

Journal Pre-proofs

Operation and energy flexibility evaluation of direct load controlled buildings equipped with heat pumps

Gerard Mor, Jordi Cipriano, Benedetto Grillone, Frédéric Amblard, Ramanunni Parakkal Menon, Jessen Page, Marcus Brennenstuhl, Dirk Pietruschka, Ruben Baumer, Ursula Eicker

PII: S0378-7788(21)00768-4
DOI: <https://doi.org/10.1016/j.enbuild.2021.111484>
Reference: ENB 111484

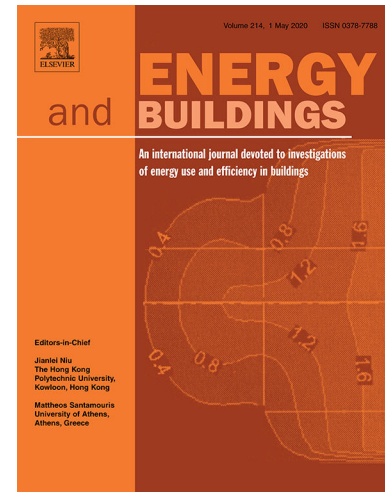
To appear in: *Energy & Buildings*

Received Date: 3 August 2021
Accepted Date: 16 September 2021

Please cite this article as: G. Mor, J. Cipriano, B. Grillone, F. Amblard, R.P. Menon, J. Page, M. Brennenstuhl, D. Pietruschka, R. Baumer, U. Eicker, Operation and energy flexibility evaluation of direct load controlled buildings equipped with heat pumps, *Energy & Buildings* (2021), doi: <https://doi.org/10.1016/j.enbuild.2021.111484>

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2021 Published by Elsevier B.V.



Operation and energy flexibility evaluation of direct load controlled buildings equipped with heat pumps

Gerard Mor^a, Jordi Cipriano^{a,*}, Benedetto Grillone^b, Frédéric Amblard^c, Ramanunni Parakkal Menon^d, Jessen Page^c, Marcus Brennenstuhl^e, Dirk Pietruschka^e, Ruben Bäumer^f, Ursula Eicker^d

^aCentre Internacional de Mètodes Numèrics a l'Enginyeria. Building Energy and Environment Group. Pere de Cabrera 16. Office 2G. 25001. Lleida. Spain

^bCentre Internacional de Mètodes Numèrics a l'Enginyeria. Building Energy and Environment Group. Rbla. Sant Nebridi 22, 08222 Terrassa, Spain

^cHes-So Valais-Wallis. Route du Rawyl 47. 1950 Sion, Switzerland

^dNext Generation Cities Institute, Concordia University, 1455 Boulevard de Maisonneuve O., EV-6.111 Montréal H3G 1M8, Canada

^eStuttgart University of Applied Sciences, Schellingstraße 24, 70174 Stuttgart, Germany

^fCentrica Business Solutions Belgium NV. Roderveldlaan 2/b2, 2600 Antwerp, Belgium

Abstract

To date, the assessment of the energy flexibility to be delivered by existing buildings and by their legacy HVAC systems is hindered by a lack of commonly agreed-upon methodologies. There are many research works in the field; however, many of them are focused on the design stage or, in case of addressing building operation, they are based on controlled experimental set-ups. The novelty of this paper lies in the fact that it develops and validates an original methodology for the Flexibility Function estimation to evaluate the delivered energy flexibility of several Automated Demand Response services applied on different heat pump systems working under real operations. The active interaction with several electricity markets, ranging from the Spanish day-ahead market to the German and Swiss ancillary services markets, have also been evaluated during the winter and spring seasons. The method results showed that heat pumps could offer a significant potential of flexibility in the analysed countries. Nevertheless, it has also been envisaged that some restrictions concerning reaction times and reliability may affect its readiness for certain ancillary services markets.

Keywords: demand response, HVAC, flexibility, model predictive control

*Corresponding author

Email addresses: gmor@cimne.upc.edu (Gerard Mor), cipriano@cimne.upc.edu (Jordi Cipriano), bgrillone@cimne.upc.edu (Benedetto Grillone), frederic.amblard@hevs.ch (Frédéric Amblard), ramanunniarakkal.menon@concordia.ca (Ramanunni Parakkal Menon), jessen.page@hevs.ch (Jessen Page), marcus.brennenstuhl@hft-stuttgart.de (Marcus Brennenstuhl), dirk.pietruschka@hft-stuttgart.de (Dirk Pietruschka), Ruben.Baumer@centrica.com (Ruben Bäumer), ursula.eicker@concordia.ca (Ursula Eicker)

Nomenclature

Acronyms

aFRR automatic Frequency Restoration Reserve Market

AR AutoRegressive

ARX AutoRegressive with eXogenous

ASHRAE American Society of Heating, Refrigerating and Air-Conditioning Engineers

BaU Business as Usual

BES Building Energy Simulation

CM Cluster Manager

CVRMSE Coefficient of Variation of the Root Mean Squared Error

DA Day-Ahead Electricity Price

DER Distributed Energy Resource

DHW Domestic Hot Water

DR Demand Response

DSO Distribution System Operator

ECM Energy Conservation Measures

EU European Union

FCR Frequency Containment Reserve

FF Flexibility Function

HP Heat Pump

IoT Internet of Things

MAPE Mean Absolute Percentage Error

mFRR Manual Frequency Restoration Reserve

MPC Model Predictive Control

NEMO Nominated Electricity Market Operator

NLME Non Linear Mixed Effect

RMSE Root Mean Squared Error
RR Reserve Replacement
SAR Seasonal Auto Regressive
SH Space Heating
SVM Support Vector Machines
TS Time Series
TSO Transmission System Operator

Subscripts and superscripts

b baseline
bd building number
e active
f trace to be tracked
o outdoor
opt optimized
t time t

Variables

A Percentage of activation time within a time step
B Backward shift operator
i flexibility evaluation period
n number of time steps
P Power
RC Resistor and Capacitator
T Temperature

1. INTRODUCTION

Renewable energy sources like solar panels and wind turbines are invaluable for transitioning to a fossil-free energy system to mitigate climate change impacts. However, their natural fluctuations introduce significant uncertainty in the power grid. In addition, they transform the present unidirectional centralized system into a bi-directional decentralized system with smaller units and multiple prosumers, increasing the difficulty to achieve power balance [1]. This leads to an increased need for flexibility on the demand side [2, 3] and for new storage capacity [4, 5]. One attractive solution identified to support the transition of power systems is to manage not only the energy supply but also the demand via Demand Response (DR) programs [6, 7]. The principle behind it is to use various economic incentives to shift the electrical loads of end-use customers from times with a high wholesale market price or when the system's security is threatened to other time periods. As has been pointed out in [8], there are predominantly two types of DR programs: i) explicit DR (also called incentive-based); ii) and implicit DR (also called price-based). In Implicit DR, a price signal is sent to the prosumers to motivate their user behaviour change. Explicit DR involves the participation of a third party, who takes action on behalf of a customer by sending an activation signal such that the system behaviours are directly modified. In both DR programs, and considering that nearly 50 % of the total energy consumption of buildings comes from Space Heating (SH)/Cooling (SC) and domestic hot water (DHW), as stated by [9], there is definitely a role that electrically driven Heating, Ventilation and Air Conditioning (HVAC) systems can play.

Although the installation of control devices, communication, control protocols and standardization have improved, DR is currently still rarely implemented in the commercial and even less in the residential sector in Europe [10]. Serale et.al [11] reviewed 161 papers on Model Predictive Control (MPC) in buildings, and revealed that only a fourth considered residential buildings and only a bit more than a fifth compared experimental cases to simulated cases. Kohlhepp et.al. [12] performed a thorough review of 16 projects of field tests and demonstrations of applied DR from around the world. Only four projects had more than 100 households, a size large enough to represent load diversity and test resource competition. A singular case of commercially applied DR to large scale residential buildings is run by the French company Voltalis, which manages one of the biggest portfolios of explicit DR services in the world. They follow a strategy of DR based on service curves [13, 14]. To our knowledge, they have not published peer-reviewed papers analysing the impacts of this DR strategy or provided a general methodology to evaluate the delivered energy flexibility. In general, there is a lack of test case benchmarks. Comparing the results among case studies with different goals, addressed electricity markets and technology environments is still very challenging.

The few real case applications of DR have brought forth a wide diversity of methodologies to evaluate the energy flexibility that individual or clustered buildings can provide. In many cases, assessment methodologies are focused on the potential energy flexibility at the building design stage. Arteconit et. al [15] is a clear example of defining an indicator of flexibility labelling at the de-

48 sign stage. Finck et. al [16] performed a very detailed analysis of the demand
49 flexibility that power-to-heat systems can deliver. Several flexibility indicators
50 such as available storage capacity and efficiency are enhanced with a flexibil-
51 ity factor, which relates electricity costs in the lower price and higher price
52 periods in a day-ahead electricity market DR scenario. A thermal instanta-
53 neous power flexibility indicator is also described. These indicators have a
54 great potential to evaluate the energy flexibility in DR services addressing the
55 ancillary markets. The only weakness is that they were demonstrated in a the-
56 oretical simulated environment. Moreover, that research was more focused on
57 developing control strategies and not on the flexibility evaluation itself.

58 In their hands-on review, Reynders et.al [17] made a valuable contribution
59 in reviewing prior research dealing with definitions and quantification of en-
60 ergy flexibility. One of their main conclusions was that a large share of the
61 performed research practices did not explicitly define or were not focused on
62 quantifying energy flexibility. Yet, they dealt with the development of con-
63 trol strategies and algorithms for specific case studies. They also stated that
64 most of the studies had in common the identification of three general proper-
65 ties of energy flexibility: i) the potential flexibility in several time horizons; ii)
66 the load which can be shifted; and iii) the cost of this flexibility. The authors
67 also deduced that methodologies aimed at quantifying the energy flexibility
68 by analyzing triggered events at specific times have greater strengths when
69 dealing with the flexibility to be delivered by the thermal mass of buildings
70 or energy storage systems. In contrast, methodologies which relied on dif-
71 ferences in the accumulated energy profiles are difficult to interpret because
72 they treat systems driven by multiple time constants as a single state system.
73 El Geneidy and Howard [18] performed a detailed analysis of the categories
74 of characteristics that constrain the contracted flexibility potential in homes.
75 Although their results are valuable for defining further DR strategies, they are
76 limited by simplified assumptions and exclusively based on simulated scenar-
77 ios. Bampoulas et.al [19] conducted a more detailed recent review on studies
78 aiming at defining suitable flexibility indicators. They highlighted that most
79 of these studies were limited to evaluating control strategies and assessing the
80 activating and deactivating of the building's thermal mass. Still, they did not
81 clearly quantify the flexibility potential of HVAC systems.

82 Following these remarks, Junkers et.al [20] developed a novel methodol-
83 ogy to characterize the energy flexibility as a dynamic function named the
84 Flexibility Function (FF). This FF enables a Flexibility Index, which describes
85 how a building can respond to certain activation signals. The FF is a step-
86 response function that assumes that the relation between the penalty signal
87 and the power load is linear and time-invariant. Several theoretical cases were
88 presented to validate this proposed FF , demonstrating how the FF enables the
89 quantification of the energy flexibility in different types of buildings. This pa-
90 per represents a valuable contribution to the field since it establishes a robust
91 methodology to represent, in a normalized manner, the correlation between
92 the penalty signal and the load response. The concept of the FF applies to
93 several building typologies and DR scenarios but specifically addresses im-
94 plicit DR services. However, the assumption that the dependence of the active

95 power and the activation variable is linear limits its applicability to DR ser-
96 vices which can fulfil this requirement. Recently, Junkers et.al [21] published
97 a paper presenting a new generic method capable of overcoming the linearity
98 and time dependency of the correlation between the flexibility and the penalty
99 signal. This new method follows the principles of the *FF*, but it changes the
100 perspective. They developed a non-linear dynamic model based on stochas-
101 tic differential equations. It is applied to price-based controlled buildings and
102 water towers, showing high robustness, accuracy and scalability to similar
103 business cases. One limitation is that these methods are developed to specifi-
104 cally address implicit DR services driven by penalty signals triggered by one
105 of the stakeholders of the electricity sector. This is very common in many
106 electricity markets, such as the spot electricity market, the intra-day market,
107 or certain ancillary services markets. However, in some explicit DR services,
108 where the activation variable is a power trace to be followed, such as when
109 a commercial aggregator makes bilateral agreements with their Balance Re-
110 sponsive Parties (BRP), both the *FF* and the flexibility characterization model
111 defined in [20, 21] need to be modified or extended to adapt them to these
112 different kinds of activation variables.

113 In our research, an extension of the previously developed flexibility charac-
114 terization procedures is performed, which is the main novelty of the research
115 work. Based on the background knowledge developed by Junkers et.al [20],
116 and further improved in [21], new linear regression-based models, designed
117 to characterize the energy flexibility delivered by blocks of buildings, are de-
118 veloped and validated in real cases. These new flexibility models address
119 different implicit and explicit DR scenarios. For example, the activation vari-
120 able can be the spot market price, the percentage of power to be activated, or a
121 power trace to be tracked. This is also an extra contribution to the paper. One
122 last novelty of the research lies in the fact of developing and applying these
123 flexibility models on clustered residential buildings, ranging from high energy
124 performance detached houses (Germany) to building blocks connected to low-
125 temperature district heating (Switzerland) or a group of buildings formed by
126 small shops, a food market and residential units (Spain). In all the scenar-
127 ios, the methods were applied to remote-controlled heat pumps with different
128 system configurations.

129 The rest of the paper is organized as follows. Section 2 describes the devel-
130 oped methodology, identifies potential flexibility markets, presents a common
131 methodology for quantifying energy flexibility and describes the models and
132 the new *FF* formulations. Hereinafter, the three case studies (Spain, Germany
133 and Switzerland) are presented in Section 3. They comprise three clusters of
134 buildings with heat pumps remotely driven by MPC procedures. The oper-
135 ation of the DR services and the results are summarized in Section 4, where
136 details of the outcomes of the different direct load control tests are presented.
137 The energy flexibility is assessed, through the derived Flexibility models and
138 Flexibility Functions, in this section. **Finally, the findings are extensively dis-**
139 **cussed in Section 5, and summarised in Section 6. Furthermore, some perspec-**
140 **tives for future work are also envisaged in Section 7.**

141 2. Methodology

142 2.1. Identification of the addressed flexibility markets

143 Different markets exist for the trading of electricity between buyers and
144 sellers. In the day-ahead market, products are traded for delivery on the fol-
145 lowing day. The intraday market trades products to balance possible devi-
146 ations from the day-ahead forecast. Balancing or control reserves markets
147 are needed to balance electricity generation and consumption in the short
148 term. Three different types of control reserves markets are available: i) Fre-
149 quency Containment Reserve (FCR), ii) Automatic Frequency Restoration Re-
150 serve (aFRR), and iii) Manual Frequency Restoration Reserve (mFRR). They
151 differ according to the principle of activation, to their bid minimum size and
152 symmetry, and their activation speed. The last category of markets is the Re-
153 serve Replacement (RR) market. These capacity mechanisms aim at ensuring
154 the security of supply from a long-term perspective.

