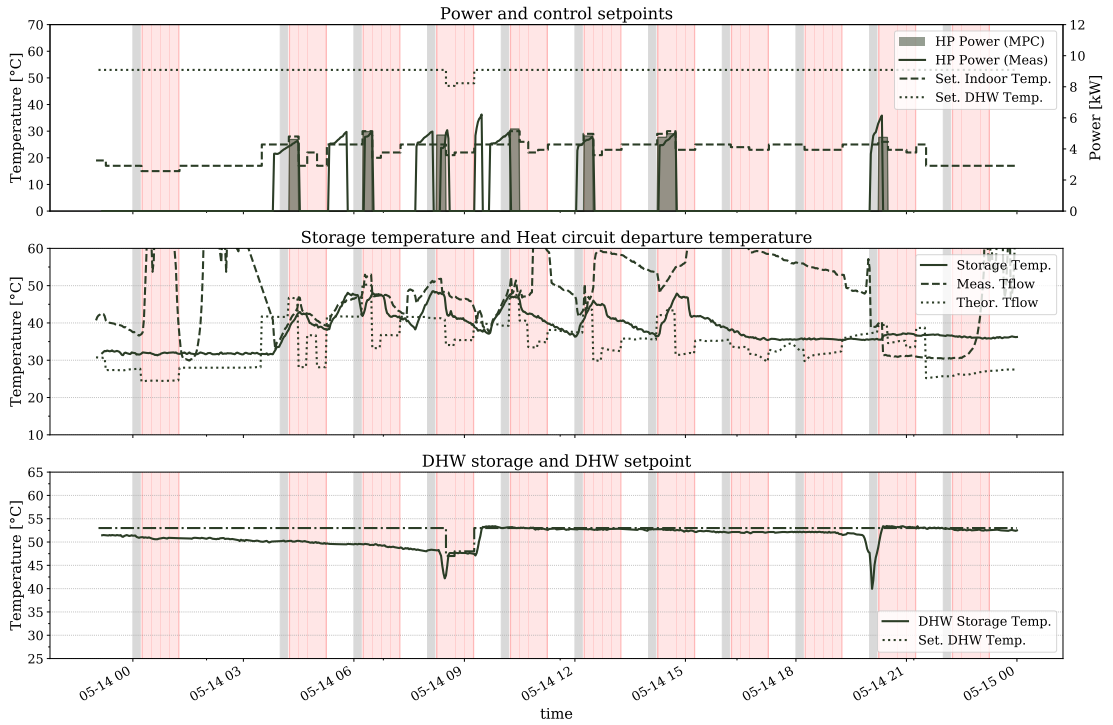


Figure 6.11 – Power and Temperature variation resulting from direct load control for Montebelo

Looking at the top panel, we can see that power consumption predicted by the tracking MPC is often underestimated. Even worse, the power given as a solution of the MPC sometimes decreases for the second time steps. Increasing the operation temperature in reality reduces the COP of the heat pump and increases its electricity consumption, which is clearly visible on the graphs. This peculiar behaviour is the result of the power formulation as a function of the storage temperature used in the MPC. A DHW cycle will almost always create a drop in temperature for a few minutes before we can see an increase. If values are not correctly updated at the start of the next period, the optimisation will assume a lower temperature and therefore a lower power consumption.

Figure 6.11 displays the results for the building Montebello which has the particularity of returning wrong measurement values for the departure temperature (Meas. T^{flow}). These failures happen at random times and for random durations. In the early hours of the day, measured T^{flow} jumps from 30°C to 60°C without any setpoint modification nor heat pump start-up, which is an artefact. They are extremely difficult to filter out as they are not typical outliers (i.e extreme value, maximal value, negative value) but tend to be just shifted by 10°C to 20°C. Twice (at 12:00 and 20:00) the heat pump switches ON during the negotiation phase. The first time it turns on the SH but does not affect the results, as the heat pump is able to continue its SH cycle. However, the second time it turns on the DHW, which affects the results, as the heat pump is not able to switch to SH on turns OFF.

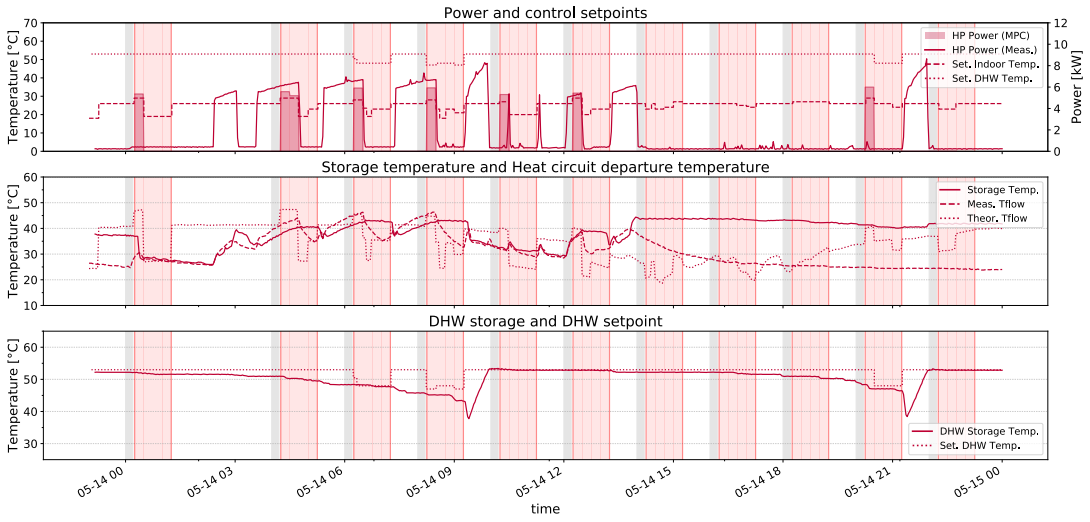


Figure 6.12 – Power and Temperature variation resulting from direct load control for Oberhofen

Figure 6.12 displays the only building where the power is not computed based on the operation temperatures but directly measured by a power meter. Power meters for other buildings were all changed shortly before we started the day-long tests. This building has a second specific feature: the external temperature sensor is exposed to solar radiation, leading to values frequently exceeding 30°C in the afternoon even in winter. For this reason, even when fixing "Room temperature setpoint" to the maximal value (30°C), the theoretical T^{flow} barely reaches above 30°C and the heat pump cannot be started.

This building has the highest consumption for SH. Thanks to direct load control the durations of the three SH cycles from 03:30 to 08:15 are successfully extended and stopped when needed. However, 3 out of the 8 upward load control predicted by the MPC failed to happen on the site. At 00:15, "Room temperature setpoint" is increased, which leads to an increase of the departure temperature and the emptying of the storage. Normally the heat pump should have started as was the case in similar conditions during previous tests. This appears to be an unforeseen failure that is not explained so far. In the afternoon, for the 16:00 and 18:00 direct control service, the heat pump is kept shut down without decreasing the setpoint because this would have been the baseline.

For the DR services at 10:00 and 20:00, we see two perfect examples of successful delaying of a DHW cycle. The setpoint is set to a reduced value for all the DR call duration and the heat pumps do not start although it would have. But this is also a perfect demonstration of the famous "rebound effect" that can happen after direct load control. In both cases, The heat pump runs for 45 minutes strait after setpoints are reset to their default values.

The model of the MPC used to estimate the power consumption is a piecewise linearisation of manufacturer datasheets based on the buffer temperature. For high temperatures it tends to underestimate the power consumption by 15%.

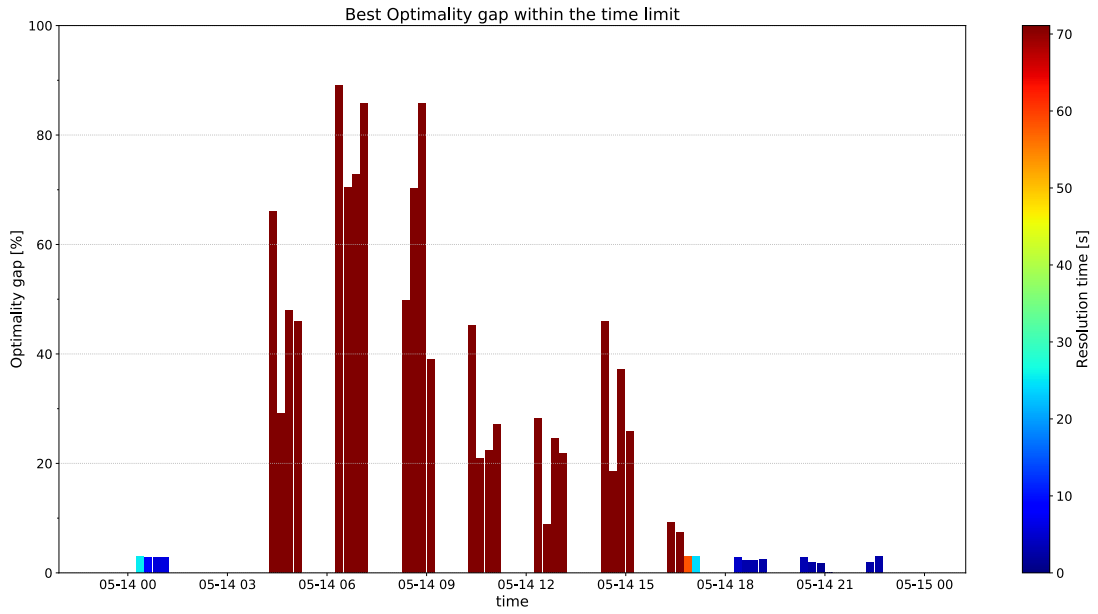


Figure 6.13 – Solver performance: optimality gap in the y-axis and resolution time in the color map

6.3.3 Solver performance

The solver used for the scheduling and the tracking MPC is gurobi (version 8.1), a leading commercial software. It can solve Mixed Integer Linear Problems (MILP), which correspond to the formulation used for the MPC at the pilot site. Models for the building and the storage are linearised but integers are required for choosing between SH and DHW and for the piecewise linearisation of the heat pump model. Normally, the optimality is proven if the difference between the upper bound (found by heuristic methods) and the lower bound (found by partial exact method) is equal to 0, i.e. an optimality gap of 0%. For real time control this is too extreme, so we assume a solution to be optimal from a relative MIP gap < 3%. The solver is stopped as soon as the gap between the best possible and the best-found integer solution drops below this value. We cannot, however, wait indefinitely for even this sub-optimal solution. Having a solution is better than having no solution and we want to have it every 15 minutes. Therefore, we set a time limit of 70 seconds after which the best solution is accepted for the new actuation. The quality of the solutions of the solver is presented in Figure 6.13.

Figure 6.13 displays the relative MIP gap tolerance percentage as well as the associated solving time. Each bar corresponds to one tracking problem. As DR calls are 1 hour long with 15 minute timesteps there are four bars per hour. A first thing to notice is that more than half of the optimisations fail to reach the relative MIP gap during the 70 seconds time limit leading to apparent sub-optimal solutions and relatively high MIP gap percentages. For MILP, the relative MIP gap can be reduced by finding better solutions through a combination of heuristic and partially exact methods or by cutting the solution space, increasing the lower bound. By

waiting until the 3% tolerance was reached, we could have maybe found that a solution with a relative optimality gap of 80% would rather have increased its lower bound than reduce its upper bound.

Solutions reaching the 3% tolerance often take less than 15 seconds and are more frequent during the evening and night. This seems to be related with some SH constraints that are more active when the mean external temperature is high, and SH is in summer mode. The solver struggles more to find a solution when it must choose between SH and DHW.

6.4 The Flexibility Function

This subsection presents a method to estimate the Flexibility Function (FF) of the pilot site, based on the work from (Junker et al., 2018). The authors proposed to characterise the energy flexibility as a dynamic function. They describe the relation between a penalty signal λ_t (e.g. from an aggregator) and the response Y_t of the system controlled as

$$Y_t = \sum_{k=0}^{\infty} h_k \lambda_{t-k} + R_t, \tag{6.4}$$

where R_t is the non responsive load. The term h_k corresponds to the inpulse response function. They define the FF as the step function by "finding the expectation at time t when $\lambda_k = 0$ for $k < 0$ and $\lambda_k = 1$ for $k \geq 0$.

$$FF(t) = \sum_{k=0}^{t} h_k \tag{6.5}$$

However, their work is based on a penalty signal that does not exist in our case and so the FF methodology defined by (Junker et al., 2018) is not directly applicable. An alternative method is developed by partners of the project (Gerard Mor and Jordi Cipriano from the CIMNE institute) and applied to the resulting loads of the tests. This method also accounts for the effect of exogenous variables like the outdoor temperature.

6.4.1 Modeling

Figure 6.14 depicts the performance of the heat pumps tested from April 3rd to May 15th for direct load control. The granularity of the data is 2 minutes and it is aggregated from the five responsive buildings. During the field test, multiple activation traces are tested, as explained in Chapter 5. The heat pumps' operation is affected by these traces. Therefore, differences between realised P^e and baseline P^b averaged electricity power are expected during activation periods. The alternative method used to characterise the flexibility of this 5-building cluster accounts for the theoretical power difference between the activation trace and the baseline, and the real power difference between the actual trace and the baseline. The real power difference is modelled in terms of the theoretical difference and its autoregressive terms as

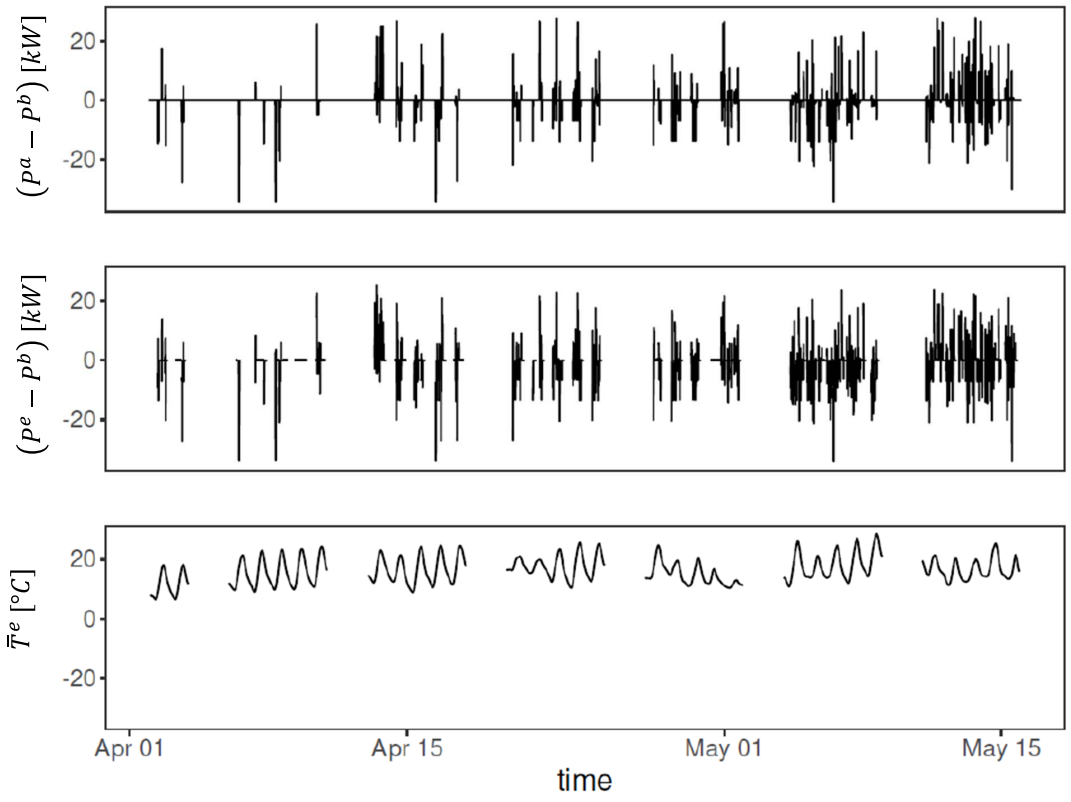


Figure 6.14 – Difference between the activation trace P^a and total baseline power P^b, versus the difference between the total actual power P^e and P^b, and the 4-hours moving-averaged outdoor temperature

presented in Equation 6.6.

$$\Phi^{s,T^e}(B)\left(P_t^e - P_t^b\right) = \omega^{T^e}(B)\left(P_t^a - P_t^b\right) + \varepsilon_t \tag{6.6}$$

Here, P_t^e is the measured power of the heat pumps, P_t^b is the baseline power when the control is used with BaU schedules and P_t^a is the "theoretical" activated power trace of the heat pumps. $\Phi^{s,Te}$ and ω^{Te} are the autoregressive terms depending on two categorical variables which affect the estimation of the terms according to the value they take. These categorical variables are: (s) which refers to the sign of the activation trace whether it is positive (1), negative (-1) or zero (0) compared to the baseline power P^b, and T^e which refers to the outdoor temperature levels. Based on the 4-hour moving-averaged outdoor temperature during the test periods, we split the results into two groups of temperature level: [6.5 °C, 15.7 °C) and [15.7 °C, 28.5 °C]. Therefore, neither (s) or T^e are used as exogenous variables of the model. The backward shift operators B are defined as $B^k y_t = y_{t-k}$, where y_t represents each considered series $(P_t^e - P_t^b)$ and $(P_t^a - P_t^b)$ at time t and $k \in [0, j]$. Here, j refers to the maximum order allowed to that backward shift operator B and ε_t is the white noise residual of the model at time t.

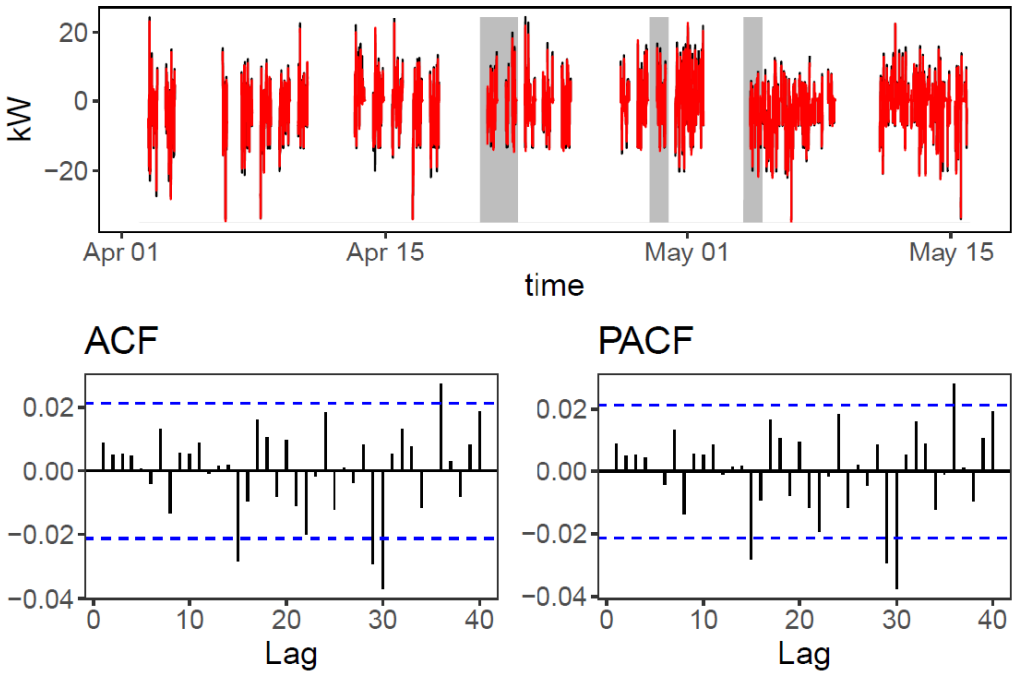


Figure 6.15 – Flexibility model for the Swiss pilot site; the upper graph shows a comparison of $(P^e - P^b)$ (black line) and the predicted one performed with the flexibility model (red line); the lower graphs show the auto correlation functions of the training period residuals. The Auto-Correlation Function (ACF) and Partial Auto-Correlation Function (PACF) are within the blue dated lines (the residuals are not auto-correlated)

To avoid over-fitting the training set, the data is split into two parts. 90% of the data is used to train the model and the other 10% to validate it. The top panel in Figure 6.15 depicts the training period (white) and the validation period (grey). The realised load (black line) is compared to the predicted load using the FF model (red), corresponding to the difference in power between actual consumption and the baseline as described in Equation 6.6. There are no significant differences between the training and the validation periods, so over-fitting issues are avoided. Furthermore, the AutoCorrelation Function (ACF) and Partial AutoCorrelation Function (PACF) plots do not indicate any auto-correlation in the residuals. Hence, the training residuals can be considered as white noise, which is a necessary step in order to validate the model as capable of estimating the FF.

6.4.2 Estimation of the Flexibility Function

In this estimation of the FF, a 100% activation signal of length n = 15, 30 and 60 minutes is evaluated during a period of i = 120 minutes. We use a multi-step prediction method to estimate the FF, predicting the expected response of ±1 kW activation trace, where the previous estimate of the flexibility function $\hat{P}^e$ is used for the new prediction step. We set the baseline power $P^b_t = 0$ for $t \in (0, 1, ..., i)$. Here, s is equal to 1 if the activation is positive and -1 if negative.

$$\left(P^{e}_{t\leq 0} - P^{b}_{t\leq 0}\right) = 0 \tag{6.7a}$$

$$\left(P^{a}_{t} - P^{b}_{t}\right) = \begin{cases} 0, & \text{if } t \leq 0 \\ \left(\underbrace{(s,...,s)}_{n \text{ times}}, \underbrace{(0,...,0)}_{i\text{-}n \text{ times}}\right), & \text{otherwise} \end{cases} \tag{6.7b}$$

$$\Phi^{s,T^{e}}(B)\left(P^{e}_{t} - P^{b}_{t}\right) = \omega^{T^{e}}(B)\left(P^{a}_{t} - P^{b}_{t}\right) \tag{6.7c}$$

$$\Phi^{s,T^{e}}(B)_{k=0}\left(P^{e}_{t} - P^{b}_{t}\right) = -\Phi^{s,T_{o}}{}_{k\geq 1}(B)\left(P^{e}_{t} - P^{b}_{t}\right) + \omega^{T^{e}}(B)\left(P^{a}_{t} - P^{b}_{t}\right) \tag{6.7d}$$

$$\Phi^{s,T^{e}}(B)_{k=0} = 1 \tag{6.7e}$$

Considering the FF model in Equation 6.6 and the set up describe in previous equations, the FF is defined as:

$$\begin{aligned} FF_{t} &= \left(P^{e}_{t} - P^{b}_{t}\right) \\ &= -\Phi^{s,T^{e}}(B)\left(P^{e}_{t-1} - P^{b}_{t-1}\right) + \omega^{T^{e}}(B)\left(P^{a}_{t} - P^{b}_{t}\right) \end{aligned} \tag{6.8}$$

6.4.3 Application of the Flexibility Function

Figure 6.16 shows the FF for the pilot site . The left Y-axis describe the change in power ($P^{e} - P^{b}$) and the right Y-axis describes the change in power due to the activation negotiated with the aggregator ($P^{a} - P^{b}$). The flexibility is analysed for two different outdoor temperature levels; low-to-mid range [6.5 ºC, 15.7 ºC) in yellow and mid-to-high [15.7 ºC, 28.5 ºC] in black. Two types of activation (e.g. red signal [-1, 0, 1]) are also tested: (1) *Negative*, when the consumption is lower than the baseline, and (2) *Positive*, when the consumption is higher than the baseline. The terms “Negative” and “Positive” used here have to be differentiated from the existing positive (Upward) and negative (Downward) reserve services defined in market regulation and provided by conventional generators. In this methodology, the term “Positive" refers to an increase in power consumption compared to the baseline, which, from a market perspective, is equivalent to a decrease in power production (negative reserve).

At low-to-mid range outdoor temperature

In the case of tracking *negative* activation traces (left panels), the actual power follows 80-90% of the theoretical activation load during the first 15 minutes, reaching the maximum deactivation peak (98%) after 13 minutes. Then, the deactivation decreases to 75% after 30 minutes, maintaining this percentage over the 60 minutes.

When tracking a *positive* activation trace, the actual power follows 80-90% of the theoretical activation during the first 15 minutes, to decrease linearly to 50% after 30 minutes and maintain this activation with a small rebound (+10%) up to the 60 minutes. This means that buildings can provide the amount of flexibility estimated by the models during the first 15 minutes but then, the limited availability of thermal energy storage in the building (either for SH or DHW)

does not allow for full activation compliance. In both cases, the rebound effect of up to 30% change in power, starts just after the activation/deactivation of the trace and its peak is after approximately 13 minutes.

At mid-to-high outdoor temperatures

In the case of mid-to-high outdoor temperatures, the flexibility peak of the first 15 minutes no longer exists. This can be explained mainly because the buildings have even less thermal storage capacity if the thermal comfort should to be maintained. In this case, the main control variable used to shift energy consumption is the DHW temperature setpoints. The average fulfillment of activation traces is 60% in the case of positive activation traces and and 75% in the case of negative traces. The rebound effects follow the same path as in the lower temperatures case but with smaller peaks.

The drawback of the method is that it is only as good as the training set provided. Non-occurring events will be poorly addressed. The advantage of the method is that it can be used for characterising the flexibility of a single building or the aggregated flexibility of several buildings together. The FF of clusters can then simply be added together. Although the size of the cluster is too small to draw a strong conclusion, the FF of a group of buildings can give us a rough idea of how many buildings are required to reach a certain power and energy deviation. In the case of the five buildings cluster and mid-to-high outdoor temperatures, 6.16 shows us that more than $1/0.6 = 1.67$ equivalent clusters should have been activated to meet the activation goal.