155 In this paper, four of the above-mentioned markets are selected to be ad-
156 dressed through direct load control DR services: i) the Spanish wholesale elec-
157 tricity market (day-ahead); ii) the German operating reserve; iii) the German
158 intraday spot market and; iv) the Swiss imbalance market (aFRR).

159 In Spain, OMIE is the nominated electricity market operator (NEMO) for
160 managing the Iberian Peninsula's day-ahead and intraday electricity markets.
161 The delivery takes place on the day after the trading day (incl. weekends or
162 holidays), and trading sessions take place in one daily auction 365 days/year.
163 Sale and purchase bids can be made considering between 1 and 25 energy
164 blocks in each hour, with power and prices offered in each block. In the case
165 of sales, the bid price increases with the block number; in purchases, the bid
166 price decreases with the block number. The minimum size is 0.1 MW. The
167 Spanish TSO, Red Eléctrica Española, has developed an information system
168 known as 'System Operator Information System (esios)', specially designed
169 to run all the necessary processes to ensure economic and reliable exploitation
170 of the Spanish Power System in real-time. The esios portal offers an open API
171 where the wholesale electricity prices for the next 24 h are published once the
172 spot market is closed (at 13 h of every day). These electricity prices become
173 the control variable for the direct load control services implemented in the
174 Spanish use case.

175 Unlike the day-ahead spot market in Germany and Switzerland, the intra-
176 day market can be described as a corrector market because the time intervals
177 between trade and activation and the activation period are significantly lower.
178 Thereby, electrical energy is traded in intervals of one hour for Switzerland or
179 15 min for Germany. In Germany, trades for 15 min intervals can be completed
180 between 15:00 (CET) of the previous day until 5 min before activation [22].

181 In Germany, four different TSOs are responsible for the reserve markets,
182 and around 60 companies are pre-qualified to deliver operating reserves. There-
183 fore, compared to the spot trade market, there is a highly reduced field of
184 actors. The FCR activation time of a few seconds is very short term. aFRR
185 requires an activation time of less than 30 s and 5 min to reach full power. RR
186 requires 5 min for activation. mFRR and RR are traded daily and bids can be
187 provided in blocks of 4 h. Negative and positive reserve power is traded. As

188 a first instance, positive or negative power is offered with different assigned
189 prices. If an offer is accepted, a working price (e.g. EUR/MWh) is also offered,
190 and the activation occurs according to the working price within a merit order
191 list. The main drawbacks of mFRR and RR are that at least 1 MW of power
192 must be certified. Thereby, an aggregated larger pool operation is necessary.
193 The German operating reserve market, especially mFRR has seen dropping
194 costs within the last years [23, 24], whereas in comparison the amount of en-
195 ergy traded at the EPEX Intraday market has almost doubled from 2014 – 2019
196 (from 47 TWh to 91.6 TWh) [25], shifting the favourability more to intraday
197 trade. In the German pilot site, activations were carried out by Centrica, an
198 aggregator company situated in Belgium, according to available market data
199 from Belgium. This is justifiable due to the fact, that the spot market products
200 are tradable in between Germany and Belgium [26] as well as the operating
201 reserve market conditions are comparable [27].

202 The Swiss operating reserve markets are managed only by one TSO (Swiss-
203 grid). Compared to Germany, the minimum certified bid of the aFRR and RR
204 markets is 5 MW, making them even less accessible for residential buildings,
205 as a vast pool of assets would be needed. For aFRR, the trading is automated.
206 The products traded are asymmetric and must be available 30 s to 5 min after
207 the notification for a duration of up to 15 min. The size of the aFRR in Switzer-
208 land in 2017 was ± 380 MW [28]. The high participation of hydropower supply
209 in the reserve markets limits residential DR. The high number of DSOs present
210 in Switzerland [28], each of them with a limited asset pool, also hinders the
211 development of DR services by the DSOs. For the field tests of the Swiss pilot
212 site, the targeted reserve market was the aFRR, as its market constraints are
213 the most accessible for heat pumps. By combining a pool of batteries with
214 fast activation time and heat pumps whose power availability lasts longer,
215 an aggregator could theoretically fulfil the market constraints. The trading in
216 this work was done by Centrica (aggregator), and HES-SO Valais-Wallis car-
217 ried out the activations in Switzerland. Real trading could not be tested, as it
218 would have required 200 times the capacity offered by the pilot site to reach
219 the minimum bid of 5 MW.

220 2.2. *New reference methodology to assess energy flexibility*

221 The methodology to characterize the energy flexibility in a more standard-
222 ized way follows the initial methodology set out by [20]. This methodology
223 defines a dynamic function, named the flexibility function FF , which charac-
224 terizes the energy flexibility of any device through the use of penalty signals.
225 In our research, the analysed use cases do not strictly follow the activation
226 of the energy flexibility through penalty signals since they respond to other
227 DR schemes. To address these different DR schemes, we took a broader ap-
228 proach than [20] and implemented a methodology to include other kind of sig-
229 nals and activation variables which are more realistic for the analysed energy
230 flexibility markets. The proposed methodology follows the process shown in
231 Fig. 1. As can be seen, the initial point starts with setting up the baseline mod-
232 elling, which corresponds to the energy performance model of the buildings
233 in a Business as Usual (BaU) scenario. This baseline model is then used to
234 forecast the building energy consumption for the time horizon defined by the

235 activation period. This energy forecasting is integrated into a model predictive
 236 control optimization where the activation variable is the output. The cost func-
 237 tion depends on the flexible electricity market to be addressed. The activation
 238 period is different for each use case and flexible electricity market. It is driven
 239 by the optimized activation variable, ranging from a penalty signal, such as
 240 the day-ahead price, a percentage of power activation time, or a power trace
 241 to be tracked. The active power consumed throughout of activation period
 242 is registered. This time series is considered as the dependent variable within
 243 the flexibility model. The baseline forecasting and the activation variable time
 244 series are defined as the independent variables. The flexibility model is then
 245 formulated also to include the corresponding autoregressive terms. The next
 246 step consists of training this flexibility model with historical data of the acti-
 247 vation period. Once the flexibility model is trained and validated, the i -step
 248 prediction is used to define the flexibility function, FF .

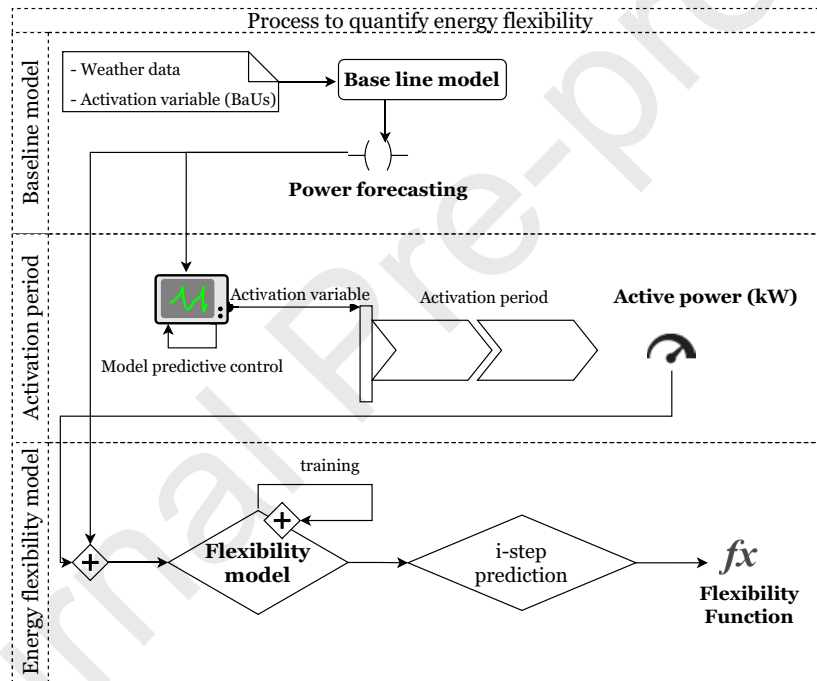


Figure 1 – The general process to quantify the energy flexibility

249 2.2.1. Baseline modelling

250 Since the energy flexibility cannot be directly measured, as it represents
 251 the activation or deactivation of power usage, it is determined by compar-
 252 ing measured power during the activation period and forecasting the power
 253 consumed by the building as if the activation had not taken place. This sup-
 254 posed scenario is called the Business as Usual (BaU) scenario. To determine
 255 the energy load forecasting under the BaU scenario, a model of the thermal
 256 dynamics and the energy consumption of the building, prior to the activation
 257 period, needs to be developed. This model is called the baseline model. The
 258 baseline model can be defined as the energy characterization of the starting

259 situation and has a fundamental role in the determination of energy flexibil-
 260 ity. In fact, the baseline model allows isolating the effects of the activation
 261 variables from the effects of other parameters that can simultaneously affect
 262 the energy consumption. To obtain the baseline model, several approaches
 263 can be followed:

- 264 • Empirical modelling based on a system of differential equations and heat
 265 transfer functions
- 266 • Grey box modelling based on state-space models
- 267 • Data-driven modelling based on transfer function models or machine
 268 learning techniques

269 In this research, the three approaches have been used for the different use
 270 cases. The first approach requires detailed models with several monitored
 271 variables and a calibration stage to fit with the monitored data. An example of
 272 these kinds of calibration processes can be found in [29]. The second approach
 273 requires monitoring the state variable (indoor temperature or water tank tem-
 274 perature) and a precise process to identify the unknown parameters. A[30, 31]
 275 detailed description of the identification procedure applied over suitable grey
 276 box building heat dynamics models is presented. The third approach requires
 277 good data quality of a minimum historical period and the measurement of the
 278 control variable. Several authors applied this last approach to determine the
 279 heat dynamics of buildings. In [32], some of the most common data-driven
 280 methods used to develop baseline models are reviewed. The baseline mod-
 281 els developed in each use case are described in detail and referenced in the
 282 corresponding subsection of Section 3 of this document.

283 2.2.2. Flexibility models

284 A flexibility model is a regression-based model which aims at finding the
 285 correlation among the active power, the activation variable and the power
 286 under the BaU scenario. In this research, a data-driven approach is followed
 287 based on Autoregressive (AR) models. As previously mentioned, the initial
 288 modelling technique is defined by [20] is modified to adapt it to the specific
 289 constraints of the different activation variables. [The original model by Junker](#)
 290 [et al. assumes that the active load when exposed to a the penalty signal can](#)
 291 [be separated into two parts; the load that dynamically responds to the the](#)
 292 [the penalty, and the non-responsive load \(baseload power, in our equations\).](#)
 293 [However, in our case the dynamics due to the active and baseload signal itself](#)
 294 [are added. Thus, an ARX model is considered based on the initial equation](#)
 295 [presented in \[20\]. This allows to better estimate the amount of time and load](#)
 296 [that can be flipped once an activation, a change of price, or a trace to follow is](#)
 297 [received. Additionally, it helps to the proper estimation of the rebound effect](#)
 298 [caused by a change in the penalty, as it considers the thermal inertia available](#)
 299 [in the system.](#)

300 In the use cases when the activation variable is the day-ahead electricity
 301 price of the wholesale spot market, the model formula is described in Eq. 1.

$$\phi_{T_0}(B)P_t^e = \omega_{T_0}(B)P_t^b + \Psi_{T_0}(B)DA_t + \varepsilon_t \quad (1)$$

302 The auto regressive terms $\phi_{T_o}(B)$, $\omega_{T_o}(B)$ and $\Psi_{T_o}(B)$ are the parameters of
 303 the model. The sub index T_o represents their dependence with one categorical
 304 variable, the outdoor temperature. In order to better express this dependency,
 305 a 4 hours moving-averaged transformation is applied over the outdoor tem-
 306 perature for the testing periods. This averaged temperature is further split in
 307 two levels: [6.67 °C - 12.3 °C] and [12.3 °C - 21.5 °C]. Therefore, the T_o is not
 308 used as a exogenous variable of the model. The backward shift operators, B ,
 309 are defined as $B^k y_t = y_{t-k}$, where y_t is the considered variable (P_t^e , P_t^b, DA_t)
 310 at time t and $k \in [0, j]$. Here, j refers to the maximum order allowed to that
 311 backward shift operator, B . DA_t corresponds to the activation variable, the
 312 day-ahead electricity price. The ε_t corresponds to the white noise residual of
 313 the model at time t .

314 In the use cases when the activation variable is the percentage of activation
 315 time within each time step, the model formula is described in Eq. 2.

$$\phi_{bd}(B)P_t^e = \omega_{bd}(B)P_t^b + \Psi_{bd}(B)A_t + \varepsilon_t \quad (2)$$

316 The auto regressive terms $\phi_{bd}(B)$, $\omega_{bd}(B)$ and $\Psi_{bd}(B)$ are the parameters of
 317 the model. The sub index bd represents their dependence with one categor-
 318 ical variable, the building number. bd comprises the categorical values of
 319 the building number for this use case [20, 22, 24, 25], and a virtual build-
 320 ing that aggregates the power of all of them. Therefore, the building number,
 321 bd , is not used as a exogenous variable of the model. The backward shift
 322 operators, B , are defined as $B^k y_t = y_{t-k}$, where y_t is the considered variable (P_t^e ,
 323 P_t^b, A_t) at time t and $k \in [0, j]$. Here, j refers to the maximum order allowed
 324 to that backward shift operator, B . A_t corresponds to the activation variable,
 325 which is the percentage of time, within every time step, with active power,
 326 $A_t = [0\% - 100\%]$. The ε_t corresponds to the white noise residual of the model
 327 at time t .

328 In the use cases when the activation variable is a trace to be tracked, the
 329 power used within the activation period is no longer the model's dependent
 330 variable. In Eq. 3, the dependent variable is substituted by the difference be-
 331 tween the active power, P_t^e , and the baseline power, P_t^b . The activation vari-
 332 able in Eq. 1 is then substituted by the difference between the power trace to
 333 be tracked, P_t^f and the baseline power, P_t^b . The modified formula is shown in
 334 Eq. 3.

$$\phi_{s,T_o}(B)(P_t^e - P_t^b) = \omega_{T_o}(B)(P_t^f - P_t^b) + \varepsilon_t \quad (3)$$

335 The autoregressive terms $\phi_{s,T_o}(B)$, $\omega_{T_o}(B)$ are the parameters of the model.
 336 The sub-index T_o represents their dependence on a categorical variable, the
 337 outdoor temperature. Based on a 4 hours moving-averaged transformation
 338 of the outdoor temperature, for the test periods, the results are split into two
 339 groups of outdoor temperature levels: [6.5 °C - 15.7 °C] and [15.7 °C - 28.5 °C].
 340 The sub-index s refers to the sign of the trace to be tracked in relation to the
 341 baseline power, being equal to 1 when it is positive, equal to 0 when there is no
 342 difference with the baseline power, and equal to -1 when it is negative. There-
 343 fore, neither the T_o nor the s are used as exogenous variables of the model. The

344 backward shift operators, B , are defined as $B^k y_t = y_{t-k}$, where y_t is the consid-
 345 ered variable (P_t^e, P_t^b, X_t) at time t and $k \in [0, j]$. Here, j refers to the maximum
 346 order allowed to that backward shift operator, B . The ε_t corresponds to the
 347 white noise residual of the model at time t .