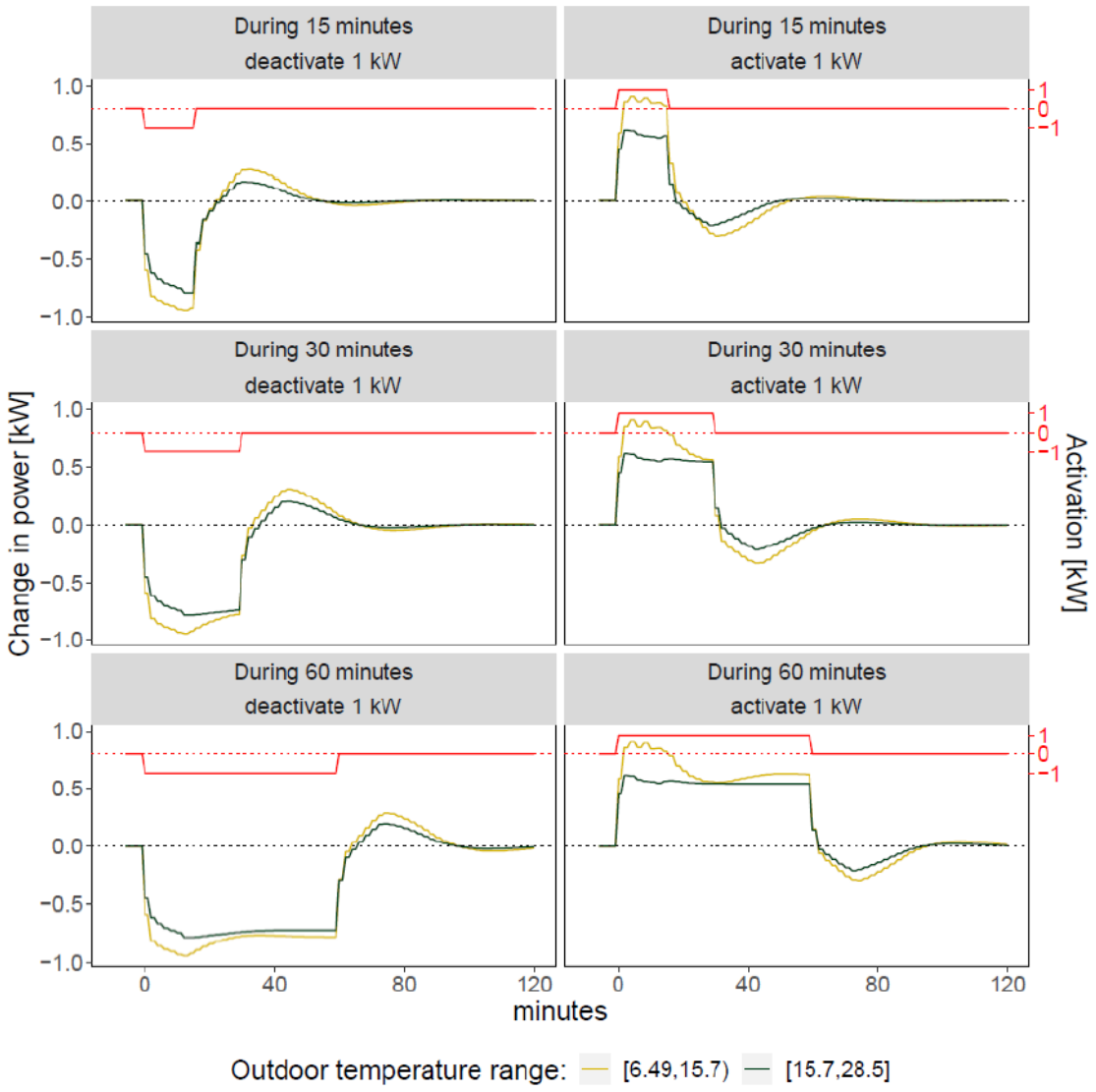


Figure 6.16 – Flexibility Function (FF) of a 5-buildings cluster

7 Heat Pump for Demand Response: Lessons learnt

The goal of this thesis was to assess the benefits and challenges of a flexible control of heat pumps in residential buildings. In a first step (Chapter 2), we used a simulation-based setting to test our building and heat pump models and Model Predictive Control (MPC) algorithm. We then moved to a pilot site of twelve inhabited residential building blocks in Valais (Switzerland), which we equipped for monitoring and remote control of their heat pumps (Chapter 3). Next, we assessed how well the pilot site operates once connected to our Demand Response (DR) control system using monitoring data and stand-alone tests (Chapter 4). We showed that we are able to perform DR services at a local scale. In Chapter 5, we showed that our DR control system can also be used at a larger scale that includes a third party to access energy markets. Finally, we assess the reliability and performance of DR transactions between the pilot site and an aggregator in Chapter 6, and show that we achieve notable 91% heat pump activation success and 50-95% fulfillment of the activation traces.

There are many propositions for advanced DR algorithms that are mostly applied only to simulated data, thus never testing if the assumptions made *in silico* also apply to the real world, e.g. (Reynders et al., 2018). Further, very few applied research projects report the success but also the challenges faced when implementing advanced control methods on real systems. One example is Cigler, Gyalistras, et al. (2013), who reported their lesson learnt from field tests concerning the application of Model Predictive Control in academic buildings. Since one of the major challenges of this thesis was to implement a DR control system that works in real life, we here present our lessons learnt from implementing a DR system in real residential buildings.

This chapter is based on the work that will be presented at the next 13th International Energy Agency Heat Pump Conference (HPC2020) postponed to April 2021 (Menon et al., 2021). It is structured as follows: Section 7.1 summarises the problems encountered during the project for DR participation with residential heat pumps. Section 7.2 builds on the lesson learnt to propose solutions and highlight future outlooks in the field. Finally, section 7.3 concludes our work.

7.1 Problems encountered for Demand Response participation with heat pumps

Heat pumps have proven to be difficult to use in DR applications. Conventional heat pumps have a non-linear behavior, which makes it difficult to model them for advanced predictive control methods, as presented in Chapter 2. The existing internal logic within heat pumps, and certain industrial installation practices of heat pump systems, make it even more complex. The problems we encountered during this project can be divided into four main categories: (i) installation issues, (ii) communications issues, (iii) control algorithm design issues, and (iv) market issues.

7.1.1 Installation issues

For most existing buildings, certified installers install the heat pumps. Each manufacturer manages the existing certification framework. The manufacturers provide a set number of configurations for certain types of buildings and the installer has to conform to these configurations. Unfortunately, the blueprint of the thermal installation often does not conform to the certified configurations (e.g. pipe connection, number and location of the sensors).

The monitoring issues reported in Chapter 3 and Chapter 4 result from the sudden inter-connectivity of systems that were previously not connected. Many industrial fields (e.g. building physics, control, power system) have their own specific standards and protocols. However, these industry standards do not translate well in the age of the Internet of Things (IoT) and for use with DR. Proprietary standards, and the lack of conversion interfaces, are the biggest technical obstacles preventing devices like heat pumps from being used for DR services.

Of the existing heat pump manufacturers, most offer interfaces that are capable of converting the internal logic and standards to inter-operable standards like ModBus or BACNet. This is the first step in the right direction. However, quite a few manufacturers provide this option solely for monitoring and not for remote reading and writing of variables for control purposes. Out of the twelve buildings at the pilot site, only seven buildings are equipped with heat pumps that enable remote control for DR purposes. The other five have heat pumps that provide data only for reading purposes. In addition, even the number of monitored values available varies from manufacturer to manufacturer from more than hundreds of data points to less than a dozen. This makes it particularly difficult to develop generic models and control frameworks that work for all heat pumps.

7.1.2 Communication Issues

During this project, we have seen the importance of reliable and extensive monitoring data at each fundamental phase of modeling, validation and real time control. As presented in

Table 7.1 – Communication failure: Number and duration of monitoring interruptions (missing data) for the pilot site from January 2019 to September 2020.

Monitoring disruption	Lenzburg	Oberhofen	Jegenstorf	Montebello	Grandson
interruption < 1h	7	6	17	8	7
1h < interruption < 24h	10	9	12	9	8
interruption > 24h	5	5	2	5	6
Total	22	20	31	22	21
Maximum duration	18d 0:36	56d 6:54	5d 0h20	32d 3h44	40d 7h26
Total duration	43d 11h56	95d 6h38	9d 3h58	54d 8h46	73d 21h44

Chapter 3, data integrity and data consistency need to be ensured. This requires a large data screening effort to keep the pertinent data and identify faulty behaviours.

In the case of residential DR, multiple buildings have to be aggregated. If these buildings are equipped with smart meters or gateways, the different measurement signals need to be centralised and sent to the supervision unit via an internet connection. This connection is, for the moment, rather slow and does not allow a real-time control of the devices. This makes it impossible to respond to the short activation time criterion. Several issues with monitoring reliability were continuously noticed throughout the entire project. Monitoring systems can fail during a few hours and up to several weeks, as reported in Table 7.1.

Most of interruptions shown in Table 7.1 resulted from a problem on-site (e.g. IP reset of the gateway by the internet provider). The software installed at the pilot site was able to resolve some of them automatically, but for the longest disruptions the intervention of a technician on-site was necessary. Communication failure can also arise between the centralised controller and the online management platform. As presented in Section 5.2, we developed an interface capable of handling most communication congestion problems.

7.1.3 Control algorithm design issues

Heat pumps are non-linear in behavior because their Coefficient of Performance (COP) depends on the operating temperatures of the cycle (Fischer et al., 2016; Salpakari & Lund, 2016; Vrettos et al., 2013). The part-load efficiency of the compressor also varies non-linearly with the compressor speed (Kim et al., 2014). When they are associated with stratified storage tanks (i.e. non-linear) or with floor heating system, which have high inertia, it is even more complex to translate their behaviour into simple equations. As a result, it is challenging to provide time-series of the electric power consumption and thermal power output within the time constraints imposed by electricity markets. We proposed a Mixed-Integer Linear Programming (MILP) formulation, which works for both simulated data and on real heat pump systems. However, it did not scale well with increasing building number, and global optimality was rarely achieved within the time constraint.

Ideally, we want to develop a model structure that is as generic as possible, i.e. find a minimum common ground for all buildings and heat pumps. In residential buildings, especially for buildings from before the year 2000, there is often a lack of monitoring, and information about building design is often missing. Adapting the models developed for simulations to the constraints of the real system was challenging. Assessing the state of charge of a building is far from simple in reality, as indoor temperature sensors are not always available or placed at useless locations. For example, comfort can not directly be assessed if sensors are located in non-heated areas like the staircases. State-of-the-art techniques like the Kalman filter should help updating the state for closed loop control and improve the performances and stability of Model Predictive Control (MPC). However, due to the lack of variables, it was more difficult to set up than expected and was therefore not implemented.

The deployment of MPC is a long-term task that requires careful monitoring, tuning and adjustments. We showed that, depending on the characteristics of the heat pump and its associated heat distribution system, an efficient implementation of MPC can be very complex, if not even possible. This is because MPC is better at taking high level decisions (e.g. switching the heat pump ON or OFF), than at directly specifying the low-level control actions. The centralised MPC we developed (Chapter 2) has the advantage of being fairly easy to formulate, it clearly has a scaling problem with increasing building cluster size or increasing prediction horizon. Our choice to include local constraints directly in the formulation can be questioned and should be further evaluated. Adding these constraints has a direct impact on the tractability, convergence rates and resolution time required for solving the problems. The addition of these constraints has increased the convergence times from a few seconds for each time step to anything from a few minutes to the maximum allowed 10 minutes (as the electricity balance markets use 15 minutes as the actuation time limit). Finally, the tuning of the weighting parameters used in the cost function to drive the solution towards different objectives is challenging as, at our knowledge, no methodology exists to help with this crucial task.

When offering DR services at a larger scale that includes a third party to access energy markets, as evaluated in Chapter 5 and Chapter 6, it is critical to assess and share the availability of the heat pumps. The transactive DR approach between the pilot site and a third party has revealed some limitations, especially concerning the baseline power estimation. Scheduling and tracking a power signal with a small number of heat pumps for secondary control reserves did not provide good results, as any deviation has a large impact on the results. Alternative control methods relying on statistical models for power load tracking could be used, but require a more controllable devices than are available at the pilot site.

7.1.4 Market issues

In 2013, Switzerland went through the major regulatory changes of opening ancillary services to DR and decreasing the minimum bid to 5 MW or less. In addition, the symmetric bidding re-

quirement was removed from most of the markets to encourage better consumer participation. Despite these developments, there remain some barriers to DR:

- The minimum bid is still 5 MW for the secondary reserve market (identified as a target market in this thesis).
- The high number of DSOs (> 700) present in Switzerland means that each of them has a limited market. This does not incentivise the DSO's to develop DR services.
- The high participation of hydraulic power in the reserve markets limits the participation of residential DR.
- The electricity market is not liberalised for small companies or households.

Existing markets are not designed for DR with residential buildings, so any modification required to promote DR requires regulation changes. The access to electricity markets for residential heat pumps is complicated, as their individual power consumption is low and thus many installations need to be aggregated. This is a major drawback compared to hydraulic power, their direct competitor in Switzerland. The higher complexity and cost of coordinating many heat pumps is borne by the aggregator. In addition, the short activation time often required by the markets could prevent the inclusion of residential DR in the different electricity markets.

The delays in the project and the rather short duration of the DR tests on the pilot site do not allow us to estimate the potential earnings for DR within residential buildings. However, partners of the Sim4blocks project simulated an access to the day-ahead and intra-day market using a simulation tool they developed. They situated potential earnings between 30 and 50€ per year and per system for single family household. For the Frequency Containment Reserve (FCR) market, Posma et al. (2019) compared an "always reliable" biding strategy (i.e. more conservative) with an "always available" biding strategy. They achieved a maximum revenue of 1€ per heat pump per week with their "always reliable" biding strategy, highlighting that failures to correctly respond to frequency deviations are heavily penalised. We can relate these results with the reliability achieved during our field tests as presented in Section 6.2. The actuation success for the cluster of building was on average 91%. However, the average fulfillment of the activation traces, based on the Flexibility Function methodology presented in Section 6.4, was only ranging from 95% down to 50%, depending on the external temperature and the duration of the DR service.

These potential earnings can be compared to the installation cost for equipping one residential building to provide DR. A sample of these installation costs has been gathered from different projects and is presented in Table 7.2.

These costs include the hardware cost (e.g.the gateways, the extra sensor) and the installation cost for one system, both for single or multi-family houses. As most of these projects are still at

Table 7.2 – Installation cost for equipping one building to provide DR from different projects

Range of costs	Project	Source
800 - 900 CHF	(“Sim4blocks”, 2016-2020)	Interview with engineer (Swiss site)
3000 - 3500 EUR	(“Sim4blocks”, 2016-2020)	Interview with engineer (German site)
1000 - 1200 CHF	(“GOFLEX”, 2016-2020)	Interview with Coordinator (Swiss site)
1500 CHF	“Neurobat” (2014-2015)	Interview with co-founder

the research phase, the cost of the development and maintenance of the software as well as the cost for a commercial database is too high to be relevant and is not included. Therefore, each solution developed in the different projects does not include the same level of detail due to the system configurations and goals specific to the project. The high difference in price reported between the Swiss and the German pilot site of the Sim4Blocks project is mostly related to how much computational power is available in the system installed. Based on the monitoring performance reported over the last two years of the project, the German system proved to be more robust with almost zero onsite interventions needed, while the Swiss system reported between two and five interventions per system as seen in Table 7.1.