348 2.2.3. Flexibility Functions

349 The Flexibility Function (FF) can be understood as the impulse response
 350 function of each flexibility model since the flexibility models include autore-
 351 gressive terms of the dependent variables, which cause an influence over the
 352 P_t^e when $t \geq 1$. To do so, an i -step prediction is performed to estimate the
 353 impulse response of the models properly.

In the use case when the activation variable is the day-ahead price, the FF is determined based on a positive and a negative change in the day-ahead electricity price (± 0.1 €/kWh) for the time steps $n = 15, 60$ and 120 minutes and for a flexibility evaluation period of $i = 480$ minutes. When the activation variable is the percentage of time of activation within each time step, 100% activation signals for time steps of $n = 1, 2$ and 4 hours are tested along a flexibility evaluation period of $i = 12$ hours. Both use cases follow a similar procedure to determine the FF :

$$t = (0, 1, \dots, i) \quad (4a)$$

$$P_{t \leq 0}^e = 0 \quad (4b)$$

$$P_{t \in \mathbb{N}}^b = 0 \quad (4c)$$

For the day-ahead electricity price as the activation variable:

$$DA = \begin{cases} 0.1, & \text{if positive price change} \\ -0.1 & \text{if negative price change} \end{cases} \quad (4d)$$

$$DA_t = \left(\underbrace{0, (DA, \dots, DA)}_{n \text{ times}}, \underbrace{(0, \dots, 0)}_{n-i \text{ times}} \right) \quad (4e)$$

$$\Phi_{k=0}^b(B) P_t^e = -\Phi_{k \geq 1}^b(B) P_t^e + \Psi^b(B) DA_t \quad (4f)$$

$$\Phi_{k=0}^b(B) = 1 \quad (4g)$$

$$FF_t = P_t^e = -\Phi_{k \geq 1}^b(B) P_t^e + \Psi^b(B) DA_t \quad (4h)$$

For the percentage of time activation within a time step as the activation variable:

$$A_t = \left(\underbrace{0, (100, \dots, 100)}_{n \text{ times}}, \underbrace{(0, \dots, 0)}_{n-i \text{ times}} \right) \quad (4i)$$

$$FF_t = P_t^e = -\Phi_{k \geq 1}^b(B) P_t^e + \Psi^b(B) A_t \quad (4j)$$

In the use case when the activation variable is a trace to be tracked, the FF is determined by considering a 100% activation signal of time steps $n = 15, 30$ and 60 minutes for a flexibility evaluation period of $i = 120$ minutes. A multi-step prediction method is used to predict the expected response of ± 1 kW of the trace to be tracked. The previous estimate of the flexibility function, $(P^e - P^b)$, is used for the new prediction step. The baseline power is set to

$P_t^b = 0$ for $t \in (0, 1, \dots, i)$. Here, s is equal to 1 if the activation is positive and -1 if it is negative.

$$(P_{t \leq 0}^e - P_{t \leq 0}^b) = 0 \quad (5a)$$

$$(P_t^f - P_t^b) = \begin{cases} 0, & \text{if } t \leq 0 \\ ((\underbrace{s, \dots, s}_{n \text{ times}}, \underbrace{0, \dots, 0}_{i-n \text{ times}})), & \text{otherwise} \end{cases} \quad (5b)$$

$$\Phi_{s, T_0}(B)(P_t^e - P_t^b) = \omega_{T_0}(B)(P_t^f - P_t^b) \quad (5c)$$

$$\Phi_{s, T_0 k=0}(B)(P_t^e - P_t^b) = -\Phi_{s, T_0 k \geq 1}(B)(P_t^e - P_t^b) + \omega_{T_0}(B)(P_t^f - P_t^b) \quad (5d)$$

$$\Phi_{s, T_0 k=0}(B) = 1 \quad (5e)$$

354 Considering the flexibility model of Equation 3 and the set up described in
355 previous equations, the FF is defined as:

$$\begin{aligned} FF_t &= (P_t^e - P_t^b) \\ &= -\Phi_{s, T_0 k \geq 1}(B)(P_t^e - P_t^b) + \omega_{T_0}(B)(P_t^f - P_t^b) \end{aligned} \quad (6)$$

356 3. Case studies

357 The methodology to evaluate the energy flexibility is applied over three
358 case studies which have in common a direct load control of space heating
359 systems driven by heat pumps:

- 360 • Case study of the Spanish wholesale electricity market price as the ac-
361 tivation variable. Blocks of buildings placed in North-East Spain (Sant
362 Cugat)
- 363 • Case study of the percentage of activation time as the activation variable.
364 Residential households placed in South Germany (Wüstenrot)
- 365 • Case study of a trace to be tracked as the activation variable. Blocks of
366 residential buildings placed in Switzerland (Naters)

367 A new player, called the Cluster Manager (CM), is incorporated in these
368 case studies. CMs are site managers that cluster together with the local en-
369 ergy, which are remotely controlled (e.g. heat pumps). They have technical
370 knowledge of these energy systems and the connected devices (control sys-
371 tem, meters, sensors...). They manage these assets and act as the bridge be-
372 tween the aggregator, who bid in the markets, and the end-user. Thus, they
373 do not have to deal with market specifications handled by the aggregator.

374 3.1. Spanish case study: wholesale market price

375 This case study is a pilot site constituted by buildings that combine apart-
376 ments, offices, shops and a local food market. They are placed in a city called
377 Sant Cugat, in Northern East Spain. Figure 2 shows the space heating and
378 cooling system configuration. It comprises a water storage tank of 3,500 litres
379 fed by two reversible heat pumps accounting for 60 kW of electric power. The

380 heat pumps are controlled by an immersed temperature probe inserted into
 381 the bottom of the water tank. The heat pumps deliver thermal energy to the
 382 water storage tank through a primary circuit with two hydraulic pumps and
 383 external heat exchangers, which follow the same operation schedules as the
 384 heat pumps. The water tank provides hot and cold water to two different
 385 hydraulic circuits, which transfer this thermal energy to 32 offices, 3 shops
 386 and a local food market. These hydraulic circuits are managed by two 3-way
 387 motorized valves incorporating a proportional integral derivative (PID) con-
 388 trol, leading to variable water volume flow rates. The control variable of the
 389 system is the water tank setpoint temperature. Since the two heat pumps do
 390 not have variable-speed compressors, they are thermostatically controlled in
 391 ON/OFF modes.

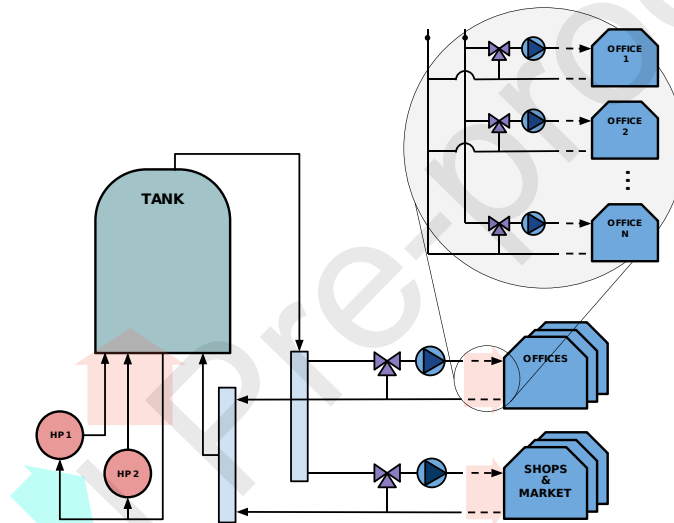


Figure 2 – Space heating configuration of the Spanish use case. The zoom shows details of the hydraulic distribution ring

392 The direct load control strategy followed in this use case is based on the
 393 augmented heat pumps performance with price information from the whole-
 394 sale market and weather forecast data for the current and following day. The
 395 heat pumps' electrical use adjusts times when the Spanish wholesale market
 396 spot price is lower (day-ahead optimization). To make these services opera-
 397 tional, a Model Predictive Control (MPC) approach is put into practice. Every
 398 day at 00:00, a Genetic Algorithm (GA) optimizes the cost function, which is
 399 the minimum daily electricity consumption cost and gets the vector of the set-
 400 point temperature of the water storage tank, T_{opt}^s , for the next 24 hours in the
 401 more cost-effective way.

402 The baseline model is developed based on the third approach mentioned
 403 in Section 2.2.1. It is a data-driven approach formed by two ARX models.
 404 They define the dynamic energy balance between the electricity load of the
 405 heat pumps, the water tank temperature and the thermal energy delivered to
 406 the offices, to the shops and the local food market, as well as the thermal losses

407 in the water storage tank and the water distribution rings. More details of this
 408 kind of model can be found in [33]. These two forecasting models need, as
 409 inputs, day-ahead predictions of the thermal energy consumed by the shops,
 410 the offices, and the local food market. Since they form a block of buildings,
 411 they can be simplified as a multi-space building formed by several thermal
 412 balance nodes. This model is expected to behave highly non-linear in relation
 413 to the external temperature and other climate-dependent exogenous variables.
 414 Therefore, data-driven models are also used to evaluate their energy perfor-
 415 mance. After a previous fine-tuning phase, where several machine learning
 416 models were evaluated, the Generalised Additive Model GAM, developed by
 417 Hastie et.al [34], provided the highest accuracy and was the selected one.

418 3.2. German case study: percentage of activation time

419 This case study is a pilot site situated in the rural municipality of Wüstenrot
 420 in southwest Germany. It consists of a newly built positive energy settlement
 421 with 18 residential single and multifamily buildings. These buildings are con-
 422 nected to a low-temperature district heating grid fed by a so-called “agrother-
 423 mal” – a large scale geothermal - collector. All buildings are equipped with
 424 decentralized heat pumps, thermal buffer storage tanks ranging from 175 to
 425 300 litres, radiant floor systems, and photovoltaic (PV) systems of installed
 426 power between 6 and 29 kWp per building. In addition, a cloud-based moni-
 427 toring system is installed for 12 buildings that include all relevant thermal and
 428 electrical energy flows. Within those 12 buildings, a local energy management
 429 system is installed to control the heat pumps. Figure 3 shows a scheme of the
 430 energy systems configuration of one of the households. Since different man-
 431 ufacturers provided the heat pumps, some connectivity problems appeared
 432 with the interfaces of some of them and the activation was only carried out
 433 for four heat pumps manufactured by Tecalor (Typ TTF 10 and TTC05). Two
 434 of these heat pumps have a maximum electrical power of 2.38 kW, and two
 435 have a maximum power of 3.82 kW. These activations aimed to test the poten-
 436 tial and challenges of flexible control of heat pumps from the viewpoint of a
 437 flexible service provider.

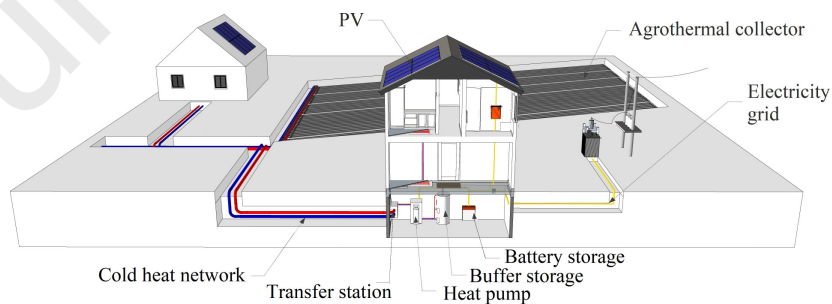


Figure 3 – Energy systems configuration of one of the single households of Wüstenrot pilot site

438 The development of the baseline model followed the first approach men-
 439 tioned in Section 2.2.1. For four of the selected households, a white-box model
 440 of each building was generated. More details of the models can be found in

441 [35] and in [36]. They include heat pumps, buffer storage water tanks and
442 control systems. To increase the model's accuracy, a calibration on parameters
443 changeable by the users (indoor setpoint temperature and air exchange rate)
444 with measured data was carried out. Given the unavailability of a baseline
445 for the fifth household, due to inadequate monitoring data, this baseline has
446 been derived from another house which was most similar (same heat pump
447 type and no heating buffer) applying a linear extrapolation based on the his-
448 torical consumption difference of both. Input parameters for the heat pump
449 control are active power, DHW temperatures and floor heating temperatures.
450 The control strategy is a direct load control over heat pumps on/off.

451 3.3. *Swiss case study: trace to be tracked*

452 This case study is a pilot site placed in the municipality of Naters, in South-
453 ern Switzerland. It comprises 12 residential multi-family buildings connected
454 to a centralized low-temperature district heating network (anergy network).
455 It represents 166 residential units. The size of the buildings ranges from 4 to 36
456 residential units per building. The buildings' construction years range from
457 1919 to 2015. Thus their envelopes have different thermal efficiencies and have
458 either radiators or floor heating systems. Each building is equipped with one
459 or two fixed speed compressor heat pumps, thermal buffer storage tanks for
460 SH and DHW. Hardware components called 'gateways' are installed in each
461 building. They collect, process and export data from the building devices (e.g.
462 heat pumps, electricity meters) to a cloud-based platform that enables remote
463 control of the heat pumps. The gateways installed in this use case do not have
464 the same level of internal intelligence as the management system installed in
465 the German use case. Due to some restrictions in the control interfaces with
466 the heat pumps, only five out of fourteen heat pumps were intensively tested,
467 accounting for a maximum aggregated electricity power of 34.3 kW. Figure 4
468 shows and scheme the energy systems configuration of the multi-apartment
469 buildings.

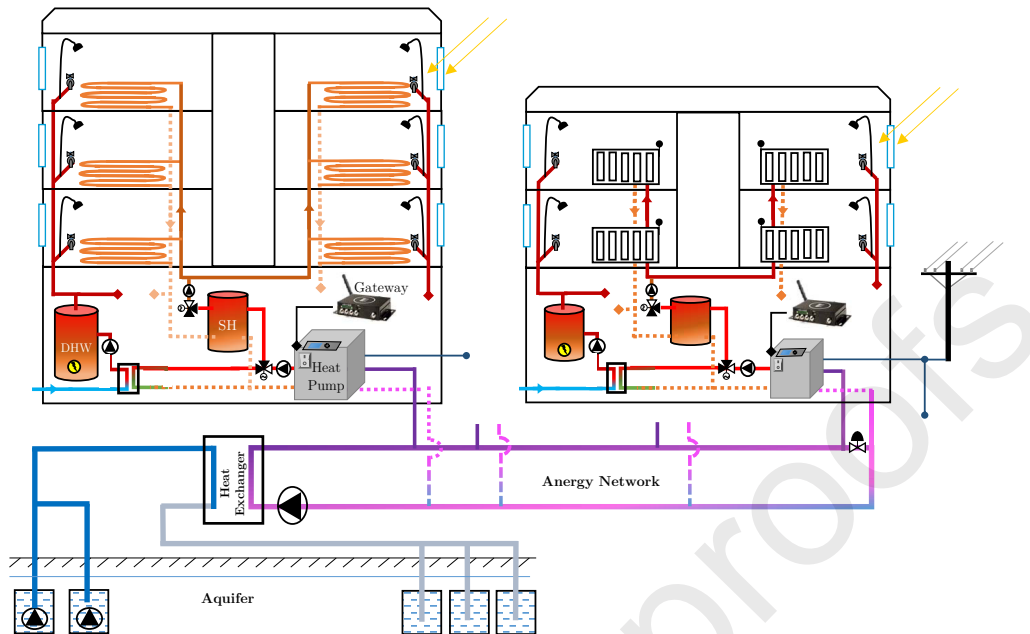


Figure 4 – Energy systems configuration of one of the multi-apartment buildings of Nater's pilot site

470 The test aims to confirm the potential and challenges of flexible control of
 471 heat pumps in residential buildings from the viewpoint of a flexible service
 472 provider. A transactive DR approach was tested (a two-way communication
 473 system). Its reliability and performance over consecutive days with multiple
 474 DR-events per day was also assessed. The framework can be divided into
 475 three steps: i) the site is waiting to provide DR services by running BaU; ii) the
 476 aggregator starts negotiating power traces with the CM; iii) once a trace has
 477 been agreed on, the CM tracks it with an MPC adapted from the formulation
 478 developed by [37]. The baseline trace is modelled based on the third approach
 479 mentioned in Section 2.2.1. It is a data-driven approach formed by a Seasonal
 480 Autoregressive model (SAR) for each building using the past 3 days' power
 481 data. The aggregated baseline for the site is computed by summing up the
 482 estimated baseline of each building. The other traces are generated by solving
 483 scheduling optimization problems. The control variables of the heat pumps
 484 are the SH and DHW temperature set points, which are increased/decreased
 485 based on the new values optimized by the MPC.

486 4. Results

487 4.1. Operation of the Spanish case study

488 Figure 5 depicts the results of the direct load control applied in the case
 489 study where the Spanish wholesale spot market price acts as the activation
 490 variable. An MPC optimization was applied during the activation period,
 491 which comprised from March 29th to April 12th 2020.

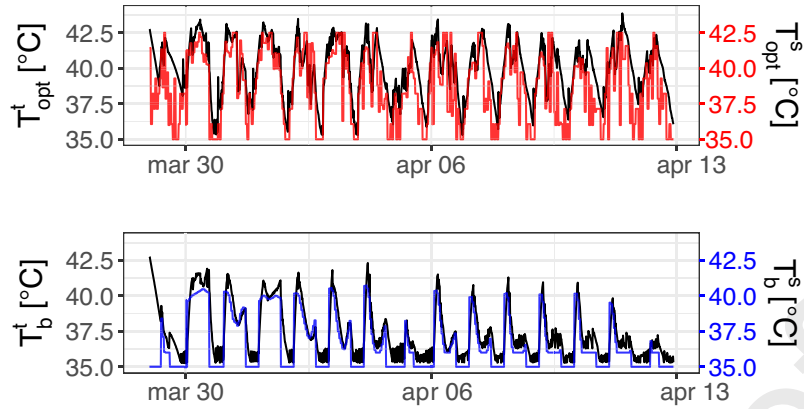


Figure 5 – Results of the direct load control of the use case of the Spanish wholesale spot market price as activation variable

492 The upper figure shows, in a black-coloured line, the monitored active wa-
 493 ter storage tank temperature, T_{opt}^t along the activation period. It is compared
 494 with the water storage tank setpoint temperature, T_{opt}^s , the red-coloured line,
 495 obtained as the output of the day-ahead optimization performed every day.
 496 The T_{opt}^s is the direct control variable that drove the heat pumps performance
 497 along the activation period. As can be seen, the water tank temperature fol-
 498 lows the optimized setpoint temperature very well. The lower plot shows the
 499 simulated baseline forecasting of the water storage tank temperature (black-
 500 coloured line), T_b^t , and the corresponding setpoint temperature (blue-coloured
 501 line), T_b^s , in the BaU scenario, which is minimum operational temperature
 502 level required by the offices, shops and local food market to keep the comfort
 503 requirements. The differences in both plots show the effect of the activation.
 504 It can be seen that the baseline forecasting usually has two temperature peaks
 505 and a second smaller temperature level. In contrast, the optimized temper-
 506 ature shows a single peak that is slightly lagged in time. This time lagging
 507 shows the MPC is shifting the higher setpoint temperature values to the peri-
 508 ods with lower electricity prices.

509 4.1.1. Time series inputs for the flexibility model development

510 In Figure ,6 the day-ahead signal price, DA , the forecasting of the baseline
 511 power load, P^b , and the active power of the heat pumps, P^e , are shown. Com-
 512 paring the two time series of power, the differences due to the MPC are ap-
 513 preciable. The bigger differences can be seen for the first days of April, where
 514 the active power is concentrated in the lower price hours while the baseline
 515 forecasting also consumes in higher prices periods.

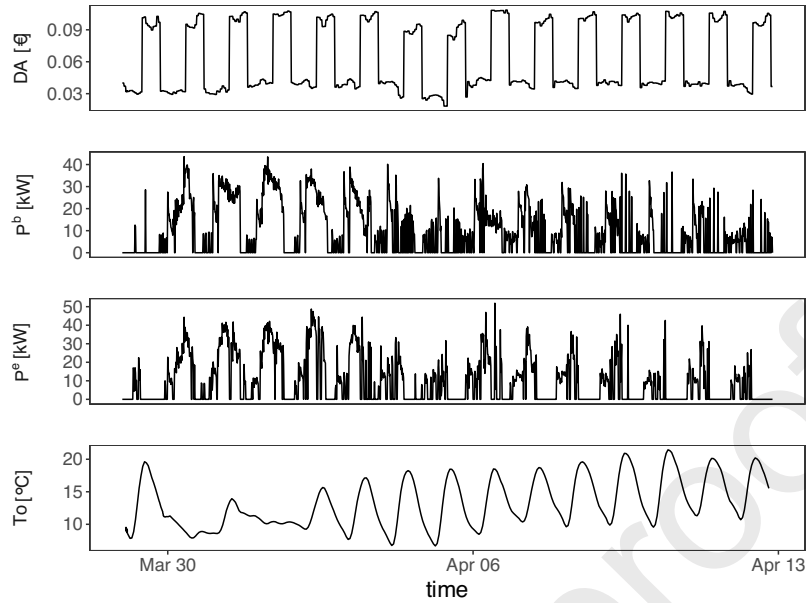


Figure 6 – Active power, P^e , forecasting of the power baseline, P^b , day-ahead electricity price, DA and outdoor temperature, T_o of the heat pumps of the Spanish use case, during the direct load control operation period

516 Since the objective of this use case is to reduce the cost of the energy con-
 517 sumption of the heat pumps, Figure 7 depicts the accumulated cost difference
 518 achieved between the active optimized energy performance (black line) and
 519 the BaU scenario (red line). The reduction of cost reaches 18 % at the end
 520 of the field test operation period. This is an auspicious outcome to consider
 521 day-ahead price optimization as an important way to optimize the operational
 522 costs of heat pumps systems while offering flexibility to the electricity system.

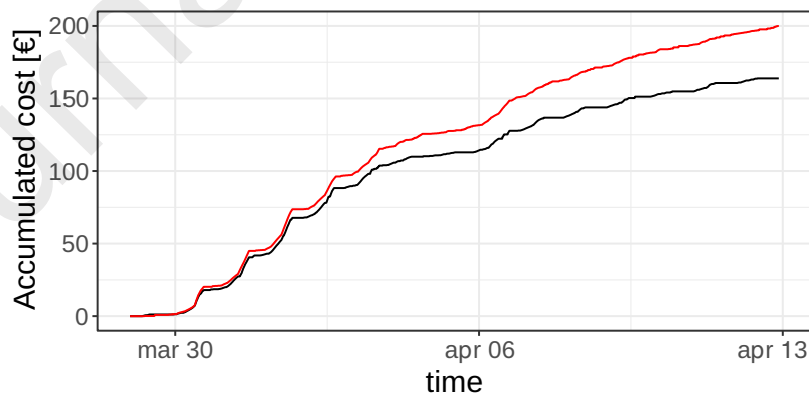


Figure 7 – The accumulated cost of the active optimized energy performance (black line) compared to the BaU scenario (red line)

523 4.2. Operation of the German case study

524 Before operating the Tecalor heat pumps, different tests were conducted
 525 to verify their control capabilities. An upwards signal of 100 % activation for

526 30 minutes, followed by a stop of 10 minutes and the second activation of 15
 527 minutes was sent to the heat pump controller. The result is shown in Figure 8.
 528 The time to start up was 56 seconds from the setpoint to on. Besides, the heat
 529 pump needed 15 minutes to reach 75 % of the maximum power. It can also be
 530 seen that the activation profile started with a first step increase, followed by a
 531 roughly linear ramp. The time to shut down was 1 min 22 seconds, whereas
 532 the shutting down profile was a decreasing step function. There is a 20-minute
 533 recovery time between switching off and switching the heat pump on again.
 534 These factors determine how a flexibility service provider can control the heat
 535 pump flexibly and integrate it into a virtual power plant.

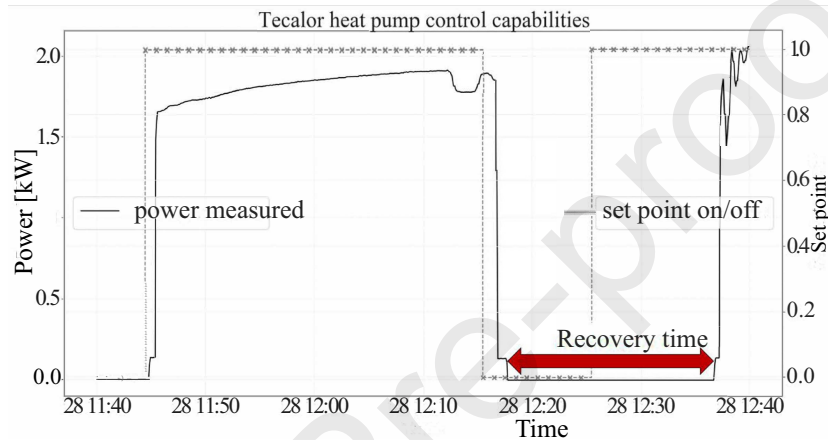


Figure 8 – Heat pump control capabilities analysis

536 Another test was performed to assess a stepwise activation. For certain
 537 flexibility services, a heat pump may have to deliver a linear increasing or de-
 538 creasing power curve (e.g. track the TSO's aFRR signal). Since the modulation
 539 of the power output of the heat pumps was not possible, the test performed to
 540 deliver a linear ramp was based on stacking the deactivation of heat pumps.
 541 In this test, 1 minute between each heat pump switching on/off was set up,
 542 and a variation time of switching on between 5 to 30 minutes. During this
 543 test, 3 heat pumps were available at the case study pilot site. Temperature
 544 measurements of both the DHW and the floor heating system were available,
 545 allowing us to estimate the available flexibility in the system. The ranking of
 546 the heat pumps to switch them on and off was based on the measured tem-
 547 perature in the floor heating circuit, which turned out to be the limiting factor.

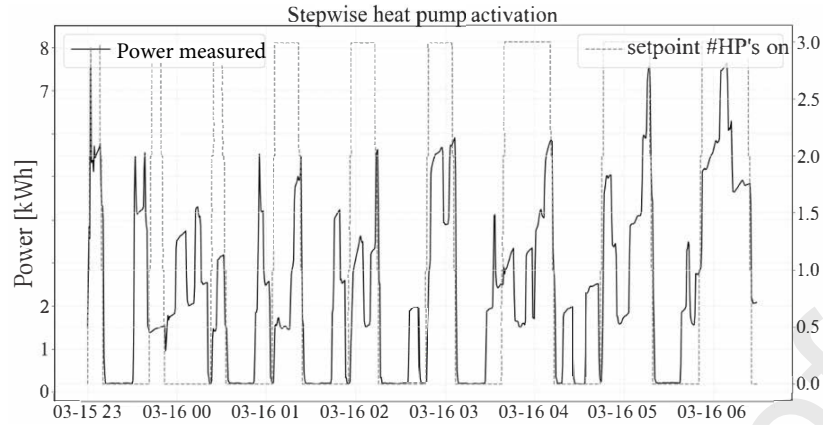


Figure 9 – Stepwise action of 3 heat pumps in the German pilot site

548 This test is shown in Figure 9. The results were not satisfactory to de-
 549 liver a service such as aFRR standalone. This can be explained by the small
 550 pool (3 units) and the fact that the heat pumps were often unavailable for
 551 (de)activation due to comfort/safety constraints. Furthermore, since the heat
 552 pumps controls are driven by load curves that are dependent on the indoor
 553 and outdoor temperatures, and the latest was high for the testing period, the
 554 heat pump power demand was lower than initially expected. In Figure 10 a
 555 deeper zoom on the (un)availability causes of one of the heat pumps is shown.
 556 Number 1 indicates forced on the situation, which means unavailability of the
 557 heat pump. This is due to the DHW temperature dropping below a threshold,
 558 forcing the heat pump to switch on for comfort reasons. Number 2 indicates
 559 a forced off situation, which means the heat pump is unavailable because the
 560 floor heating temperature exceeds the threshold temperature, forcing the heat
 561 pump to switch off for comfort/safety reasons. After the temperature drops
 562 again below the low threshold, the heat pump can be activated again, as can
 563 be seen from the graph.

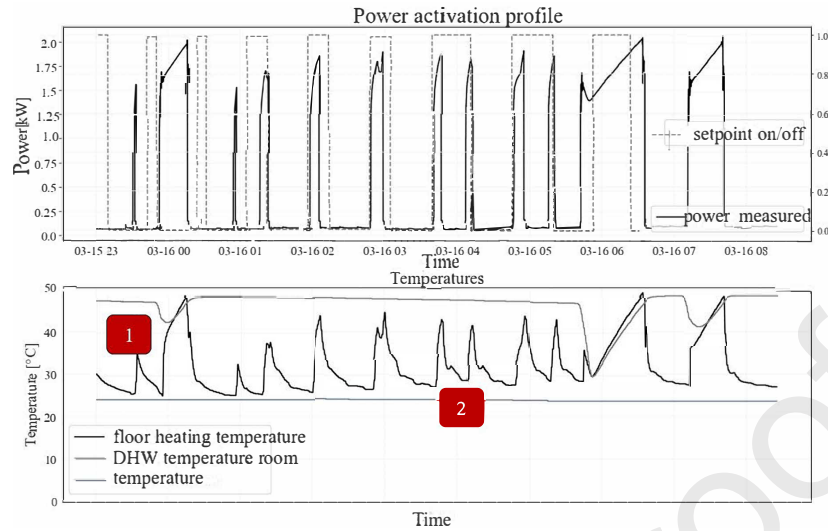


Figure 10 – Power and temperatures of the heat pump systems along the activation period

564 Looking at the overall results of the performed tests, it has been demon-
 565 strated that the flexible operation of heat pumps in the cases study is possi-
 566 ble and can be leveraged for multiple flexibility services or energy markets.
 567 Nevertheless, important points of attention are: i) the latency to ramp up to
 568 full power to ramp down to switch it off, which is around 1 minute; ii) and
 569 the recovery time, which is around 20 minutes. Furthermore, the comfort set
 570 points and the available storage in hot water tanks or the inertia of the build-
 571 ing clearly determine the duration for which the heat pump can be switched
 572 on or off.