Even if the costs reported in Table 7.2 are associated to different system configurations, they are still relevant to give an order of magnitude of the costs. This shows that the potential of these markets is overall relatively small for the moment. The installation costs being high, residential DR is not guaranteed to be profitable. Even if industrialised production is expected to divide the installation cost by more than a factor two, the whole business model of accessing the considered market is still challenged by the maintenance cost. Of course, the latter depends on the country, as labor costs vary significantly, but if a technician needs to intervene several times within the twenty years of operation of the system, the cost drastically increases. Thus, residential DR seems not mature enough for existing wholesale or reserve markets.

7.2 Potential solutions for Heat Pumps

With what we have seen from the previous section, it seems that the first thing to address would be to improve the access to monitoring data from heat pumps. At the moment, the use of proprietary protocols prevent heat pumps from being accessed and remotely controlled in a reasonable manner. One way of doing this would be to not only define the transmission protocols or the heat pump’s status (e.g. smart grid ready), but also to define the heat pump’s reaction for a given control. Heat pump manufacturers could include the extra costs associated to better remote control within the price of their products.

One of the main challenges is to provide off-the-shelf solutions for the communication unit (i.e. gateways) that is economical and applicable to a large number of energy conversion units (i.e. a heat pump). At the pilot site, the heat pumps were manufactured by three different companies. Additionally, self-learning installation software should be capable to

identify and address the values/signals of the controller's interface, hence, reducing the set-up costs for connecting the communication unit to the cloud. In order for DR services to be economically interesting for a wide variety of buildings (typically individual buildings and small collective buildings equipped with simple control systems), the costs for associated hardware requirements must be kept as low as possible. The affordable solutions developed for heat pumps at our pilot site aim to remove one of the greatest remaining barriers to the wide-spread implementation of DR applied to thermal loads of buildings, thereby creating a market for such services at a large-scale.

The quality of the collected data and the consistency of the monitoring needs to be ensured. Our work reveals that a significant number of installations are not performing to the mark that they were built and installed for. There needs to be a framework to study the behavior of the systems over a set period of time and make changes to the system if need be, so that they work as they are listed. The installation of smart meters, the check of the heating system and its control could be sold together as a service, which would reduce the costs and enable a cost-effective deployment of residential DR.

Social survey within the Sim4blocks project on the question of DR service reveals that tenants want "control over their heating system without doing anything". User involvement was kept to a minimum throughout the project and during the entire project duration we only got one comfort complaint consecutive to a DR test. However, solutions allowing feedback from tenants concerning their thermal comfort would definitely increase the flexibility range of residential buildings as it could replace or help calibrate the indoor temperature measurements.

The aggregation of heat pumps alone does not meet the activation time and reliability requirements of most existing markets. The solution probably rather lies in combining systems composed of power-to-heat (e.g. heat pumps, cogeneration systems), electricity storages (e.g. batteries), and electricity production units (e.g. solar panels), as has been demonstrated by project partners, especially in the case of auto-consumption. These systems could also be artificially coupled together by the aggregator.

In this thesis, we highlighted the key variables and constraints that need to be accounted for in the control framework of heat pumps for successfully offering flexibility on energy markets. Special care should be put on accurately modeling the latency to ramp to full power, the minimum activation time or the minimum shut down duration. These constraints are often specific to the heat pump. They can be obtained from manufacturer data-sheets or by screening the monitoring data. The latter can be made automatic. We also showed the negative effect that space heating buffer tanks can have when offering flexibility. Compared to directly feeding heat to the house, a buffer tank offers a small storage capacity to avoid too many activations. However, the control logic associated tends to blocks or delay the activation of heat pumps. This diminishes the load tracking reliability of the heat pumps (Chapter 6 Figures 6.8 to 6.12). A simple while still accurate modeling of the interaction between the

heat pump, the heat buffer and the house needs to be investigated in more detail in future research.

The access to electricity markets for residential DR reveals to be complicated due to regulation constraints. However, future changes of legislation and technical improvements could make it become interesting and realistically applicable. The removal of the barriers and the harmonisation of the electricity markets is thus primordial. At the same time, incentives and regulatory frameworks should be developed for the deployment of local balance markets. Most of the DSO's interviewed during the project showed great interest in local markets, as they could help them to reduce the penalties they pay to the balance responsible parties for inaccurate forecasts.

Finally, there needs to be concerted effort and collaboration between the market actors, heat pump manufacturers and the designers of the latest machine learning and control algorithms. There is a major gap between the solutions presented in research papers and their implementation on real systems. Software and hardware that are flexible enough to adapt to the needs of the markets and the needs of the users in residential buildings have yet to be developed. This will also inform the market policy makers as to how to relax or tighten the rules to make it easier for heat pumps to participate.

7.3 Concluding remarks

With this thesis we demonstrated that the flexible operation of a cluster of residential heat pumps in real inhabited buildings is technically possible. It can be leveraged by local (e.g. Cluster Manager) and commercial aggregators for Demand Response (DR) services locally or in energy markets. However, the application of DR services in residential buildings requires extensive knowledge in the areas of building physics and simulation, cyber-physical systems (hard-and software), communication, data processing, business markets and optimal control. As there is no single person with such a large range of expertise, the application of DR services in residential building requires the collaboration of multiple players. We believe that concerted efforts between engineers, scientists and policy makers across different fields would benefit the development of DR in residential buildings.

As residential buildings are often poorly monitored, the installation of additional devices and their maintenance implies considerable supplementary cost. The remaining question is if the investment and operating costs of such systems can be sufficiently reduced in future installations, and if they will be compensated by the forecast revenues, which currently are still low and uncertain.

The diversity of the systems to be equipped is one of the main technical problems in the adoption of residential DR within the market. Since flexibility potentials show huge variation across buildings, simple methods based on monitoring should be developed to quickly estimate the predictive control potential for a given building.

Model Predictive Control handles multi-variable control problems with complex objectives and has allowed us to explicitly enforce constraints and to account for actuator limitations. We believe that the work presented in this thesis gives a good idea of the effort required to implement MPC in pre-existing residential buildings. We showed that depending on the characteristics of the heat pump and its associated heat distribution system, an efficient implementation of MPC can be very complex. Advanced controllers are as good as the system they operate. Further improvement can often be achieved by first fixing the existing systems.

We showed that DR for residential buildings with heat pumps is a more complex topic than often depicted and is, as of now, still immature. The flexibility offered by buildings is not "free", and successful aggregation of the flexibility is still quite tedious. This is confirmed by the bankruptcy of multiple companies in the field, by the strictly constrained electricity markets and by the absence in Europe of large scale residential DR service. Nevertheless, various options for residential DR technically exist and with the right incentives could make residential DR become competitive in a future with increasing renewable energy supply.

A Python class and function for the DR controller

```
import requests, logging, time
import os, sys
import datetime as dt
import pandas as pd
import numpy as np
from pytz import timezone
import json

if sys.version_info[0] ==2:
    import ConfigParser
    configuration = ConfigParser.RawConfigParser()
else:
    import configparser
    configuration = configparser.RawConfigParser()


class EODATA:
    def __init__(self):
        self.config = configuration
        self.config.read('config.ini') #config file contains passwords, username and is initiated when the EODATA is called.
        self.max_retry = 3
        self.token = None

  '''=========================================================
    ================= STRUCTURE DEFINITIONS ==================
    =========================================================   '''
    #===Token for retrieving data from the energysystem or asset ===
    def _token(self):
        payload = {
            'username': self.config.get('DEFAULT','username'),
            'password': self.config.get('DEFAULT','password'),
            'client_id': 'direct-access',
            'grant_type': 'password'}

        r = requests.post(
            url= 'https://eos.misurio.com/auth/realms/eos/protocol/openid-connect/token', #self.config['token_url']
            data=payload,
            timeout=(10, 10))
            # retries to connect for 10s before timeout error

        if r.ok:
            self.token = r.json()['access_token']
            return self.token
        else:
            logging.error('Problem with rest interface to get token.')
            logging.error(r.reason)
            raise LookupError()

    #=== Time serialisation  ===
    def time_serialise(self, t1):
        zurich = timezone('Europe/Zurich')
        # time has to be in format yyyy-mm-dd HH:MM
        tt1 = zurich.localize(dt.datetime(int(t1[0:4]), int(t1[5:7]), int(t1[8:10]), int(t1[11:13]), int(t1[14:16])))

        timestamp_tt1 = int((time.mktime(tt1.timetuple()) + tt1.microsecond/1000000.0)*1000)
        return timestamp_tt1