573 4.2.1. Time series inputs for the flexibility model development

574 The operation of the case study in Wüstenrot, Germany, was a direct load
 575 control of four of the available heat pumps considering activation signals sent
 576 by a commercial aggregator. When activation was sent, the heat pumps had
 577 to operate for as long as possible during the whole activation period. In this
 578 case study, the control variable is the percentage of activation time (ON/OFF)
 579 of each building or heat pump (named 20, 22, 24 and 25). The energy flexi-
 580 bility is also analysed from this point of view. Figure 11 depicts the operation
 581 performance of the heat pumps from February 15th to March 31st. During
 582 those days, some activation signals were sent by the commercial aggregator.
 583 Therefore, as the actual heat pumps operation was affected by these signals,
 584 large differences between active power, P^e , and the forecasting of the baseline
 585 power in BaU scenario, P^b , can be appreciated for the activation period.

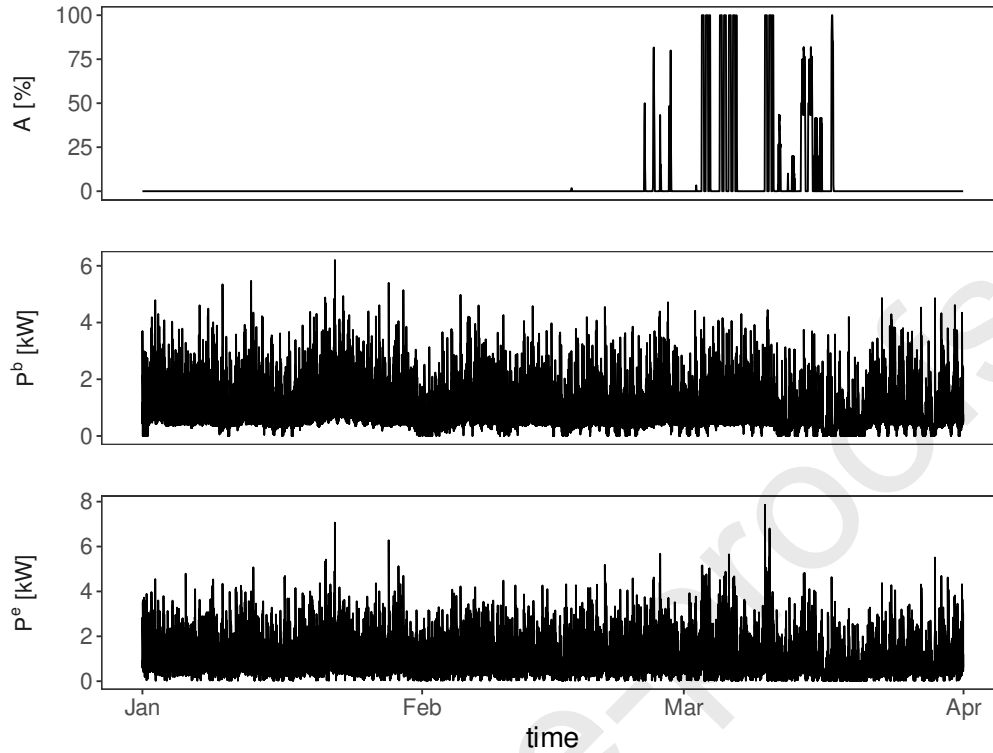


Figure 11 – Active heat pumps power (P^e), forecasting of baseline power in BaU scenario (P^b) and percentage of activation time in each hour (A) of four households in Wüstenrot pilot site

586 4.3. Operation of the Swiss case study

587 The use case in Naters, Switzerland, consists of a direct load control of
 588 five HPs that consider activation traces negotiated between the CM and the
 589 commercial aggregator. When an activation trace is accepted, the heat pumps
 590 should track the trace during the whole activation period.

591 Figure 12 represents the results of a day from a week-long test of direct
 592 load control services, detailed at the building level. The light grey vertical ar-
 593 eas display the 15 minute negotiation periods between the aggregator and the
 594 CM. The light red vertical areas display the direct load control periods per-
 595 formed on-site as solutions of the tracking MPC optimization. It is not always
 596 easy to assess what a system would have done without direct load control, but
 597 coupling set points, temperature and power measurements can visually help.
 598 As a reminder, HP's local control works with hysteresis on the temperature of
 599 each storage. When the storage temperature drops too far below the setpoint
 600 value of the hysteresis, the compressor starts, and the HP runs until the upper
 601 value of the hysteresis is met. This is, of course the theory, but unforeseen
 602 events can sometimes change this behaviour.

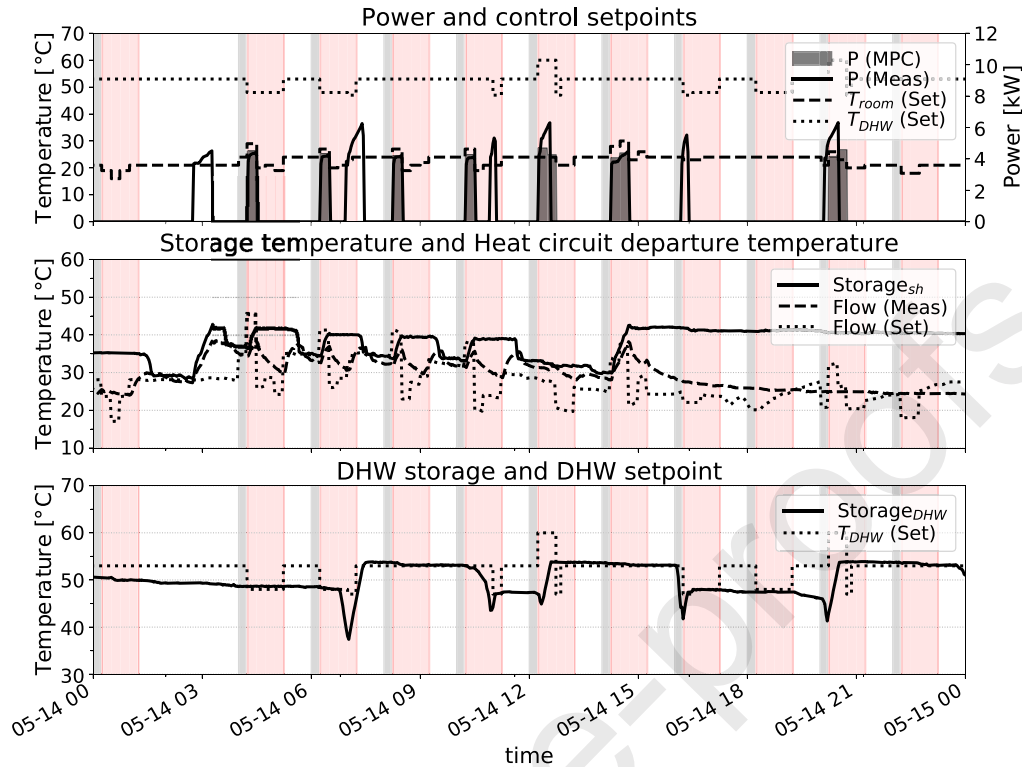


Figure 12 – Power and temperature variation resulting from direct load control for one building in the Swiss pilot site

603 The top panel of Figure 12 shows both the temperature setpoints used for
 604 controlling the HP and the power measurements. The dotted lines correspond
 605 to the setpoint values for SH and DHW, respectively. Outside the direct load
 606 control periods, the values of those set points are set back to their default val-
 607 ues. The solid coloured line displays the measured power consumed by the
 608 compressor of the HP. The solid coloured bars are the power consumption
 609 given as the solution of the tracking MPC. The middle panel represents the
 610 effect of direct load control on SH. The dashed line corresponds to the mea-
 611 sured departure temperature of the heating circuit after the 3-way valve. The
 612 dotted line represents the theoretical departure temperature of the circuit as
 613 given by the heat curve of the HP. It is modelled as a function of the SH set-
 614 point displayed in the top panel and T_o averaged over 3 hours. The bottom
 615 panel represents the effect of direct load control on DHW. In Figure 12, it can
 616 be seen that direct load control of SH perfectly matches the results of the track-
 617 ing MPC. Instead, for the DHW load, it appears to be more difficult. Having
 618 only one sensor to assess the energy state inside the DHW storage tank makes
 619 it difficult to predict when a new cycle will occur. For comfort reasons, DHW
 620 is always prioritized and setpoints are only reduced to a minimum of 47 °C.
 621 Therefore, delaying a DHW cycle for more than 30 minutes is not always pos-
 622 sible, as demonstrated for the DR call at 06:00. In the bottom panel, we can see
 623 that the storage temperature at the start of the period is low. This is because
 624 the setpoints are set to the lowest possible value. At 06:40, a DHW consump-
 625 tion brought the storage temperature below the lower bound of the hysteresis,

626 which starts a new DHW cycle. The DR called at 10:00 is a good example of
 627 the usefulness of MPC when dealing with direct load control. When the power
 628 traces are generated, the storage tank temperature is maximal. There is only a
 629 small chance that a DHW cycle will happen in the next hour. However, within
 630 the third 15 minutes interval, a sudden high DHW consumption puts the stor-
 631 age temperature below the lower bound of the hysteresis, and the heat pump
 632 starts a new DHW cycle. At 11:00, to avoid deviating further from the trace,
 633 the DHW setpoint is reduced, which directly stops the heat pump.

634 4.3.1. Iterative tracking performance

635 Figure 13 presents the results of a day from a week-long test of direct load
 636 control services over all the HPs. The light grey vertical areas display the
 637 15 minute negotiation periods between the aggregator and the CM. The light
 638 red vertical areas display the direct load control periods performed on-site as
 639 solutions of the tracking MPC optimization.

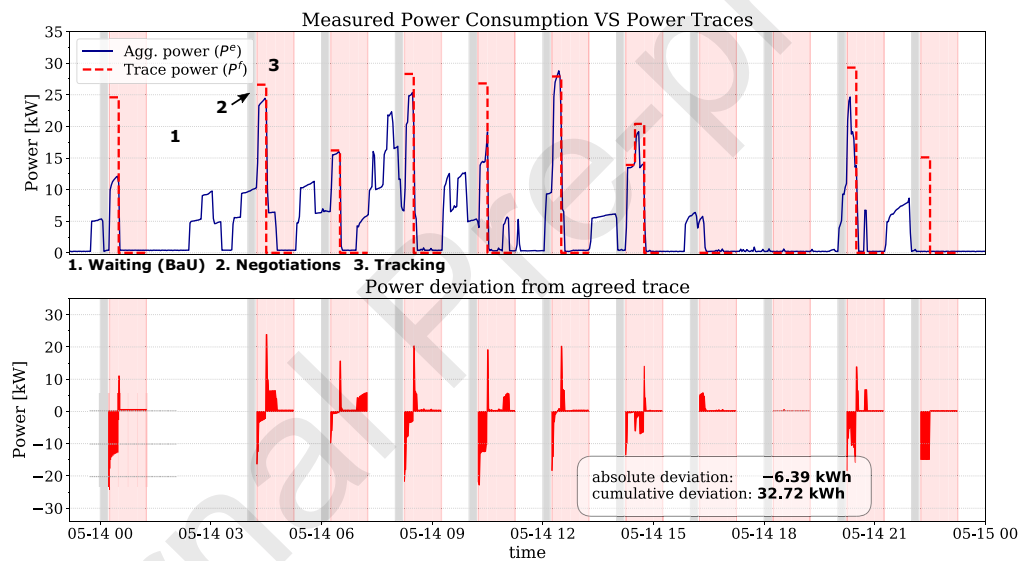


Figure 13 – Power deviation compared to the agreed-upon traces resulting from the DR calls over a day for a weekly test in the Swiss pilot site

640 The top panel of Figure 13 displays the aggregated power (blue) of five
 641 participating HPs on May 14th 2020. The daily average outside temperature
 642 is 18 °C with temperatures above 20 °C from 12:00 to 20:00. Therefore, most
 643 HP consumption occurs during the early hours of the day when the outside
 644 temperature is still cold. The dashed red lines are the power loads P^f agreed
 645 upon by the aggregator and the CM. The selected traces are assumed to be
 646 constant over the sampling period of 15 minutes. Each one corresponds to
 647 a power trace resulting from a 6-hour forecast scheduling optimization prob-
 648 lem proposed by the CM and selected by the aggregator. The bottom panel
 649 of Figure 13, represents the power deviation ($P^f - P^e$). When the values are
 650 negative, it means that the on-site power was lower than the expected trace,
 651 and when they are positive, it means that the power was higher. The relative

652 deviation over the day is -6.4 kWh and the cumulative deviation, computed
653 as the sum of all the absolute deviations, is equal to 32.7 kWh. When high
654 power change occurs as a result of direct load control, high deviation spikes
655 can be observed. The negative spikes correspond to an activation delay of the
656 HPs: Even when conditions for the local controller are met, HP compressors
657 are only started after a 2-minute delay by the local controller. To compen-
658 sate, the tracking MPCs are launched two minutes before the new actuation
659 periods. As soon as an optimal solution is found, the new setpoints are sent.
660 Setpoints to switch off HPs are sent at the actuation time. HP compressors
661 directly stop when conditions are met, except when an explicit minimum run-
662 ning time is implemented by the local controller. The positive spikes observed
663 can be the result of the monitoring sampling rate of 2 minutes and of the way
664 power is measured: The power consumption of four out of five HPs is not
665 directly measured but reconstructed from operating temperature time series
666 and manufacturer datasheets. The interpolation and the model formulation
667 can sometimes create mismatches.

668 4.3.2. *Time series inputs for the flexibility model development*

669 In this use case, the objective of the flexibility function is to characterize
670 how flexible the HP consumption was due to the activation trace accepted
671 by both entities in terms of amount and shift in time. In this case study, the
672 control variables are the DHW and SH setpoint temperature of five multi-
673 household buildings. The entity that controls these variables is the CM, which
674 proposes feasible traces that can be fulfilled. Figure 14 depicts the perfor-
675 mance of the HPs from April 3rd to May 15th. The granularity of the monitored
676 data is two minutes, and the power is aggregated over the individual readings
677 of the five available buildings. For this field operation period, multiple activa-
678 tion traces were tested in several operation tests. They are represented separ-
679 ated by gaps in Figure 14. In the top panel, the difference between the power
680 trace to be tracked and the baseline forecasting is represented. As expected,
681 significant differences between these two time series are clearly appreciated.
682 The middle panel shows the differences between the active power, P^e , and
683 forecasting of the baseline power, P^b . As in the other graph, the differences
684 show that the heat pumps are following the trace up to a certain level and that
685 these traces have very different patterns than the BaU scenario.

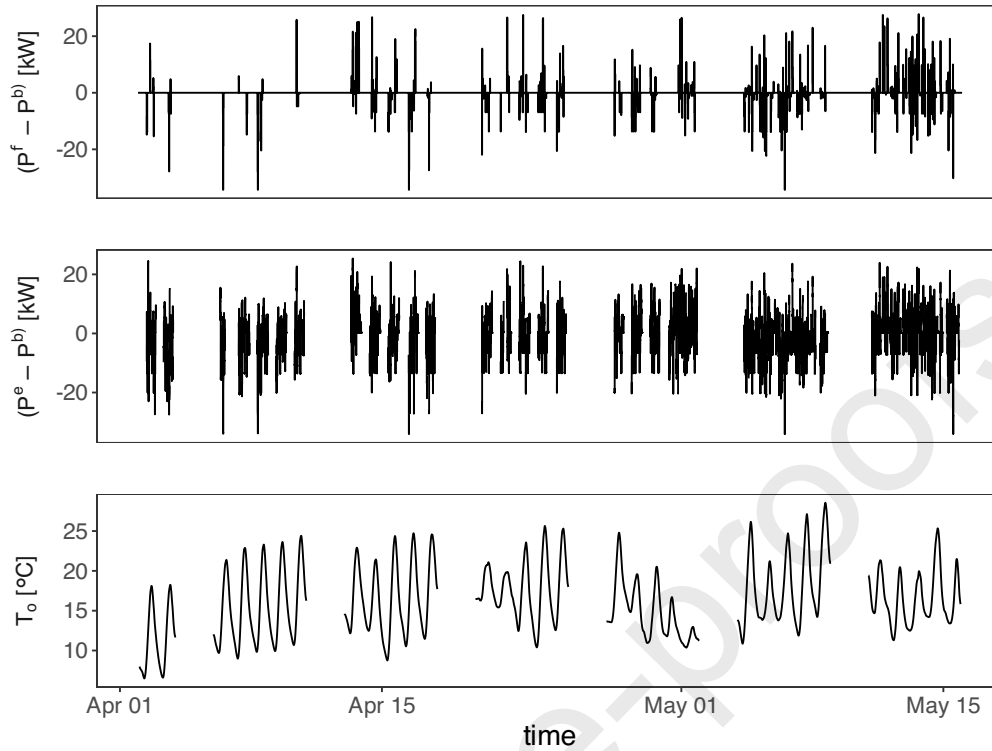


Figure 14 – Difference between the trace to be tracked, P^f , and the prediction of the baseline, P^b , versus the difference between the active power P^e and the baseline prediction of the baseline, P^b , and the 4-hour moving-averaged outdoor temperature in the Swiss pilot site. The tests in May were week-long tests

686 4.4. Energy flexibility evaluation and quantification

687 4.4.1. Training and validation of the flexibility models

688 For the case study where the activation variable is the Spanish day-ahead
 689 electricity price, a training and validation activation period was set up from
 690 March 29th to April 12th 2020. The flexibility model of this case study is de-
 691 fined in Equation 1. The training of the model to identify the regression pa-
 692 rameters was carried out using 90 % of the data. The remaining 10 % of data
 693 was used to validate the model with new data and then avoid model over-
 694 fitting. The Flexibility Function (FF) is finally inferred from this model. The
 695 top plot of Figure 15 depicts the training and validation periods with white
 696 and grey backgrounds, respectively. In this plot, the active power, P^e , is repre-
 697 sented by a black coloured line. The forecasting based on the flexibility model
 698 is represented by a red coloured line. It can be seen that no significant dif-
 699 ferences in residuals between the two periods are appreciated; therefore, it is
 700 confirmed that overfitting issues were avoided. Additionally, from the two
 701 bottom plots, the Auto Correlation Function, ACF, and the Partial Autocorre-
 702 lation Function, PACF, of the residuals of the training period, do not indicate
 703 autocorrelation in residuals. Therefore, they can be considered i.i.d, and the
 704 white noise condition is fulfilled. This is the requirement for a model to be
 705 considered valid.

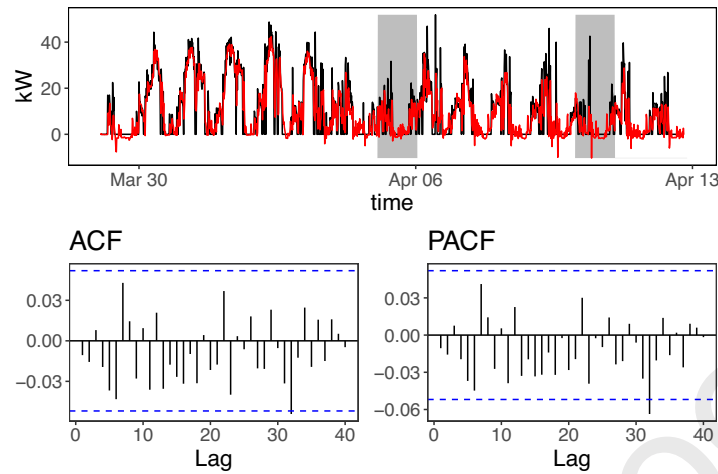


Figure 15 – Flexibility model for the Spanish pilot site: the upper graph is a comparison of the active power (black line) and the predicted one (red line); the lower graphs show the autocorrelation functions of the training period residuals

706 For the case study where the activation variable is the percentage of activa-
 707 tion time, in the German pilot case, a training and validation activation period
 708 was set up from February 15th to March 31st 2020. The flexibility model of this
 709 case study is defined in Equation 2. The training of the model was carried out
 710 using 90 % of the data. The remaining 10 % of the data was used to validate
 711 the model. In Figure 16, the upper graph depicts the training and validation
 712 periods with white and grey backgrounds, respectively. In this graph, the ac-
 713 tive power, P^e , is represented by a black coloured line. The forecasting based
 714 on the flexibility model is represented by a red coloured line. It can be seen
 715 that no significant differences in residuals between the two periods are appre-
 716 ciated. Although there are two significant spikes in time lags 3 and 12 in the
 717 bottom plots of the ACF and PACF, there is no clear indication of autocorrela-
 718 tion in residuals of the training period. Therefore, they can be considered as
 719 i.i.d. and then, the white noise condition is fulfilled for this flexibility model.

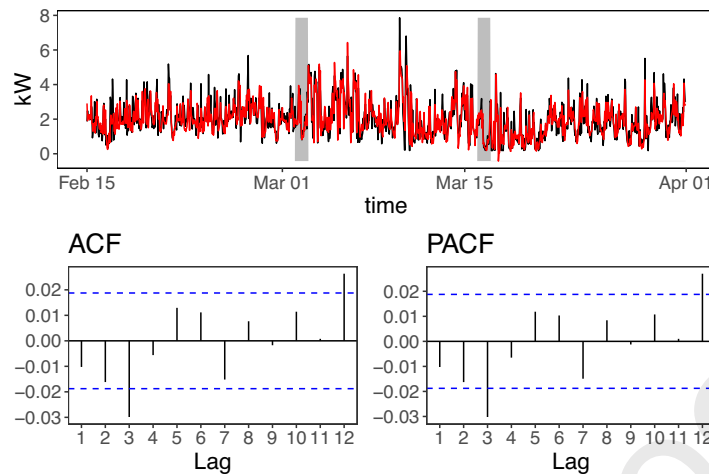


Figure 16 – Flexibility model for the German pilot site; the upper graph shows a comparison of the active power (black line) and the predicted one (red line); the lower graphs show the autocorrelation functions of the training period residuals

720 For the case study where the activation variable is the trace to be tracked,
 721 the Swiss pilot case, a training and validation activation period was set up
 722 from April 3rd to May 15th in the Swiss pilot site case study. The flexibility
 723 model of this case study is defined in Equation 3. The training of the model
 724 was carried out using 90 % of the data. The remaining 10 % of data was used
 725 to validate that the model. In Figure 17, the upper graph depicts the train-
 726 ing and validation periods with white and grey backgrounds, respectively. In
 727 this graph, the difference between the active power and the prediction of the
 728 baseline power in BaU, ($P^e - P^b$), is represented by a black coloured line. The
 729 forecasting based on the flexibility model is represented by a red coloured line.
 730 It can be seen that no significant differences in residuals are appreciated. Al-
 731 though there is one significant spike in time lag 15, in the bottom plots of the
 732 ACF and PACF, there is no clear indication of autocorrelation in residuals of
 733 the training period. Therefore, they can be considered i.i.d. The white noise
 734 condition is fulfilled for this flexibility model.

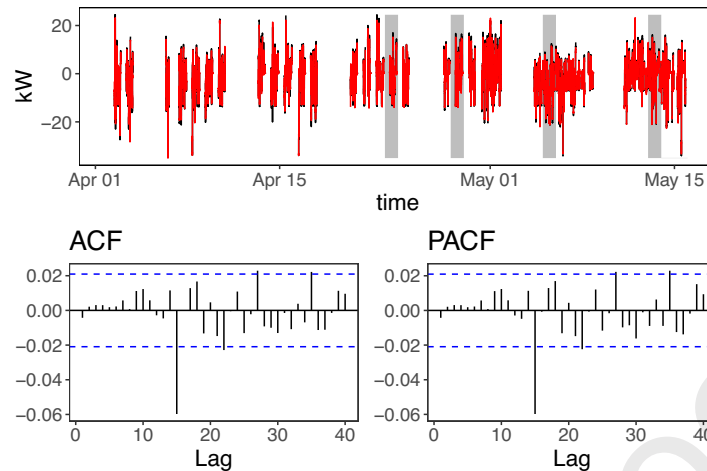


Figure 17 – Flexibility model for the Swiss pilot site; the upper graph shows a comparison of $(P^e - P^b)$ (black line) and the predicted one performed with the flexibility model (red line); the lower graphs show the autocorrelation functions of the training period residuals

735 4.4.2. Flexibility functions

736 Figure 18 and Figure 19 show the obtained flexibility functions, *FFs*, for the
 737 Spanish case study, where the activation variable is the electricity day-ahead
 738 Spanish spot market. The activation variable, the day-ahead electricity price,
 739 is normalized to activation and deactivation signals of 10 cents. Figure 18
 740 shows the obtained *FFs* due to positive signals of different lengths and two
 741 different outdoor temperature levels. The left column shows the *FFs* for out-
 742 door temperature ranges between 6.67 °C and 12.3 °C. The right column shows
 743 the *FFs* for outdoor temperature ranges between 12.3 °C and 21.5 °C. It can be
 744 seen that the flexibility decreases for low outdoor temperature ranges. When
 745 outdoor temperatures are between 6.67 °C and 12.3 °C, the average maximum
 746 deactivated power reaches -7 kW, and it remains for the first 30 minutes. Then,
 747 it increases to -3.5 kW from 30 to 45 minutes, and finally, it linearly increases
 748 to -1 kW after 100 minutes of the initial price change. Whereas, when out-
 749 door temperatures are between 12.3 °C and 21.5 °C, the maximum deactivated
 750 power reaches -11 kW for the first 15 minutes; it decreases to -14 kW after
 751 30 minutes, and finally, it increases up to -1 kW after 100 minutes of the price
 752 change. The rebound effect achieves the same maximum power levels for both
 753 temperature ranges but in positive. They start just when the activation signal
 754 finishes and reach the maximum level within the first 30 minutes after the ac-
 755 tivation signal ends. Considering the maximum available power of the two
 756 heat pumps of 60 kW, this represents maximum flexibility between 11 % and
 757 23 %, with a rebound of the same level, for low and high outdoor temperature
 758 ranges, respectively. It can also be concluded that the estimated period where
 759 major energy shifts could be done is the starting 30 minutes after the price sig-
 760 nal is triggered, in both outdoor temperature ranges. This conclusion is closely
 761 related to the thermal capacity of the water storage tank, which is 3,500 litres,

762 and the permitted water tank temperature variation, which is constrained by
 763 the indoor comfort conditions in the offices, shops and the local food market.

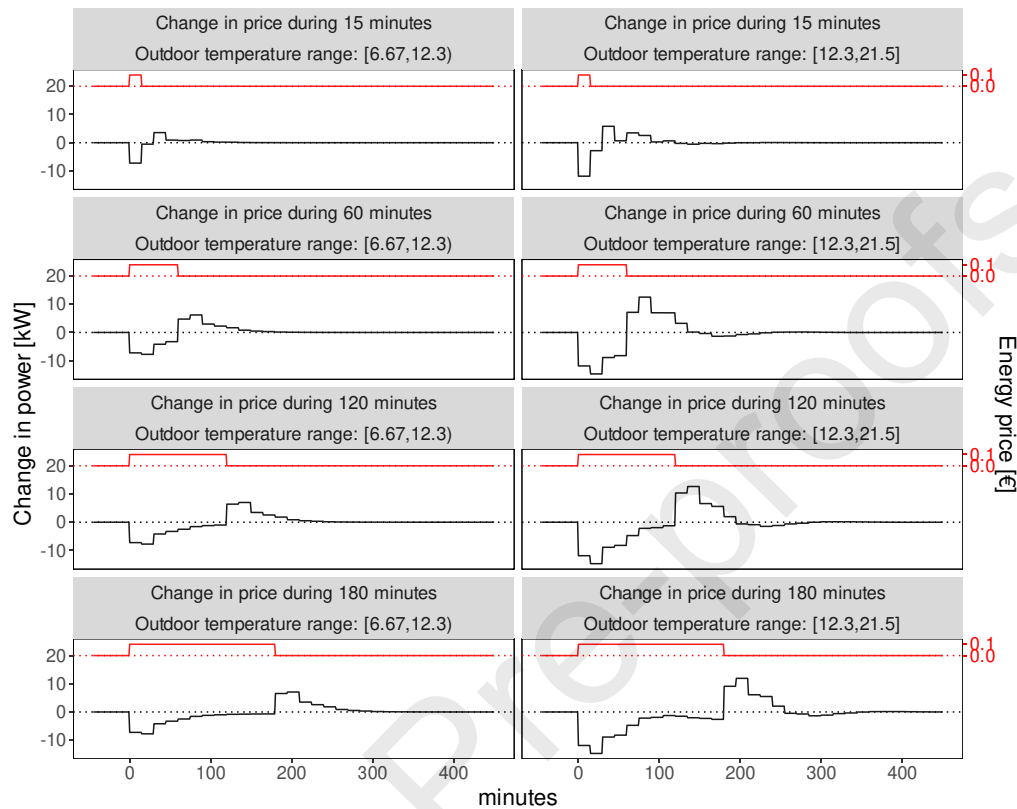


Figure 18 – *FFs* of the Spanish case study for positive changes of the spot market price

764 Figure 19 depicts the obtained *FFs* due to negative signals of different
 765 lengths and the same outdoor temperature levels. The flexibility performance
 766 is identical to the case of positive activation but another way around. The re-
 767 bound effect achieves the same maximum power levels for both temperature
 768 ranges but in negative. The same conclusions as in the case of positive signals
 769 can be deducted.