```

```
    def api_info(self, type_command): #type_command = facilities OR timeseries
        self._token()
        url = 'https://eos.misurio.com/api/'+str(type_command)+'/info'
        headers = {'authorization': 'Bearer ' + self.token,
            'content-type': 'application/json'}

        r = requests.get(
            url=url,
            headers=headers,
            timeout=(10, 10))

        if r.ok:
            return r.json()
        else:
            logging.warning('REST: failure in get request.')
            logging.warning(r.json()['errors'][0]['detail'])
            return pd.Series(np.nan)

    #=== Asset Definitions ===
    def asset_groups(self):
        max_retry = self.max_retry
        retry = 0

        while (retries <= max_retries):
            self._token()
            url = 'https://eos.misurio.com/api/facilities/asset-groups?relationCollection=tenants&relationKey=bcf69430-e948-4f3f-
    b86d-96bb0b842d65'
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                #return r.json()
                DFrame = pd.DataFrame()
                for k in range(len(r.json()['assetGroups'])):
                    X = pd.Series(r.json()['assetGroups'][k])
                    DFrame = DFrame.append(X, ignore_index=True)
                return DFrame
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json())#['errors'][0]['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== Return assets keys in each Building ===
    # Outputs a dataframe with the keys of the assets in each building
    def asset(self, building_key):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/facilities/assets?relationCollection=assetGroups&relationKey='+str(building_key)

            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                df = pd.DataFrame()
                for k in range(len(r.json())):
                    X = pd.Series(r.json()[k])
                    df = df.append(X, ignore_index=True)
                df = df.drop(['address','dataPoints','location'], axis=1)
                return df
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json()['errors'][0]['detail'])
                time.sleep(2)
                retries += 1
```

```
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    def asset_all(self): #Outputs dataframe of all assets
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/facilities/assets?relationCollection=tenants&relationKey=bcf69430-e948-4f3f-b86d-96bb0b842d65'
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                df = pd.DataFrame()
                for k in range(len(r.json())):
                    X = pd.Series(r.json()[k])
                    df = df.append(X, ignore_index=True)
                df = df.drop(['address','dataPoints','location'], axis=1)
                return df
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json()['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== Return all variables for a particular asset ===
    #Outputs a dataframe of variables for a particular asset
    def asset_variables_list(self, building_key, asset_key):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/facilities/assets?relationCollection=assetGroups&relationKey='+str(building_key)
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                df = pd.DataFrame()
                for k in range(len(r.json())):
                    X = pd.Series(r.json()[k])
                    df = df.append(X, ignore_index=True)
                df = df.drop(['address','dataPoints','location'], axis=1)
                ind = df.loc[df['key']==asset_key].index
                ind = ind[0]
                df.iloc[ind,2]
                dfv=pd.DataFrame.from_dict(df.iloc[ind,2],orient='index')
                dfv.columns = ['Variable names']
                return dfv
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json()['errors'][0]['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== Return the connection keys linking the buildings-assets ===
    def asset_var_name_connect_list(self, building_key, asset_key):
        dfv = self.asset_variables_list(building_key, asset_key)
        value = self.asset_last_value(asset_key)

        for x in value.columns:
            dfv[str(x)] = np.nan
```

```
        for y in value['timeSeriesName']:
            for x in dfv['Variable names']:
                if str(x) in str(y):
                    val = dfv.loc[dfv['Variable names'] == x].index
                    dfv.loc[val[0], 'timeSeriesName'] = y
                    val1 = value.loc[value['timeSeriesName']==y].index
                    dfv.loc[val[0], 'displayName'] = value.loc[val1[0], 'displayName']
                    dfv.loc[val[0], 'unit'] = value.loc[val1[0], 'unit']
        dfv = dfv.drop(['value'], axis = 1)

        return dfv

    #=== GET the last state of all timeseries in a list of building ===
    # It is used to get an update of the last values of all datapoints
    # If one building is disconnected, it returns nothing for it
    def all_asset_spec(self, Building_name):
        Asset_group = self.asset_groups()
        Building_ID = []

        for x in Building_name:
            for y in Asset_group['name'].values:
                if x==y:
                    Building_ID.append(Asset_group.loc[Asset_group['name'][Asset_group['name'] == y].index[0],'key'])

        Building = Asset_group[['key','name']].loc[Asset_group["name"].isin(Building_name)]
        all_as = []

        for k in Building.index:
            df = self.asset(Building['key'][k])
            as_all = []

            for i in df.index:
                as_dp = pd.Series(df['properties'][i], name= 'name')
                mask = as_dp.index.isin(['gatewayId'])
                as_dp =as_dp.loc[~mask].reset_index(drop=True).to_frame()
                as_dp['Building.name'] = Building['name'][k]
                as_dp['Building.ID'] = Building['key'][k]
                as_dp['Asset.name'] = df.name.values[i]
                as_dp['Asset.ID'] = df.key.values[i]
                as_dp['timeSeriesName'] = df.key.values[i] +'_specific:'+ as_dp + ':input'
                as_all.append(as_dp)

            as_all = pd.concat(as_all,axis=0,ignore_index=True)
            all_as.append(as_all)

        all_as = pd.concat(all_as,axis=0,ignore_index=True)
        lastvalues = self.datapoints_last_value(all_as['timeSeriesName'].tolist())
        all_as = all_as.join(lastvalues.set_index('timeSeriesName'), on='timeSeriesName')
        all_as = all_as[all_as['displayName'].notnull()]
        all_as = all_as.drop(['value.dataPointValue.status', 'displayName'], axis=1)
        all_as['value.timestamp'] = all_as.apply(lambda row: dt.datetime.fromtimestamp(row['value.timestamp']/1000), axis=1)

        return all_as

    '''=====================================================
    ================= TIMESERIES DEFINITIONS =================
    =====================================================   '''
    #=== GET last value of an asset ===
    # Return the last value of a particular asset
    def asset_last_value(self, asset_key): # Gives you the last value of a particular asset
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/timeseries/metric/'+str(asset_key)+'/lastvalue'
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                DF = pd.DataFrame()
                for x in range(len(r.json())):
                    X = pd.Series(r.json()[x])
                    DF = DF.append(X, ignore_index=True)
                return DF
```

```
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json()['errors'][0]['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== GET last value of assets (POST method) ===
    # Get the last values of assets.
    # Assets can be mixed because timeseries id include assets id
    def datapoints_last_value(self, asset_key):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/timeseries/datapoints/lastvalue'
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            payload = asset_key

            r = requests.post(
                url=url,
                headers=headers,
                data=json.dumps(payload), # info has to be json-encoded
                timeout=(10, 10))

            if r.ok:
                DF = pd.DataFrame()
                for x in range(len(r.json())):
                    DF = DF.append(pd.json_normalize(r.json()[x]), ignore_index=True)
                return DF
            else:
                logging.warning('REST: failure in post request.')
                logging.warning(r.json()['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== GET the list of all metrics ===
    # Return the list of all timeseries on the site
    def asset_all_metric_list(self):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/timeseries/metric/'
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                return r.json()
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json()['errors'][0]['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

#=== GET the list of metrics for an asset ===
# Return the list of all timeseries for a particular asset
    def asset_particular_metric_list(self, asset_key):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/timeseries/metric/'+str(asset_key)
```

```
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.get(
                url=url,
                headers=headers,
                timeout=(10, 10))

            if r.ok:
                return r.json()
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json()['errors'][0]['detail'])
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== Load datapoints from the platform between two dates ===
    # If time difference is more than a day, average values are return
    def asset_datapoints(self, asset_key, time1, time2):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url='https://eos.misurio.com/api/timeseries/datapoints/query'
            asset_key if type(asset_key) == list else [asset_key]
            payload = {'start':self.time_serialise(time1),
            'end':self.time_serialise(time2),
            'timeSeriesNames':asset_key if type(asset_key) == list else [asset_key]}
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.post(
                url=url,
                headers=headers,
                data=json.dumps(payload), #or json = payload, this info has to be json-encoded
                timeout=(10, 10))

            if r.ok:
                dp = r.json()
                list_of_df = []
                df = pd.DataFrame()
                for x in dp:
                    name = []
                    unit= []
                    values = []
                    name.append(x['name'])
                    unit.append(x['unit'])
                    values.append(x['values'])
                    dfs = pd.DataFrame()

                    for Y in range(len(values[0])):
                        dfs = dfs.append(pd.Series([dt.datetime.strptime(str(dt.datetime.fromtimestamp(int(values[0][Y]['timestamp'
])/1000)), '%Y-%m-%d %H:%M:%S'), values[0][Y]['dataPointValue']['value']]), ignore_index=True)
                    if len(dfs.index) == 0:
                        pass
                    else:
                        dfs.columns = ['Timestamp', name[-1]]
                        dfs['Timestamp'] = pd.to_datetime(dfs['Timestamp'],format='%Y-%m-%d %H:%M:%S').dt.tz_localize(timezone('
Europe/Zurich'))
                        dfs.set_index('Timestamp', inplace = True)
                        list_of_df.append(dfs)

                if not list_of_df:
                    pass
                else:
                    df = pd.concat(list_of_df, axis=1, sort=False)
                return df
            else:
                logging.warning('REST: failure in get request.')
                logging.warning(r.json())
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)

    #=== Load datapoints from the platform with 2 minutes timestamp ===
```

```
    # This def calls several times asset_datapoints with 24h duration.
    # With longer intervals, the platform returns mean average values
    def load_datapoints(self, key, start_period, stop_period):
        tz = timezone('Europe/Zurich')
        df = pd.DataFrame()

        start_time = start_period
        dt_object = dt.datetime.strptime(start_period, '%Y-%m-%d %H:%M')

        while dt_object < dt.datetime.strptime(stop_period, '%Y-%m-%d %H:%M'):
            stop_time = (dt_object + dt.timedelta(days = 1)).strftime('%Y-%m-%d %H:%M')

            dp = self.asset_datapoints(key,str(dt_object),str(stop_time))
            try:
                df = pd.concat([df,dp])
            except:
                print('Error!')
            dt_object = (dt_object + dt.timedelta(days = 1))

        df=df[~df.index.duplicated(keep='first')][start_period:stop_period]

        return df

    '''=======================================================
    ================= SETPOINTS DEFINITIONS =================
    =======================================================   '''
    # === Send setpoints to the platform ===
    # Send setpoints value contained in payload to the different assets
    def asset_set_setpoints(self, payload):
        max_retry = self.max_retry
        retry = 0

        while (retry <= max_retry):
            self._token()
            url = 'https://eos.misurio.com/api/setpoints/single/'
            headers = {'authorization': 'Bearer ' + self.token,
                'content-type': 'application/json'}

            r = requests.post(
                url=url,
                headers=headers,
                data=json.dumps(payload),
                timeout=(10, 10))

            if r.ok:
                return r.ok
            else:
                logging.warning('REST: failure in post request.')
                if r.status_code == 400:
                    logging.warning('General Problem with the request')
                elif r.status_code == 403:
                    logging.warning('The user does not have the rights to send at least one of the setpoints')
                elif r.status_code == 404:
                    logging.warning('At least one setpoint does not exist')
                else:
                    logging.warning('Unknown error')
                time.sleep(2)
                retries += 1
                print("Connection attempt {}/{}".format(retry,max_retry))

        return pd.Series(np.nan)
```

Listing A.1 – EnergyOn.py: class to interface with the online platform

```python
import os
import numpy as np
import pandas as pd

from .helpers import *

__all__ = ["model"]

here = os.path.abspath(os.path.dirname(__file__))
DATA_FOLD = "data"
DATA_FILE = "data_hp.csv"   # files with datasheet from different heat pump and manufacturer
DATA = pd.read_csv(os.path.join(here, DATA_FOLD, DATA_FILE), index_col=0)


class Model:
    """ -------- Examples --------
    Get cold source power from the BW.A17 model with source and sink temperature respectively at 12.7 and 43.2 deg.C
    >>> from hpmodels import *
    >>> model(12.7, 43.2, "BW.A17", "Q_srce")
    """
    def __init__(self):
        pass

    def __call__(self, t_src, t_snk, model, value="COP", stage=1):
        """ Return piece-weise 2D interpolation of a choosen value for given source and sink temperatures """
        if value != "ALL":
            hp_data = DATA.loc[(DATA.Model == model) & (DATA.stage == stage)]
            assert len(hp_data) != 0, "Model={} / stage={} no found !".format(model, stage)

            t_srces = hp_data.T_srce.values
            t_sinks = hp_data.T_sink.values
            zvalues = hp_data[value].values

            dict_val = {(t1, t2): z for t1, t2, z in zip(t_srces, t_sinks, zvalues)}

            n_tsrces = sorted(set(t_srces))
            n_tsinks = sorted(set(t_sinks))

            n_values = np.zeros((len(n_tsrces), len(n_tsinks)))

            for ix, tx in enumerate(n_tsrces):
                for iy, ty in enumerate(n_tsinks):
                    n_values[ix, iy] = dict_val[(tx, ty)]

            return round(interpolate(t_src, t_snk, n_tsrces, n_tsinks, n_values), 2)
        else:
            vals = ["Q_srce", "Q_sink", "P_elec", "COP"]
            return [self(t_src, t_snk, model, value=v, stage=stage) for v in vals]

    def get_available_models(self):
        """ Return a dict mapping manufacturer with associated available models """
        manufacturers = DATA.Manufacturer.unique()
        return {manuf: DATA.loc[DATA.Manufacturer == manuf, "Model"].unique().tolist() for manuf in manufacturers}
```

Listing A.2 – hpmodel.py: Returns the electric power P, the heat from the source Q^{source}, the heat to the sink Q^{sink} and the COP based on the source and sink temperature for a specific model of heat pump. The inputs are the source and the sink temperatures.

```
import os
from scipy.interpolate import interp2d


def interpolate(x, y, nx, ny, nz):
    ''' Create grid as tuples from the look up table'''
    xitv = [(nx[i - 1], nx[i]) for i in range(1, len(nx))]
    yitv = [(ny[i - 1], ny[i]) for i in range(1, len(ny))]

    ''' Get index of the tuples containing x and y'''
    try:
        xidx = [u <= x <= v for u, v in xitv].index(True)
        yidx = [u <= y <= v for u, v in yitv].index(True)
    except ValueError:
        print('Temperatures are outside the bounds!')

    nz = nz[xidx : xidx + 2, yidx : yidx + 2].transpose()

    ''' Interpolate over a 2-D grid '''
    f = interp2d(xitv[xidx], yitv[yidx], nz, kind="linear")
    return f(x, y)[0]
```

Listing A.3 – helpers.py: reshape and interpolate over a 2-D grid the data

B Additional Figures

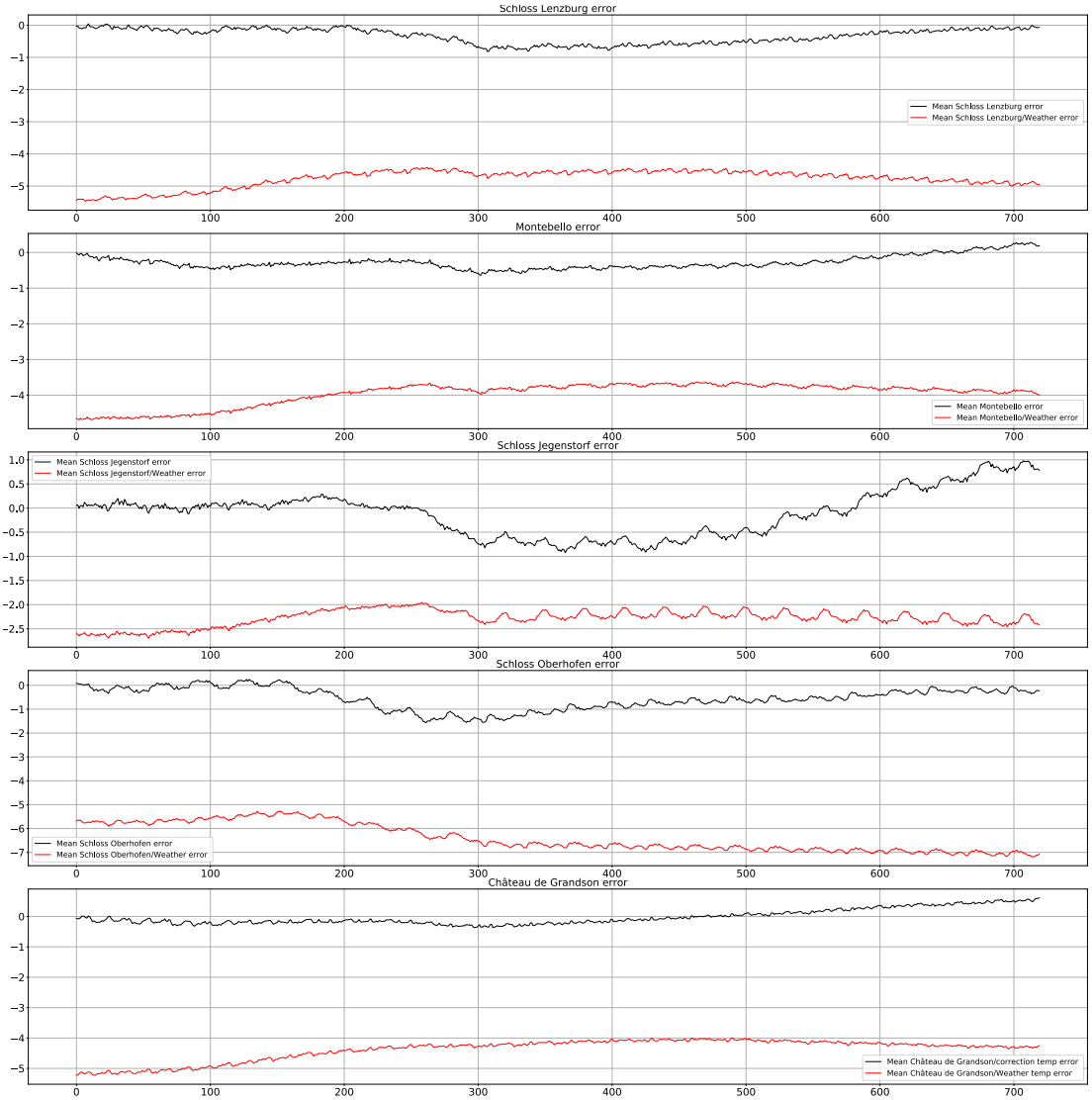
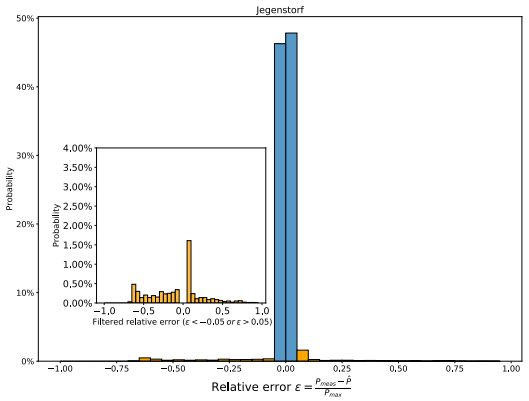
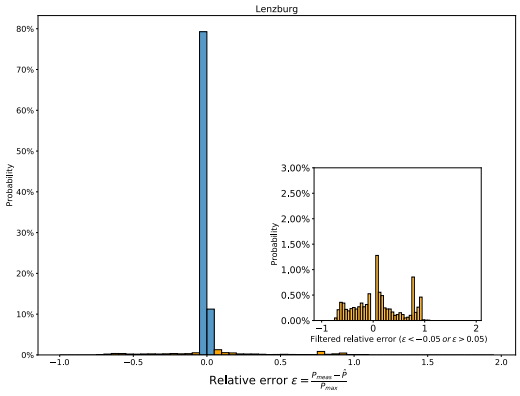