770

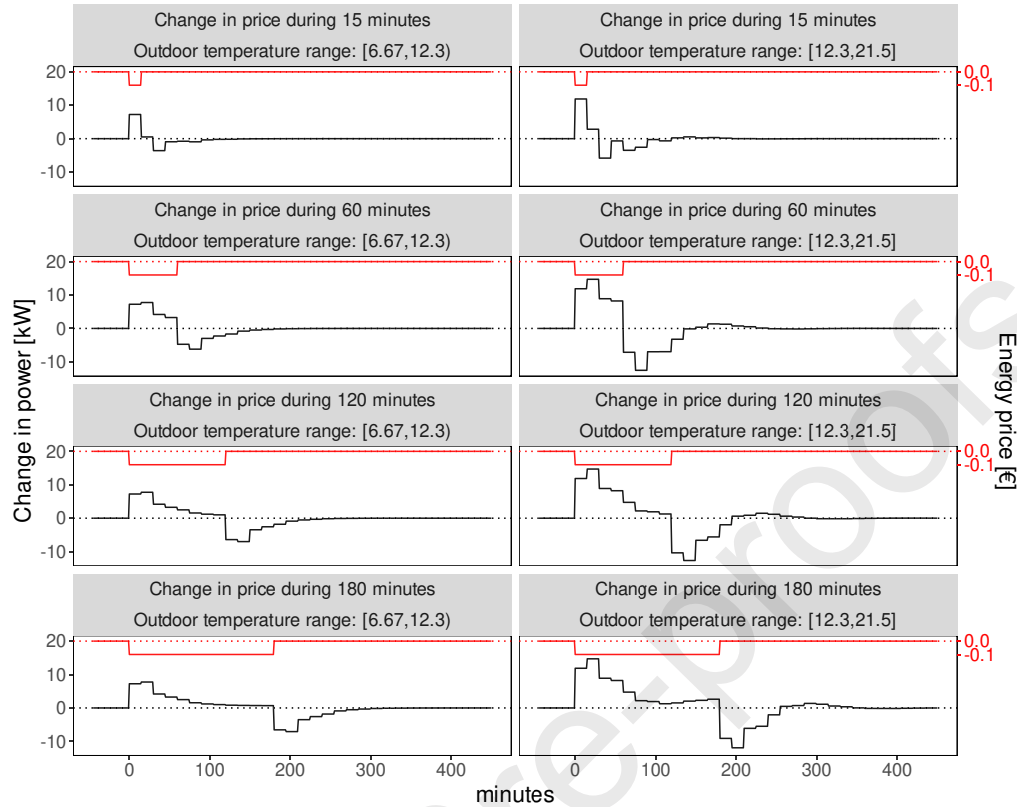


Figure 19 – FFs of the Spanish case study for negative changes of the spot market price

771 Figure 20 shows the Flexibility Functions, *FFs* of four heat pumps and their
 772 aggregated power, of the case study where the activation signal is the percentage
 773 of activation within an activation period. The activation variable has been
 774 normalized to 100 % activation time. The Figure 20 represents the *FFs* of each
 775 heat pump/building, named as 20, 22, 24 and 25 in the legend, and the ag-
 776 gregated flexibility of all of them, named as "all" in the legend. Every plot
 777 shows a *FF* for several activation periods ranging from 1 h to 4 h. From this
 778 Figure, multiple insights in relation to the achieved flexibility of a cluster of
 779 heat pumps can be extracted. The total amount of power flexibility for the
 780 cluster of 4 buildings reaches 2.8 kW -on average- for the first hour of acti-
 781 vation. And from there, it decreases to 2.3 kW for the second and the third
 782 hours of activation. If the activation period is extended to four hours, maxi-
 783 mum flexibility decreases to 2 kW. Considering a maximum available power
 784 of the four heat pumps of 10.9 kW, this represents maximum flexibility of 25
 785 % for the first hour, 20 % for three hours and 18 % for four hours. After the
 786 activation periods, Figure 20 depicts a long wave rebound effect of about 20
 787 % of the total active power. Nonetheless, around 70 % of this rebound takes
 788 place within the first 3 h after the activation period ends.

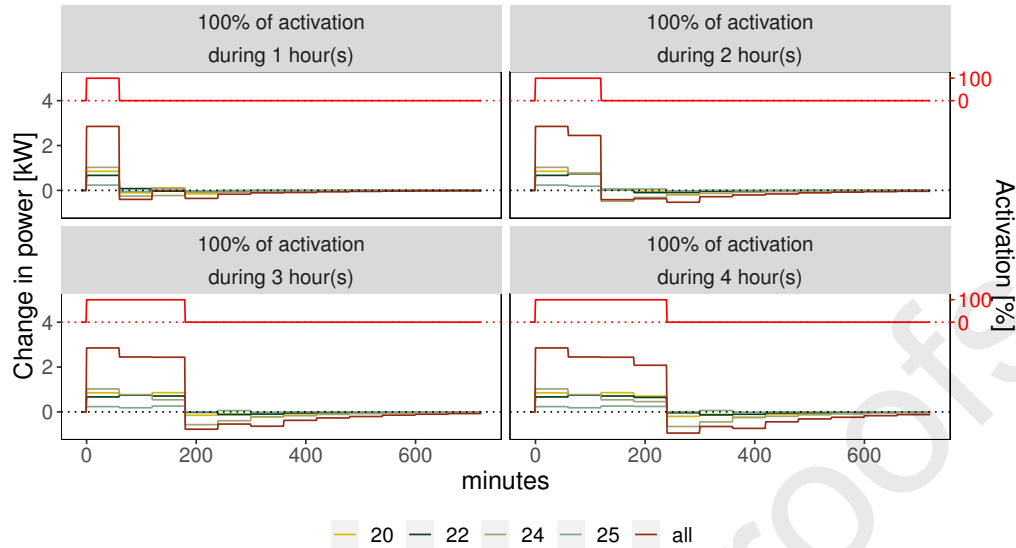


Figure 20 – FFs of 4 heat pumps of the German case study

789 In Figure 20, it can also be seen that the reactions of buildings 20, 22 and
 790 24 are quite similar and also very similar to the aggregated FFs. However, a
 791 very different behaviour happens in building 25 since it seems this heat pump
 792 is not activated. This may be due to less flexible indoor comfort conditions.

793 Figure 21 shows the FFs for the swiss case study. In this case, the activa-
 794 tion variable is a power trace that should be tracked, and the flexibility is as-
 795 sessed as the deviation towards the traces and towards the predicted baseline
 796 in the BaU scenario. In Figure 21, the left Y-axis describes the change in power
 797 ($P^e - P^b$) and the right Y-axis describes the change in power due to the trace
 798 negotiated with the commercial aggregator ($P^f - P^b$). The flexibility is anal-
 799 ysed for two different outdoor temperature levels; low-to-mid range [6.5 °C,
 800 15.7 °C] in yellow and mid-to-high [15.7 °C, 28.5 °C] in black. Two types of nor-
 801 malized activation traces of 1 kW (e.g. red signal [-1, 0, 1]) are tested: (1) *Neg-*
 802 *ative*, when the consumption is lower than the baseline, and (2) *Positive*, when
 803 the consumption is higher than the baseline. The terms *Negative* and *Positive*
 804 used here have to be differentiated from the existing positive (Upward) and
 805 negative (Downward) reserve services defined in the market regulation and
 806 provided by conventional generators. In this methodology, the term *Positive*
 807 refers to an increase in power consumption compared to the baseline, which,
 808 from a market perspective, is equivalent to a decrease in power production
 809 (negative reserve).

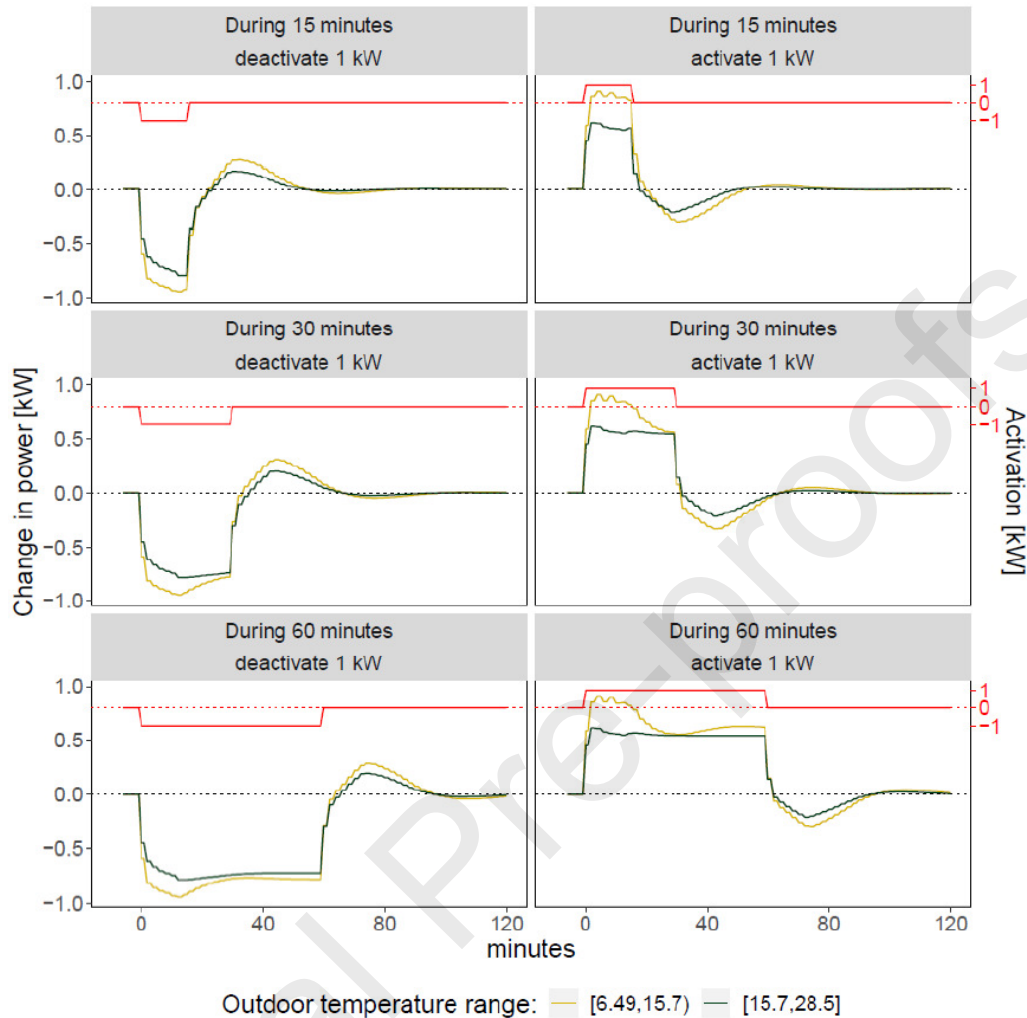


Figure 21 – Flexibility Function (FF) of a 5-buildings cluster in Naters

810 In the case of tracking *negative* activation traces (left panels) in low-to-mid
 811 outdoor temperature levels (yellow lines), the active power follows 80 to 90
 812 % of the power trace to be tracked for the first 15 minutes, reaching the max-
 813 imum deactivation peak (98 %) after 13 minutes. Then, the deactivation de-
 814 creases to 75 % after 30 minutes, maintaining this percentage for 30 minutes
 815 more. When tracking a *positive* activation trace, the actual power follows 80-
 816 90% of the theoretical activation for the first 15 minutes, then, it linearly de-
 817 creases to 50 % after 30 minutes and maintains this percentage, with a small re-
 818 bound (+10 %), up to the 60 minutes. This means that heat pumps involved in
 819 this case study, when the outdoor temperature is in the low-to-mid range, can
 820 provide the amount of flexibility required by the commercial aggregator for
 821 the first 15 minutes. Still, then, the limited availability of thermal energy stor-
 822 age in the building (either for SH or DHW) does not allow for full activation
 823 compliance. In both outdoor temperature levels, the rebound effect reaches
 824 up to 30 % change in power. It starts just after the activation/deactivation of
 825 the trace, and its peak is after approximately 13 minutes. In the case of mid-

826 to-high outdoor temperatures levels, the flexibility peak of the first 15 minutes
 827 no longer exists in both *negative* and *positive* traces to be tracked. This can be
 828 explained mainly because at these temperature levels; the buildings have less
 829 thermal storage capacity and hence less energy flexibility to keep the indoor
 830 comfort within the user-defined comfort boundaries. In this case, the system
 831 which can still provide a certain level of flexibility is the DHW system, which
 832 is thermostatically controlled by the water tank temperature set points. The
 833 average compliance of tracking the trace is 60 % along the 60 minutes of ac-
 834 tivation in the case of positive and 75 % in the case of negative traces. The
 835 rebound effects follow the same path as in the lower temperatures case but
 836 with smaller peaks of around 25 % of the change in power.

837 5. Discussion

838 Some specific conclusions can be drawn from the operation of the DR ser-
 839 vices in each of the three pilot sites:

- 840 • The direct load control of the heat pumps of the Spanish pilot site achieved
 841 18 % of accumulated cost savings at the end of the testing period (2
 842 weeks). This is a promising result to demonstrate the benefits of opti-
 843 mising the operational costs of heat pumps through augmented perfor-
 844 mance with price information from the wholesale market forecast data.
- 845 • In the German pilot site, it was demonstrated that using the flexibility
 846 of the heat pumps allowed to optimize the heating energy cost on the
 847 day-ahead energy market. This flexibility also enabled balancing a BRP's
 848 portfolio and optimization on the balancing market. With a limited num-
 849 ber of heat pump assets and only ON/OFF control, it was impossible to
 850 deliver linear power ramps based on the stacking of heat pumps.
- 851 • In the Swiss pilot site, a success of 91 % heat pump activation for the
 852 transactive DR approach and 50-95 % fulfilment of the activation traces
 853 was achieved for the testing period. The results are strongly correlated
 854 with the external temperature. Mid-range outdoor temperature condi-
 855 tions offered more flexibility, as highlighted by the higher activation suc-
 856 cess and the *FF* closer to 100 % of the theoretical activation.

857 The developed standard methodology for assessing the flexibility allowed
 858 to compare results from the different DR use cases and gave the necessary
 859 support for cross-comparison of the most significant energy flexibility indica-
 860 tors. Some specific conclusions can be deduced for the achieved flexibility in
 861 each pilot site:

862 5.1. Spanish case study

863 Considering the peak power of the heat pump system, the maximum flex-
 864 ibility achieved was between 11 % and 23 %, depending on low or high out-
 865 door temperature ranges, respectively. A contrary rebound effect at the same
 866 level was achieved in both cases. Table 1 summarizes the achieved active flex-
 867 ibility for this pilot site:

Table 1 – Achieved active power and rebound effect defined by the *FF* in the Spanish use case

Activation time [min]	Maximum change in power [kW]	Maximum power rebound [kW]
Flexibility with low outdoor temperature [$6.6\text{ }^{\circ}\text{C} \leq T \leq 12.3\text{ }^{\circ}\text{C}$]		
positive/negative change of price		
$t \leq 30$	-7/7	6/-6
$30 \leq t \leq 45$	-3.5/3.5	3/-3
$45 \leq t \leq 100$	linear increase	linear decay
$t > 100$	-1/1	0.8/-0.8
Flexibility with high outdoor temperature [$12.3\text{ }^{\circ}\text{C} \leq T \leq 21.5\text{ }^{\circ}\text{C}$]		
Positive/negative change of price		
$t \leq 15$	-11/11	8/-8
$t \leq 30$	-14/14	12/-12
$30 \leq t \leq 45$	-8/8	6/-6
$45 \leq t \leq 100$	linear increase	linear decay
$t > 100$	-1/1	0.8/-0.8

868 5.2. German case study

869 In the German pilot site, considering a maximum aggregated power of 10.9
870 kW, 25 % of flexibility was achieved for the first hour. For activation of three
871 hours, it was reduced to 20 %, and it finally decreased to 18 % for activation
872 of four hours. A long wave rebound effect of about 20 % of the total activated
873 power appears in all cases. However, around 70 % of the total rebound effect
874 occurs within the first 3 h after the activation period ends. Table 2 summarizes
875 achieved active flexibility for this pilot site:

Table 2 – Achieved active power and rebound effect defined by the *FF* in the German use case

Activation time [min]	Maximum change in power [kW]	Maximum power rebound [kW]
Flexibility under a 100 % positive activation signal		
$t \leq 60$	2.8	-0.8
$60 \leq t \leq 180$	2.3	-0.5
$180 \leq t \leq 240$	2	exponential increase
$t > 240$	—	0.0

876 5.3. Swiss case study

877 The heat pumps involved in the Swiss case study can provide the amount
878 of flexibility required by the commercial aggregator for the first 15 minutes.
879 Still, then, the limited availability of thermal energy storage in the buildings
880 does not allow for full activation compliance. In both outdoor temperature
881 levels, the rebound effect reaches up to 30 % change in power. The average
882 compliance of tracking the trace is 60 % along the 60 minutes of activation in

883 the case of positive activation traces and 75 % in the case of negative ones.
884 Table 3 summarises the achieved active flexibility for this pilot site:

Table 3 – Achieved active power and rebound effect defined by the *FF* in the Swiss use case

Activation time [min]	Maximum change in power [%]	Maximum power rebound [%]
Flexibility with low outdoor temperature [6.49 °C ≤ T ≤ 15.7 °C]		
positive/negative trace to be followed		
t ≤ 15	85/-98	-40/40
15 ≤ t ≤ 30	linear decrease	linear increase / linear decay
30 ≤ t ≤ 60	60/-75	0/0
Flexibility with high outdoor temperature [15.7 °C ≤ T ≤ 28.5 °C]		
Positive/negative change of price		
t ≤ 15	60/-75	-25/25
15 ≤ t ≤ 30	60/-75	linear increase / linear decay
30 ≤ t ≤ 60	60/-75	0/0

885 6. Conclusions

886 This study confirms that thermostatically controlled heat pumps represent
887 a huge potential for DR flexibility. Furthermore, it is possible to manage clusters
888 of heat pumps to respond to requests for DR flexibility. In addition, it has
889 been proven that forecasting and optimization algorithms can be tailored to
890 the particularities of each system configuration (e.g. HP interface, HP installation,
891 and temperature sensors).

892 The operation tests performed in three European pilot sites demonstrated
893 that the flexible operation of heat pumps in the field is possible and can be
894 leveraged for multiple flexibility services or energy markets. However, several
895 problems need to be addressed with most legacy systems. In general, those
896 systems do not provide fully interoperable connectivity with the heat pump,
897 resulting in constraints to the control and less flexible systems. Additionally,
898 it has been confirmed that outdoor conditions, configured set points and the
899 available thermal storage, both in hot water tanks or inertia in the building,
900 determine the duration for which the heat pump can be switched on or off. Another
901 important conclusion from this research is that a new player, called the Cluster
902 Manager (CM), is essential to assure a successful operation of the DR services
903 in real market scenarios.

904 7. Future work

905 Although the developed methodology to assess the flexibility in the different
906 pilot sites shows promising outcomes to demonstrate its scalability and wider
907 application, some procedures' limitations to determine the *FF* need

908 further research. These limitations are mainly related to the non-accurate in-
 909 corporation of the dynamic variability of the flexibility and the dependencies
 910 between the active energy and the activation variable. Both have been ad-
 911 dressed in this research by including the autoregressive terms in the model.
 912 However, this procedure is not accurate enough and can miss some of the non-
 913 linearities. Therefore, some improvements should be addressed. As an ex-
 914 ample, recent papers [21] opened alternative methodologies to address these
 915 non-linearities in price-based DR schemes. These complementary approaches
 916 should be investigated in real practice experiences. Finally, simpler and more
 917 cost-efficient computational methods to evaluate the flexibility potential of
 918 large amounts of buildings and HVAC systems need to be further developed
 919 to assure a seamless connection with commercial practices of aggregators and
 920 cluster managers in already existing European energy flexible markets.

921 8. Acknowledgements

922 This work emanated from research conducted with the financial support
 923 of the European Commission through the H2020 project Sim4Blocks, grant
 924 agreement 695965. Dr. J. Cipriano thanks the Ministerio de Ciencia e Inno-
 925 vación for the Juan de la Cierva Incorporación grant.

926 References

- 927 [1] C. Lo and N. Ansari, "Decentralized Controls and Communications for
 928 Autonomous Distribution Networks in Smart Grid," IEEE Transactions
 929 on Smart Grid, vol. 4, pp. 66–77, Mar. 2013. Conference Name: IEEE
 930 Transactions on Smart Grid.
- 931 [2] P. D. Lund, J. Lindgren, J. Mikkola, and J. Salpakari, "Review of energy
 932 system flexibility measures to enable high levels of variable renewable
 933 electricity," Renewable and Sustainable Energy Reviews, vol. 45, pp. 785–
 934 807, May 2015.
- 935 [3] D. S. Kirschen, A. Rosso, J. Ma, and L. F. Ochoa, "Flexibility from the
 936 demand side," in 2012 IEEE Power and Energy Society General Meeting,
 937 pp. 1–6, July 2012. ISSN: 1944-9925.
- 938 [4] S. Weitemeyer, D. Kleinhans, T. Vogt, and C. Agert, "Integration of Re-
 939 newable Energy Sources in future power systems: The role of storage,"
 940 Renewable Energy, vol. 75, pp. 14–20, Mar. 2015.
- 941 [5] P. Denholm, E. Ela, B. Kirby, and M. Milligan, "Role of Energy Stor-
 942 age with Renewable Electricity Generation," Tech. Rep. NREL/TP-6A2-
 943 47187, National Renewable Energy Lab. (NREL), Golden, CO (United
 944 States), Jan. 2010.
- 945 [6] P. Siano, "Demand response and smart grids—A survey," Renewable and
 946 Sustainable Energy Reviews, vol. 30, pp. 461–478, Feb. 2014.

- 947 [7] V. Giordano, A. Meletiou, C. Felix, A. Mengolini, M. Ardelean, G. Fulli,
948 M. Sánchez, and C. Filiou, "Smart grid projects in Europe: lessons
949 learned and current developments 2012 update. .," 2013.
- 950 [8] Z. Ma, J. D. Billanes, and B. N. Jørgensen, "Aggregation Potentials for
951 Buildings—Business Models of Demand Response and Virtual Power
952 Plants," *Energies*, vol. 10, p. 1646, Oct. 2017. Number: 10 Publisher: Mul-
953 tidisciplinary Digital Publishing Institute.
- 954 [9] X. Cao, X. Dai, and J. Liu, "Building energy-consumption status world-
955 wide and the state-of-the-art technologies for zero-energy buildings dur-
956 ing the past decade," *Energy and Buildings*, vol. 128, pp. 198–213, Sept.
957 2016.
- 958 [10] T. M. Lawrence, M.-C. Boudreau, L. Helsen, G. Henze, J. Mohammad-
959 pour, D. Noonan, D. Patteuw, S. Pless, and R. T. Watson, "Ten questions
960 concerning integrating smart buildings into the smart grid," *Building
961 and Environment*, vol. 108, pp. 273–283, Nov. 2016.
- 962 [11] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad,
963 "Model Predictive Control (MPC) for Enhancing Building and HVAC
964 System Energy Efficiency: Problem Formulation, Applications and Op-
965 portunities," *Energies*, vol. 11, p. 631, Mar. 2018. Number: 3 Publisher:
966 Multidisciplinary Digital Publishing Institute.
- 967 [12] P. Kohlhepp, H. Harb, H. Wolisz, S. Waczowicz, D. Müller, and V. Hagen-
968 meyer, "Large-scale grid integration of residential thermal energy stor-
969 ages as demand-side flexibility resource: A review of international field
970 studies," *Renewable and Sustainable Energy Reviews*, vol. 101, pp. 527–
971 547, Mar. 2019.
- 972 [13] J. L. Boudec and D. Tomozei, "Demand response using service curves,"
973 in *2011 2nd IEEE PES International Conference and Exhibition on
974 Innovative Smart Grid Technologies*, pp. 1–8, Dec. 2011. ISSN: 2165-4824.
- 975 [14] C. Eid, P. Codani, Y. Perez, J. Reneses, and R. Hakvoort, "Managing elec-
976 tric flexibility from Distributed Energy Resources: A review of incentives
977 for market design," *Renewable and Sustainable Energy Reviews*, vol. 64,
978 pp. 237–247, Oct. 2016.
- 979 [15] A. Arteconi, A. Mugnini, and F. Polonara, "Energy flexible buildings: A
980 methodology for rating the flexibility performance of buildings with elec-
981 tric heating and cooling systems," *Applied Energy*, vol. 251, p. 113387,
982 Oct. 2019.
- 983 [16] C. Finck, R. Li, R. Kramer, and W. Zeiler, "Quantifying demand flexibility
984 of power-to-heat and thermal energy storage in the control of building
985 heating systems," *Applied Energy*, vol. 209, pp. 409–425, Jan. 2018.
- 986 [17] G. Reynders, R. Amaral Lopes, A. Marszal-Pomianowska, D. Aelenei,
987 J. Martins, and D. Saelens, "Energy flexible buildings: An evaluation

- 988 of definitions and quantification methodologies applied to thermal stor-
989 age," Energy and Buildings, vol. 166, pp. 372–390, May 2018.
- 990 [18] R. El Geneidy and B. Howard, "Contracted energy flexibility characteris-
991 tics of communities: Analysis of a control strategy for demand response,"
992 Applied Energy, vol. 263, p. 114600, Apr. 2020.
- 993 [19] A. Bampoulas, M. Saffari, F. Pallonetto, E. Mangina, and D. P. Finn, "A
994 fundamental unified framework to quantify and characterise energy flex-
995 ibility of residential buildings with multiple electrical and thermal energy
996 systems," Applied Energy, vol. 282, p. 116096, Jan. 2021.
- 997 [20] R. G. Junker, A. G. Azar, R. A. Lopes, K. B. Lindberg, G. Reynders, R. Re-
998 lan, and H. Madsen, "Characterizing the energy flexibility of buildings
999 and districts," Applied Energy, vol. 225, pp. 175–182, Sept. 2018.
- 1000 [21] R. G. Junker, C. S. Kallesøe, J. P. Real, B. Howard, R. A. Lopes, and
1001 H. Madsen, "Stochastic nonlinear modelling and application of price-
1002 based energy flexibility," Applied Energy, vol. 275, p. 115096, Oct. 2020.
- 1003 [22] "Was ist der Intraday-Handel von Strom?. [https://www.next-](https://www.nextkraftwerke.de/wissen/intraday-handel/)
1004 [kraftwerke.de/wissen/intraday-handel/](https://www.nextkraftwerke.de/wissen/intraday-handel/)," May 2014.
- 1005 [23] "Regelenergie & Regelleistung - was ist das?. [https://www.next-](https://www.nextkraftwerke.de/wissen/regelenergie/)
1006 [kraftwerke.de/wissen/regelenergie/](https://www.nextkraftwerke.de/wissen/regelenergie/)," Nov. 2010.
- 1007 [24] "Regelenergiemarkt im Umbruch II – von Preisspitzen zum Misch-
1008 preisverfahren. [https://blog.energybrainpool.com/regelenergiemarkt-](https://blog.energybrainpool.com/regelenergiemarkt-im-umbruch-ii-von-preisspitzen-zum-mischpreisverfahren/)
1009 [im-umbruch-ii-von-preisspitzen-zum-mischpreisverfahren/](https://blog.energybrainpool.com/regelenergiemarkt-im-umbruch-ii-von-preisspitzen-zum-mischpreisverfahren/)," Feb. 2020.
- 1010 [25] "Day-Ahead-Handel - was ist das?. [https://www.next-](https://www.nextkraftwerke.de/wissen/day-ahead-handel/)
1011 [kraftwerke.de/wissen/day-ahead-handel/](https://www.nextkraftwerke.de/wissen/day-ahead-handel/)," 2014.
- 1012 [26] "What is Intraday Trading?. [https://www.next-](https://www.nextkraftwerke.be/en/knowledge-hub/intraday-trading/)
1013 [kraftwerke.be/en/knowledge-hub/intraday-trading/](https://www.nextkraftwerke.be/en/knowledge-hub/intraday-trading/)."
- 1014 [27] F. Ocker, S. Braun, and C. Will, "Design of European balancing power
1015 markets," in 2016 13th International Conference on the European Energy
1016 Market (EEM), pp. 1–6, June 2016. ISSN: 2165-4093.
- 1017 [28] M. D. Galus, "Smart grid roadmap and regulation approaches in Switzer-
1018 land," CIREN - Open Access Proceedings Journal, vol. 2017, pp. 2906–
1019 2909, Oct. 2017. Publisher: IET Digital Library.
- 1020 [29] J. Cipriano, G. Mor, D. Chemisana, D. Pérez, G. Gamboa, and X. Cipri-
1021 ano, "Evaluation of a multi-stage guided search approach for the cali-
1022 bration of building energy simulation models," Energy and Buildings.
1023 ENB-D-14-01064R1.
- 1024 [30] P. Bacher and H. Madsen, "Identifying suitable models for the heat dy-
1025 namics of buildings," Energy and Buildings, vol. 43, pp. 1511–1522, 2011.

- 1026 [31] P. Bacher, H. Madsen, H. A. Nielsen, and B. Perers, "Short-term heat
1027 load forecasting for single family houses," Energy and Buildings, vol. 65,
1028 pp. 101–112, Oct. 2013. 00045.
- 1029 [32] B. Grillone, S. Danov, A. Sumper, J. Cipriano, and G. Mor, "A review of
1030 deterministic and data-driven methods to quantify energy efficiency sav-
1031 ings and to predict retrofitting scenarios in buildings," Renewable and
1032 Sustainable Energy Reviews, vol. 131, p. 110027, Oct. 2020.
- 1033 [33] F. Amblard, J. Page, R. Parakkal Menon, M. Brennenstuhl, and
1034 J. Cipriano, "D.3.1-Optimization strategies for the use case scenarios.
1035 SIM4BLOCKS H2020 project. Grant agreement n° 695965.," Tech. Rep.
1036 Deliverable D.3.1., Sim4Blocks project, 2018.
- 1037 [34] T. Hastie and R. Tibshirani, "Generalized additive models," p. 10, 1990.
- 1038 [35] F. D’Ettorre, M. Brennenstuhl, A. Kathirgamanathan, M. D. Rosa,
1039 M. Yadack, U. Eicker, and D. P. Finn, "A set of comprehensive indicators
1040 to assess energy flexibility: a case study for residential buildings," E3S
1041 Web of Conferences, vol. 111, p. 04044, 2019. Publisher: EDP Sciences.
- 1042 [36] L. Romero Rodríguez, M. Brennenstuhl, M. Yadack, P. Boch, and
1043 U. Eicker, "Heuristic optimization of clusters of heat pumps: A simu-
1044 lation and case study of residential frequency reserve," Applied Energy,
1045 vol. 233-234, pp. 943–958, Jan. 2019.
- 1046 [37] T. Schütz, H. Harb, R. Streblow, and D. Müller, "Comparison of mod-
1047 els for thermal energy storage units and heat pumps in mixed integer
1048 linear programming. ECOS 2015-28th International Conference on Effi-
1049 ciency, Cost, Optimization, Simulation and Environmental Impact of En-
1050 ergy Systems.," (Pau, France), July 2015.