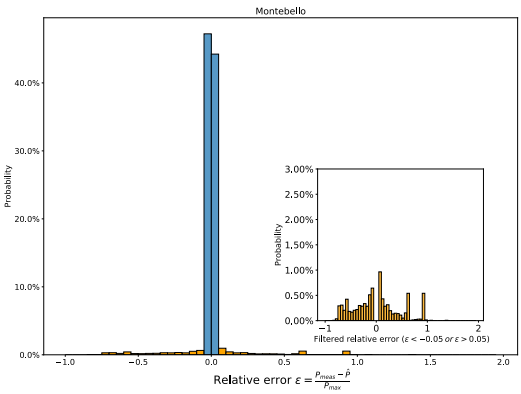
Figure B.1 – Comparison of the average error of the external temperature prediction of the online forecast with the local updated forecast using SARX models



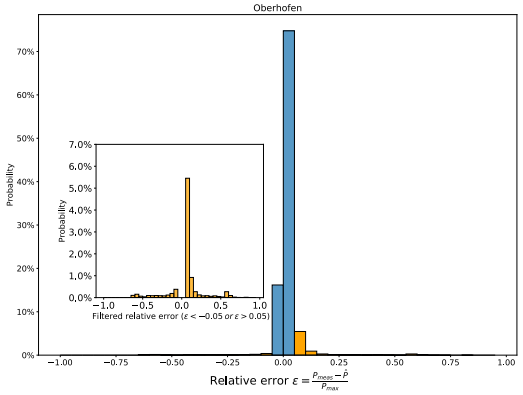
(a) Jegenstorf



(b) Lenzburg



(c) Montebello



(d) Oberhofen

Figure B.2 – Relative error distribution of the heat pump power model (based on 1) for four heat pump over 20 days. The inside plot is a zoom of the relative error when it value is outside the 5% bounds.

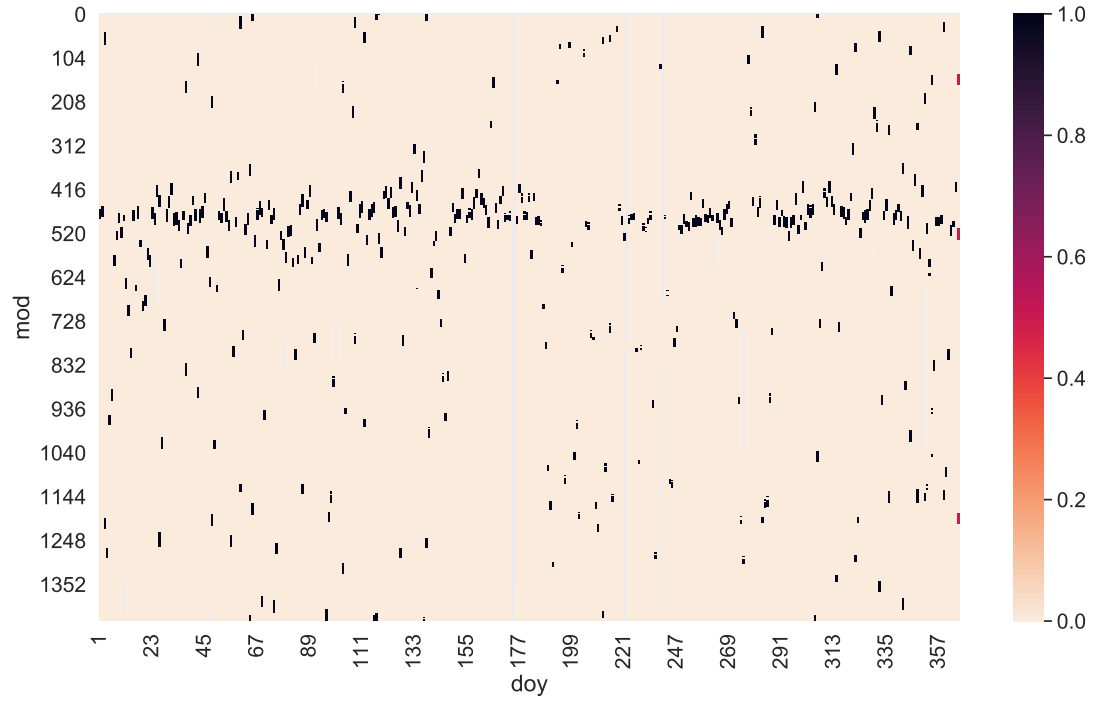


Figure B.3 – Heat map of the DHW charging cycles for Jegenstorf during a year in function of the Minute Of the Day (mod)

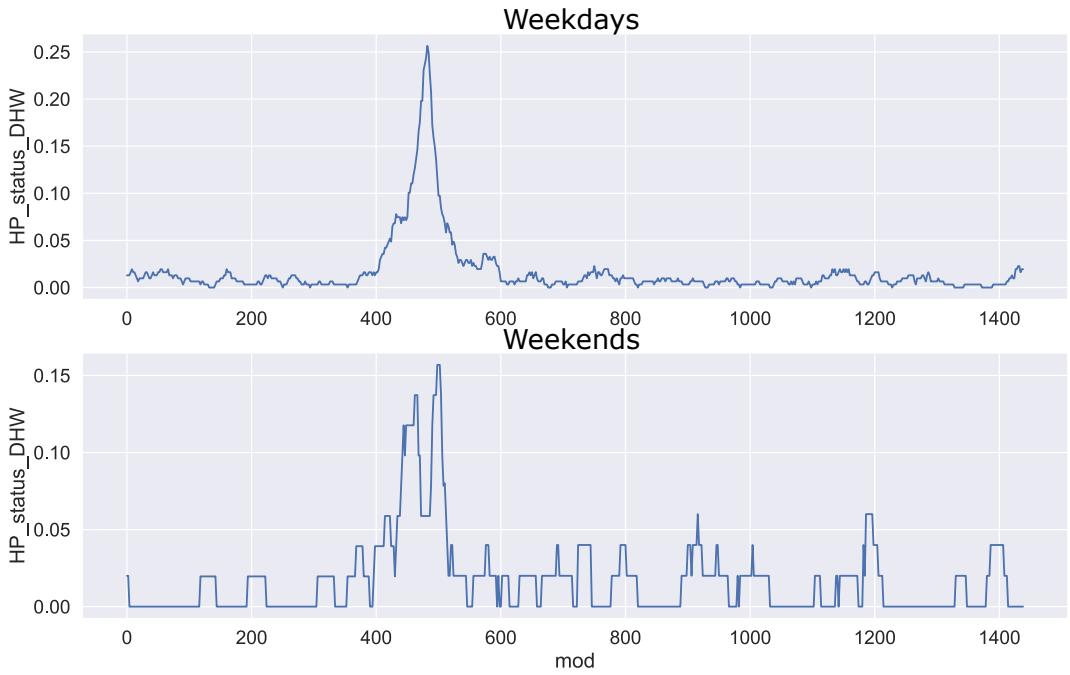


Figure B.4 – Mean activation state of a heat pump for DHW production over a year of measurement (Jegenstorf)

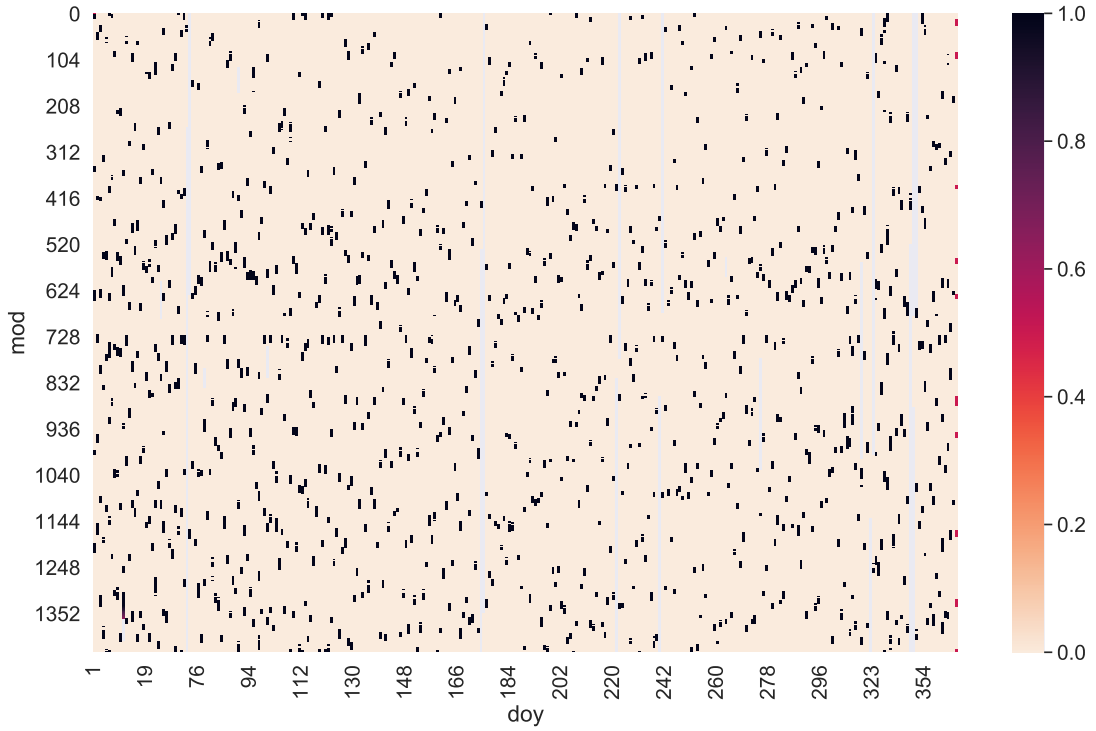


Figure B.5 – Heat map of the DHW charging cycles for Jegenstorf during a year in function of the Minute Of the Day (mod)

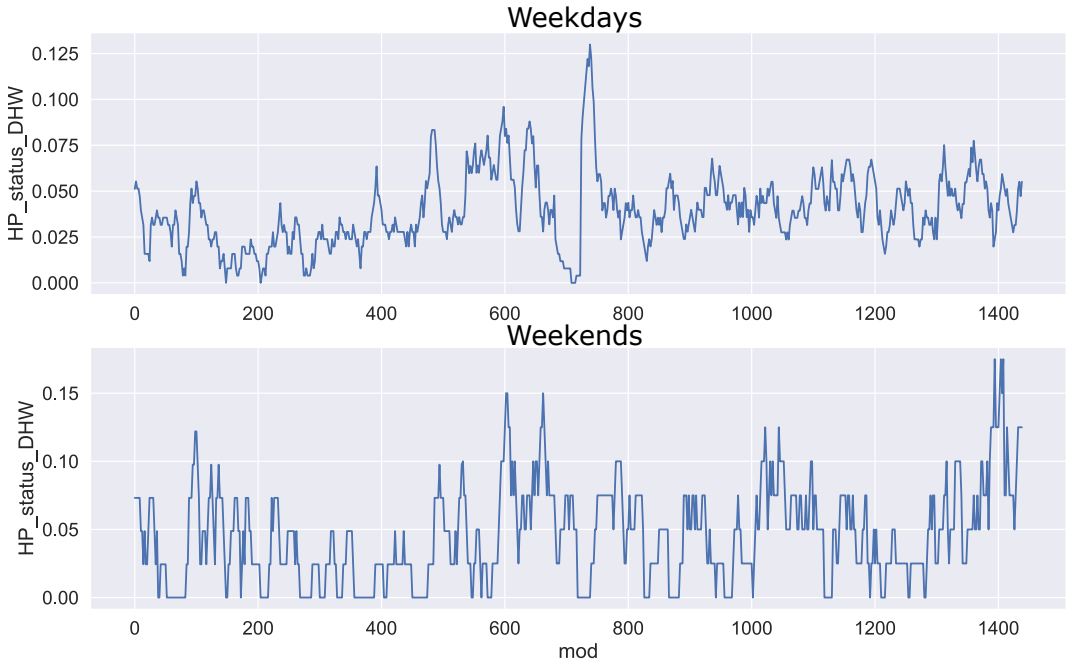


Figure B.6 – Mean activation state of a heat pump for DHW production over a year of measurement (Grandson)

Bibliography

Abrell, J. (2016). *The swiss wholesale electricity market* (tech. rep.). Swiss Competence Center for Energy Research. https://ethz.ch/content/dam/ethz/special-interest/mtec/cer-eth/economics-energy-economics-dam/documents/people/jabrell/Abrell_Swiss_Wholesale_Electricity_Market.pdf

Amblard, F., Menon, R. P., & Page, J. (2020). Flexibility characterization of a residential neighbourhood with water-to-water heat pumps using model predictive control, In *Proceedings of the 16th ibpsa conference*, IBPSA. Rome, Italy. https://doi.org/https://doi.org/10.26868/25222708.2019.210148

ASHRAE. (2019). *ASHRAE Handbook—HVAC Applications*.

Boritz, J. E. (2005). Is practitioners' views on core concepts of information integrity. *International Journal of Accounting Information Systems*, *6*(4), 260–279. https://doi.org/10.1016/j.accinf.2005.07.001

Borrelli, F., Bemporad, A., & Morari, M. (2017). *Predictive control for linear and hybrid systems* (1st). USA, Cambridge University Press.

Bradley, S. P., Hax, A. C., & Magnanti, T. L. (1977). *Applied mathematical programming*.

Cao, X., Dai, X., & Liu, J. (2016). Building energy-consumption status worldwide and the state-of-the-art technologies for zero-energy buildings during the past decade. *Energy and Buildings*, *128*, 198–213. https://doi.org/https://doi.org/10.1016/j.enbuild.2016.06.089

Cigler, J., Gyalistras, D., Siroky, J., Tiet, V.-N., & Ferkl, L. (2013). Beyond theory: the challenge of implementing model predictive control in buildings, In *Proceedings of 11th rehva world congress clima*. Prague, Czech Republic.

Cigler, J., Siroky, J., Korda, M., & Jones, C. (2013). On the selection of the most appropriate mpc problem formulation for buildings. http://infoscience.epfl.ch/record/183457

Claessens, B. J., Vrancx, P., & Ruelens, F. (2018). Convolutional neural networks for automatic state-time feature extraction in reinforcement learning applied to residential load control. *IEEE Transactions on Smart Grid*, *9*(4), 3259–3269.

Costanzo, G., Iacovella, S., Ruelens, F., Leurs, T., & Claessens, B. (2016). Experimental analysis of data-driven control for a building heating system. *Sustainable Energy, Grids and Networks*, *6*, 81–90. https://doi.org/https://doi.org/10.1016/j.segan.2016.02.002

Cui, T., Carr, J., Brissette, A., & Ragaini, E. (2017). Connecting the last mile: demand response in smart buildings [8th International Conference on Sustainability in Energy and

Buildings, SEB-16, 11-13 September 2016, Turin, Italy]. *Energy Procedia*, *111*, 720–729. https://doi.org/https://doi.org/10.1016/j.egypro.2017.03.234

De Ridder, F., Diehl, M., Mulder, G., Desmedt, J., & Van Bael, J. (2011). An optimal control algorithm for borehole thermal energy storage systems. *Energy and Buildings*, *43*(10), 2918–2925. https://doi.org/https://doi.org/10.1016/j.enbuild.2011.07.015

de Santiago, J., Rodriguez-Villalón, O., & Sicre, B. (2017). The generation of domestic hot water load profiles in swiss residential buildings through statistical predictions. *Energy and Buildings*, *141*, 341–348. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.02.045

Denholm, P., Ela, E., Kirby, B., & Milligan, M. (2010). *Role of energy storage with renewable electricity generation* (tech. rep.). National Renewable Energy Lab.(NREL), Golden, CO (United States).

D'hulst, R., Labeeuw, W., Beusen, B., Claessens, S., Deconinck, G., & Vanthournout, K. (2015). Demand response flexibility and flexibility potential of residential smart appliances: experiences from large pilot test in belgium. *Applied Energy*, *155*, 79–90. https://doi.org/https://doi.org/10.1016/j.apenergy.2015.05.101

Dias, J. M., & Costa, V. A. (2018). Adsorption heat pumps for heating applications: a review of current state, literature gaps and development challenges. *Renewable and Sustainable Energy Reviews*, *98*, 317–327. https://doi.org/https://doi.org/10.1016/j.rser.2018.09.026

Domínguez-Muñoz, F., Cejudo-López, J. M., Carrillo-Andrés, A., & Gallardo-Salazar, M. (2011). Selection of typical demand days for chp optimization. *Energy and Buildings*, *43*(11), 3036–3043. https://doi.org/https://doi.org/10.1016/j.enbuild.2011.07.024

Elsland, R., Peksen, I., & Wietschel, M. (2014). Are internal heat gains underestimated in thermal performance evaluation of buildings? [6th International Conference on Sustainability in Energy and Buildings, SEB-14]. *Energy Procedia*, *62*, 32–41. https://doi.org/https://doi.org/10.1016/j.egypro.2014.12.364

ENTSOE. (2018). *Electricity Balancing in Europe: an overview of the Europe balancing market and electricity balancing guideline*. European Network of Transmission System Operatorsfor Electricity (ENTSOE). European Copper Institute.

European parliament and council of european union regulation (eu) 2016/679 [Online accessed: 2019-12-02]. (2016).

Evins, R., Orehounig, K., Dorer, V., & Carmeliet, J. (2014). New formulations of the 'energy hub' model to address operational constraints. *Energy*, *73*, 387–398.

Fachvereinigung Wärmepumpen Schweiz. (2010-2019). Statistiken (2010-2019) [Online accessed: 2020-07-26].

Fazlollahi, S., Bungener, S. L., Mandel, P., Becker, G., & Maréchal, F. (2014). Multi-objectives, multi-period optimization of district energy systems: i. selection of typical operating periods. *Computers & Chemical Engineering*, *65*, 54–66. https://doi.org/https://doi.org/10.1016/j.compchemeng.2014.03.005

Federal Statistical Office. (2009-2015). Interactive tables: housing by canton, building category, type of heating, hot water, energy source and period of construction [Online accessed:

2020-07-26]. https://www.pxweb.bfs.admin.ch/pxweb/en/px-x-0902010000_102/px-x-0902010000_102/px-x-0902010000_102.px/?rxid=5dc7de02-9ed7-4079-af28-1bc98807e58a

Federal Statistical Office. (1990, 2000). Interactive tables: housing by canton, building category, type of heating, hot water, energy source and period of construction [Online accessed: 2020-07-26]. https://www.pxweb.bfs.admin.ch/pxweb/en/px-x-0902020100_112/px-x-0902020100_112/px-x-0902020100_112.px/?rxid=6eb9edb1-2214-40de-b591-b0d17c818d81

Fell, M. J., Shipworth, D., Huebner, G. M., & Elwell, C. A. (2015). Public acceptability of domestic demand-side response in Great Britain: The role of automation and direct load control. *Energy Research & Social Science*, *9*, 72–84. https://doi.org/10.1016/j.erss.2015.08.023

Fischer, D., Toral, T. R., Lindberg, K., Wille-Haussmann, B., & Madani, H. (2014). Investigation of thermal storage operation strategies with heat pumps in german multi family houses [Renewable Energy Research Conference, RERC 2014]. *Energy Procedia*, *58*, 137–144. https://doi.org/https://doi.org/10.1016/j.egypro.2014.10.420

Fischer, D., Lindberg, K. B., Madani, H., & Wittwer, C. (2016). Impact of pv and variable prices on optimal system sizing for heat pumps and thermal storage. *Energy and Buildings*, *128*, 723–733. https://doi.org/https://doi.org/10.1016/j.enbuild.2016.07.008

Fischer, D., & Madani, H. (2017). On heat pumps in smart grids: a review. *Renewable and Sustainable Energy Reviews*, *70*, 342–357. https://doi.org/https://doi.org/10.1016/j.rser.2016.11.182

Fischer, D., Wolf, T., & Triebel, M.-A. (2017). Flexibility of heat pump pools: the use of sg-ready from an aggregator's perspective, In *12th IEA Heat Pump Conference 2017, Rotterdam 2017*.

Fischer, D., Wolf, T., Wapler, J., Hollinger, R., & Madani, H. (2017). Model-based flexibility assessment of a residential heat pump pool. *Energy*, *118*, 853–864. https://doi.org/https://doi.org/10.1016/j.energy.2016.10.111

Florita, A. R., & Henze, G. P. (2009). Comparison of short-term weather forecasting models for model predictive control. *HVAC&R Research*, *15*(5), 835–853. https://doi.org/10.1080/10789669.2009.10390868

Foucquier, A., Robert, S., Suard, F., Stéphan, L., & Jay, A. (2013). State of the art in building modelling and energy performances prediction: a review. *Renewable and Sustainable Energy Reviews*, *23*, 272–288. https://doi.org/https://doi.org/10.1016/j.rser.2013.03.004

Fuentes, E., Arce, L., & Salom, J. (2018). A review of domestic hot water consumption profiles for application in systems and buildings energy performance analysis. *Renewable and Sustainable Energy Reviews*, *81*, 1530–1547. https://doi.org/https://doi.org/10.1016/j.rser.2017.05.229

Galus, M. D. (2017). Smart grid roadmap and regulation approaches in switzerland. *CIRED - Open Access Proceedings Journal*, *2017*(1), 2906–2909.

Geidl, M., Koeppel, G., Favre-Perrod, P., Klockl, B., Andersson, G., & Frohlich, K. (2006). Energy hubs for the future. *IEEE power and energy magazine*, *5*(1), 24–30.

Giani, A., Bitar, E., Garcia, M., McQueen, M., Khargonekar, P., & Poolla, K. (2013). Smart grid data integrity attacks. *IEEE Transactions on Smart Grid*, *4*(3), 1244–1253.

Giodano, V., Meletiou, A., Covrig, C., Mengolini, A., Ardelean, M., Fulli, G., Jiménez, M., & Filiou, C. (2013). *Smart grid projects in europe: lessons learned and current developments 2012 update*. Joint Research Center. Publications Office of the European Union.

Girardin, L. (2012). A gis-based methodology for the evaluation of integrated energy systems in urban area, 218. https://doi.org/10.5075/epfl-thesis-5287

Gkatzikis, L., Koutsopoulos, I., & Salonidis, T. (2013). The role of aggregators in smart grid demand response markets. *IEEE Journal on Selected Areas in Communications*, *31*(7), 1247–1257.

Goflex. (2016-2020). https://www.goflex-project.eu

Gorecki, T. T., Fabietti, L., Qureshi, F. A., & Jones, C. N. (2017). Experimental demonstration of buildings providing frequency regulation services in the swiss market. *Energy and Buildings*, *144*, 229–240. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.02.050

Grundfos. (2013). *Alpha2: installation and operating instructions*. Version Document number: 95047457. Grundfos.

Gurobi Optimization, L. (2018). Gurobi optimizer reference manual. http://www.gurobi.com

Haes Alhelou, H., Hamedani Golshan, M. E., Njenda, T., & Siano, P. (2019). A survey on power system blackout and cascading events: research motivations and challenges. *Energies*, *12*. https://doi.org/10.3390/en12040682

Halvgaard, R. (2014). Model predictive control for smart energy systems, 182. https://backend.orbit.dtu.dk/ws/portalfiles/portal/97868333/Model_Predictive_Control.pdf

Haves, P., & Xu, P. (2007). The building controls virtual test bed a simulation environment for developing and testing control algorithms, strategies and systems, In *Proc. of the 10-th ibpsa conference*.

Hernández Hernández, C., Rodriguez, F., Moreno, J. C., Mendes, P., Normey-Rico, J., & Guzmán, J. (2017). The comparison study of short-term prediction methods to enhance the model predictive controller applied to microgrid energy management. *Energies*, *10*, 884. https://doi.org/10.3390/en10070884

Jenkins, D., Liu, Y., & Peacock, A. (2008). Climatic and internal factors affecting future uk office heating and cooling energy consumptions. *Energy and Buildings*, *40*(5), 874–881. https://doi.org/https://doi.org/10.1016/j.enbuild.2007.06.006

Jensen, S. Ø., Marszal-Pomianowska, A., Lollini, R., Pasut, W., Knotzer, A., Engelmann, P., Stafford, A., & Reynders, G. (2017). Iea ebc annex 67 energy flexible buildings. *Energy and Buildings*, *155*, 25–34. https://doi.org/https://doi.org/10.1016/j.enbuild.2017.08.044

Jünger, M., Liebling, T. M., Naddef, D., Nemhauser, G. L., Pulleyblank, W. R., Reinelt, G., Rinaldi, G., & Wolsey, L. A. (2009). *50 years of integer programming 1958-2008: from the early years to the state-of-the-art*. Springer Science & Business Media.

Junker, R. G., Azar, A. G., Lopes, R. A., Lindberg, K. B., Reynders, G., Relan, R., & Madsen, H. (2018). Characterizing the energy flexibility of buildings and districts. *Applied Energy*, *225*, 175–182. https://doi.org/https://doi.org/10.1016/j.apenergy.2018.05.037

Killian, M., & Kozek, M. (2016). Ten questions concerning model predictive control for energy efficient buildings. *Building and Environment*, *105*, 403–412. https://doi.org/https://doi.org/10.1016/j.buildenv.2016.05.034

Kim, Y.-J., Norford, L. K., & Kirtley, J. L. (2014). Modeling and analysis of a variable speed heat pump for frequency regulation through direct load control. *IEEE Transactions on Power Systems*, *30*(1), 397–408.

Kirschen, D. S., Rosso, A., Ma, J., & Ochoa, L. F. (2012). Flexibility from the demand side, In *2012 ieee power and energy society general meeting*.

Kohlhepp, P., Harb, H., Wolisz, H., Waczowicz, S., Müller, D., & Hagenmeyer, V. (2019). Large-scale grid integration of residential thermal energy storages as demand-side flexibility resource: a review of international field studies. *Renewable and Sustainable Energy Reviews*, *101*, 527–547. https://doi.org/https://doi.org/10.1016/j.rser.2018.09.045

Kok, K., & Widergren, S. (2016). A society of devices: integrating intelligent distributed resources with transactive energy. *IEEE Power and Energy Magazine*, *14*(3), 34–45.

Lasseter, R. H., & Paigi, P. (2004). Microgrid: a conceptual solution, In *2004 ieee 35th annual power electronics specialists conference (ieee cat. no.04ch37551)*.

Lawrence, T. M., Boudreau, M.-C., Helsen, L., Henze, G., Mohammadpour, J., Noonan, D., Patteeuw, D., Pless, S., & Watson, R. T. (2016). Ten questions concerning integrating smart buildings into the smart grid. *Building and Environment*, *108*, 273–283. https://doi.org/https://doi.org/10.1016/j.buildenv.2016.08.022

Li, K., Wang, B., Wang, Z., Wang, F., Mi, Z., & Zhen, Z. (2017). A baseline load estimation approach for residential customer based on load pattern clustering [Proceedings of the 9th International Conference on Applied Energy]. *Energy Procedia*, *142*, 2042–2049. https://doi.org/https://doi.org/10.1016/j.egypro.2017.12.408

Lund, P. D., Lindgren, J., Mikkola, J., & Salpakari, J. (2015). Review of energy system flexibility measures to enable high levels of variable renewable electricity. *Renewable and Sustainable Energy Reviews*, *45*, 785–807. https://doi.org/https://doi.org/10.1016/j.rser.2015.01.057

Ma, Z., Billanes, J. D., & Jørgensen, B. N. (2017). Aggregation potentials for buildings—business models of demand response and virtual power plants. *Energies*, *10*(10), 1646.

Maciejowski, J. M. (2002). *Predictive control: with constraints*. Pearson education.

Madsen, H., Halvgaard, R., Poulsen, N., & Jorgensen, J. (2012). Economic model predictive control for building climate control in a smart grid, In *Innovative smart grid technologies, ieee pes*, Los Alamitos, CA, USA, IEEE Computer Society. https://doi.org/10.1109/ISGT.2012.6175631

Mathieu, J. L., Price, P. N., Kiliccote, S., & Piette, M. A. (2011). Quantifying changes in building electricity use, with application to demand response. *IEEE Transactions on Smart Grid*, *2*(3), 507–518.

Mayne, D. Q. (2014). Model predictive control: recent developments and future promise. *Automatica*, *50*(12), 2967–2986. https://doi.org/https://doi.org/10.1016/j.automatica.2014.10.128

Menon, R. P., Amblard, F., & Page, J. (2021). *Roadblocks to heat pump use in residential buildings with model predictive control for demand response and ancillary markets and potential solutions* [To be presented at the rescheduled IEA Heat Pump Conference 2020].

Motegi, N., Piette, M. A., Watson, D. S., Kiliccote, S., & Xu, P. (2007). *Introduction to commercial building control strategies and techniques for demand response–appendices* (tech. rep.). Lawrence Berkeley National Lab.(LBNL), Berkeley, CA (United States).

Neurobat [Accessed: 2020-09-30]. (2014-2015). https://www.energie-cluster.ch/admin/data/files/file/file/1994/02_stephen-neff_neurobat.pdf?lm=1509106879

Novak, T. (2018). *Heat pumps, integrating technologies to decarbonise heating and cooling*. European Copper Institute. European Copper Institute.

Oldewurtel, F., Parisio, A., Jones, C. N., Gyalistras, D., Gwerder, M., Stauch, V., Lehmann, B., & Morari, M. (2012). Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, *45*, 15–27. https://doi.org/https://doi.org/10.1016/j.enbuild.2011.09.022

Peterhans, S., Dott, R., Eugster, W., Freymond, M., Milton, G., Georges, G., Hubacher, P., & Muller, C. (2019). *Rapport annuel 2019 (gsp)*. Groupement professionnel suisse pour les pompes a chaleur GSP.

Posma, J., Lampropoulos, I., Schram, W., & van Sark, W. (2019). Provision of ancillary services from an aggregated portfolio of residential heat pumps on the dutch frequency containment reserve market. *Applied Sciences*, *9*(3), 590.

Raju P, E., & Jain, T. (2019). Chapter 2 - distributed energy resources and control (R. K. Chauhan & K. Chauhan, Eds.). In R. K. Chauhan & K. Chauhan (Eds.), *Distributed energy resources in microgrids*. Academic Press. https://doi.org/https://doi.org/10.1016/B978-0-12-817774-7.00002-8

Remund, J. (2008). Quality of meteonorm version 6.0.

Reynders, G., Nuytten, T., & Saelens, D. (2013). Potential of structural thermal mass for demand-side management in dwellings. *Building and Environment*, *64*, 187–199. https://doi.org/https://doi.org/10.1016/j.buildenv.2013.03.010

Reynders, G., Amaral Lopes, R., Marszal-Pomianowska, A., Aelenei, D., Martins, J., & Saelens, D. (2018). Energy flexible buildings: an evaluation of definitions and quantification methodologies applied to thermal storage. *Energy and Buildings*, *166*, 372–390. https://doi.org/https://doi.org/10.1016/j.enbuild.2018.02.040

Reynders, G., & Wetter, M. (2015). *Quantifying the impact of building design on the potential of structural storage for active demand response in resi-dential buildings* (tech. rep.). https://lirias.kuleuven.be/1729315

Salpakari, J., & Lund, P. (2016). Optimal and rule-based control strategies for energy flexibility in buildings with pv. *Applied Energy*, *161*, 425–436. https://doi.org/https://doi.org/10.1016/j.apenergy.2015.10.036

Scattolini, R. (2009). Architectures for distributed and hierarchical Model Predictive Control – A review. *Journal of Process Control*, *19*(5), 723–731. https://doi.org/10.1016/j.jprocont.2009.02.003

Scherer, H. F., Pasamontes, M., Guzmán, J. L., Álvarez, J. D., Camponogara, E., & Normey-Rico, J. E. (2014). Efficient building energy management using distributed model predictive control. *Journal of Process Control*, *24*(6), 740–749. https://doi.org/10.1016/j.jprocont.2013.09.024

Schütz, T., Harb, H., Streblow, R., & Müller, D. (2015). Comparison of models for thermal energy storage units and heat pumps in mixed integer linear programming, In *Proceedings from the 28th international conference on efficiency, cost, optimization, simulation and environmental impact of energy systems, ecos 2015, 29.06.2015-03.07.2015, pau, france*. https://publications.rwth-aachen.de/record/540090

Schütz, T., Schraven, M. H., Remy, S., Granacher, J., Kemetmüller, D., Fuchs, M., & Müller, D. (2017). Optimal design of energy conversion units for residential buildings considering german market conditions. *Energy*, *139*, 895–915. https://doi.org/https://doi.org/10.1016/j.energy.2017.08.024

Schütz, T., Streblow, R., & Müller, D. (2015). A comparison of thermal energy storage models for building energy system optimization. *Energy and Buildings*, *93*, 23–31. https://doi.org/https://doi.org/10.1016/j.enbuild.2015.02.031

Seabold, S., & Perktold, J. (2010). Statsmodels: econometric and statistical modeling with python, In *9th python in science conference*.

Serale, G., Fiorentini, M., Capozzoli, A., Bernardini, D., & Bemporad, A. (2018). Model predictive control (mpc) for enhancing building and hvac system energy efficiency: problem formulation, applications and opportunities. *Energies*, *11*(3), 631.

Sharifi, R., Fathi, S., & Vahidinasab, V. (2016). Customer baseline load models for residential sector in a smart-grid environment. *Energy Reports*, *2*, 74–81. https://doi.org/https://doi.org/10.1016/j.egyr.2016.04.003

Sim4blocks. (2016-2020). https://sim4blocks.eu

Steen, D., Stadler, M., Cardoso, G., Groissböck, M., DeForest, N., & Marnay, C. (2015). Modeling of thermal storage systems in milp distributed energy resource models. *Applied Energy*, *137*, 782–792.

Vale, Z., Pinto, T., Morais, H., Praça, I., & Faria, P. (2011). Vpp's multi-level negotiation in smart grids and competitive electricity markets, In *2011 ieee power and energy society general meeting*.

van der Heijde, B., Sourbron, M., Vega Arance, F., Salenbien, R., & Helsen, L. (2017). Unlocking flexibility by exploiting the thermal capacity of concrete core activation [11th International Renewable Energy Storage Conference, IRES 2017, 14-16 March 2017, DÃ¼sseldorf, Germany]. *Energy Procedia*, *135*, 92–104. https://doi.org/https://doi.org/10.1016/j.egypro.2017.09.490

Vrettos, E., Lai, K., Oldewurtel, F., & Andersson, G. (2013). Predictive control of buildings for demand response with dynamic day-ahead and real-time prices, In *2013 european control conference (ecc)*.

Wang, A., Li, R., & You, S. (2018). Development of a data driven approach to explore the energy flexibility potential of building clusters. *Applied Energy*, *232*, 89–100. https://doi.org/https://doi.org/10.1016/j.apenergy.2018.09.187

Weitemeyer, S., Kleinhans, D., Vogt, T., & Agert, C. (2015). Integration of renewable energy sources in future power systems: the role of storage. *Renewable Energy*, *75*, 14–20. https://doi.org/https://doi.org/10.1016/j.renene.2014.09.028

Wemhöner, C., Hafner, B., & Schwarzer, K. (2000). Simulation of solar thermal systems with carnot blockset in the environment matlab® simulink®, In *Proceedings from eurosun2000: ises europe solar congress*.

Wetter, M. (2011). Co-simulation of building energy and control systems with the building controls virtual test bed. *Journal of Building Performance Simulation*, *4*(3), 185–203.

Wijbenga, J. P., MacDougall, P., Kamphuis, R., Sanberg, T., van den Noort, A., & Klaassen, E. (2014). Multi-goal optimization in powermatching city: a smart living lab, In *Ieee pes innovative smart grid technologies, europe*.

Zhang, Y., & Hanby, V. I. (2007). Short-term prediction of weather parameters using online weather forecasts.

Frédéric Amblard

Personal Information

Address	Av. Louis-Ruchonnet 2bis 1003 Lausanne frederic.amblard@epfl.ch
Date of Birth	Feb. 23th 1991
Nationality	French

Professional Experience

May 2016 – Present **PhD scholar at the HES Sion and EPFL**
- *Software development in python:* implementing prediction models
- *Data analysis:* assessing heating and DHW profiles (3 years of monitoring data)
- *Student supervision and training:* supervising undergrad projects

Mar 2015 – Jun 2015 **Research Assistant at IPESE Laboratory, EPFL, Switzerland**
- *Software development:* Development of the lab computational platform OSMOSE
- *Data analysis:* project GEOTHERM II, Uncertainty Analysis and Robust Optimisation

Mar 2014 – Sep 2014 **Engineering internship**, Centre de recherche Observation, Impacts, Energie (OIE), MINES ParisTech, Sophia Antipolis
- *Research:* Realisation of the task III of EVALGTHDOM project. Life Cycle Assessment of prospective scenario of the energetic mix in Guadeloupe island (FR) at horizon 2035.
- *Results:* In collaboration with governmental and private partners, we defined a quantification methodology that will serve as a based for future projects.

Jun - Aug 2011, 2012 **Laboratory Assistant**, EFLUM at EPFL
- *Research:* using a lysimeter to study the fluid mechanics in the ground. Set up of an experiment with an optic fiber buried in the lysimeter and field monitoring with a Distributed Temperature Sensing (DTS)

Education

Sep 2015 – present **PhD thesis in Energy** with François Maréchal, Department of Engineering science and technology (STI), IPESE Laboratory, EPFL, Switzerland and with Jessen Page, Institute of Sustainable Energy, HES-SO Valais Wallis, Switzerland. **Title:** *"Reliability and Performance of Model Predictive Control for Demand Response with Residential Heat Pumps"*

Sep 2013 – Feb 2015 **Master studies in Environmental engineering, Minor in Energy**, EPFL, Switzerland. Thesis with François Maréchal, IPESE Laboratory, **Title:** *"Geothermal energy integration in urban system – The case study of the city of Lausanne"*

Sep 2010 – Jun 2013 **Bachelor studies in Environmental Engineering**, EPFL, Switzerland

Aug 2003 – Sep 2009 **Baccalauréat Scientifique** Lycée Georges Clemenceau, Montpellier, mention Très Bien

Aug 2003 – Sep 2009 **Allgemeine Hochschulreife**, Lycée Georges Clemenceau, Montpellier, mention Très Bien

Technical skills

Language	French (native), English (C1), German (C1)
Programming	Matlab, python, R, SQL, Ampl
OS & Tools	Linux, MS Windows, Git, Latex, Maninfold / UML editor, Ecoinvent / Simapro

Activities and interests

2011 – 2014 **Committee member for Japan Impact**: A yearly convention about Japanese culture organized at EPFL (2 days, 4500 visitors)

Publications

2021
- Menon R. P, **Amblard F**, & Page J (2021) *Roadblocks to heat pump use in residential buildings with model predictive control for demand response and ancillary markets and potential solutions* [To be presented at the rescheduled IEA Heat Pump Conference 2020]

2020
- Brès A, **Amblard F**, & Hauer H (2020) *Building simulation for control tuning: accounting for model usefulness in calibration metrics choice*, Proceedings of the BauSIM 2020 conference

2019
- **Amblard F**, Menon R. P, & Page J (2020) *Flexibility characterization of a residential neighbourhood with water-to-water heat pumps using model predictive control*, In Proceedings of the 16$^{\text{th}}$ BS2019 conference, IBPSA. Rome, Italy.
- Brès A, **Amblard F**, Page J, Hauer S, Shadrina A (2020) *Now it looks more real - A Study of Metrics and Resolution for the Calibration of Dynamic Simulation*, In Proceedings of the 16$^{\text{th}}$ BS2019 conference, IBPSA. Rome, Italy.

2015
- Moret A, Gerber L, **Amblard F**, Peduzzi E & Maréchal F (2015) *Building simulation for control tuning: accounting for model usefulness in calibration metrics choice*, Conference: 40$^{\text{th}}$ Workshop on Geothermal Reservoir Engineering, Stanford, CA, USA