Erasmus University Rotterdam
Erasmus School of Economics
Bachelor Thesis Econometrie en Operationele Research

Optimal demand response management for HEMS using real-time pricing

Sophie van Groesen (581870)

Supervisor:	(Ece) E Karakoyun
Second assessor:	R. Spliet
Date Final Version:	1st July 2024

Abstract

Given the growing complexity of modern energy systems, Demand Response (DR) models are becoming increasingly important for successful energy management systems. The effect of renewable energy sources, such as solar and wind, on market pricing emphasizes the crucial need to optimize the use of existing energy resources. This paper investigates the use of DR models in a comprehensive and modular set of Mixed-Integer Linear Programming (MILP) models integrated in the Autonomous Home Energy Management Systems (AHEMS). This allows for integrated optimization of all energy resources, including shiftable loads, electric water heaters (EWH), air conditioners (AC) based on indoor temperature dynamics, and static and electric vehicle (EV) batteries. Additionally, the use of the static and EV batteries are investigated a little further. Furthermore, we explore if the costs can be improved when we forecast the day-ahead market price of energy, taking the renewable energy generated in solar and wind parks into account for an accurate forecast. In this paper, the forecasting performance of the Auto Regressive (AR), Seasonal autoregressive integrated moving average (SARIMA), Lasso, Ridge, and Random Forest (RF) methods are evaluated using the root-mean-square error and the Mean Absolute Error (MAE). Additionally, we will do a Diebold-Mariano test, compared to the AR benchmark. The results show that consumers may lower their net power expenses or even profit by using an energy management system based on the recommended models included in the paper. The static battery is very profitable since it stores energy during low-cost periods to be used within the household during high-cost periods. Integrating the hourly forecasting of the day-ahead energy market pricing, when applying the Ridge model, significantly reduces electricity bills even further.

1 Introduction

Given the growing complexity of modern energy management systems, the significance of Demand Response (DR) models increases, as they play a vital role in the management of power networks (Siano, 2014). Furthermore, renewable energy sources, such as solar and wind energy generated in solar parks and wind farms, influences market prices since they are irregular (Winkler et al., 2016). As a result, optimizing the use of available energy resources becomes crucial.

Concerns about the possible effects of greenhouse gas emissions are increasing in our society, as is understanding of the link between energy usage and climate change. Increased energy costs, driven by increased demand and simultaneously a limiting supply, serve to reinforce this knowledge. Energy costs have risen, particularly since certain countries' energy sector deregulation, presenting new challenges and opportunities (Bhattacharya et al., 2012). This opens up new prospects for home energy management systems (HEMS) in DR markets (Beaudin and Zareipour, 2015). HEMS are DR instruments that optimize a household's energy consumption on behalf of consumers, thus shifting and reducing the demand for energy. HEMS optimize the consumption and production schedule, while taking into account various factors such as energy costs, consumer satisfaction, environmental considerations and load profiles.

Several approaches have been offered in the current literature, including optimization and heuristic solution algorithms, for developing efficient schemes and making energy consumption and production decisions. However, assessing the effectiveness of different approaches is complicated by the various modeling elements, such as device type, timing, and targets (Beaudin and Zareipour, 2015). DR improves the adaptability of the customer and thereby the engagement, offering a wide range of potential benefits for the market effectiveness and the reliability of the energy system. For instance, DR lowers peak demand to postpone the need for network upgrades, in addition DR also reduce the total cost of investment for installations and capital. Vardakas et al. (2014) demonstrates that DR is the most economically efficient and dependable method for managing sudden increases in energy demand and motivating individuals to modify their energy consumption habits in response to factors like fluctuating power costs. The categorization of DR schemes based on control mechanisms, stimuli and decision variables provides an overview of the many optimization models proposed for effective DR management.

Improved DR models are considered a key solution to society's energy issue, especially with smart grids (Ipakchi and Albuyeh, 2009). This paper investigates the use of such methods including Mixed-Integer Linear Programming (MILP) models integrated in the Autonomous Home Energy Management Systems (AHEMS). We expand the work of Antunes et al. (2022), in which they developed a comprehensive and modular set of MILP models of appliance operation to be seamlessly integrated in AHEMS. This allows for integrated optimization of all energy resources, including shiftable loads, electric water heaters (EWH), air conditioners (AC) based on indoor temperature dynamics, and static and electric vehicle (EV) batteries. Following their research, we will focus on the integrated model, using all these loads to optimize. Additionally we explore if we can improve the costs when we forecast the day-ahead market price of energy, taking the solar and wind energy generation into account for an accurate forecast. We will use hourly

data of the day-ahead market price obtained from ENTSO-E[1], and the hourly data of the solar energy of solar parks, the wind energy on land and the wind energy in the sea obtained from the "Nationale Energie Dashboard"[2]. We will compare five different forecasting methods, explained in Section 4.2. The Auto Regressive (AR) method, the Seasonal Autoregressive Integrated Moving Average (SARIMA) method, two penalized regression methods; Lasso and Ridge and as a machine learning method the Random forest. We will compare the forecasts using the root-mean-square error (RMSE) and the Mean Absolute Error (MAE). Additionally, we will do a Diebold-Mariano test to determine which forecast model to use to incorporate the hourly forecasts into the integrated model.

Effective energy management requires charging static batteries during periods of abundant energy, as indicated by low market prices, and discharging during periods of shortage, as shown by high market prices. This technique reduces dependency on coal-fired power facilities and minimizes the environmental implications of growing CO_2 emissions. For this reason, we want to explore how to optimize the use of the static battery and how the static and EV battery will affect the cost of the consumer.

With this research we want to contribute the development of reliable AHEMS that can be parameterized with user preferences in order to optimally manage all energy resources on behalf of customers who are benefiting from DR programs. Additionally, we want to find an accurate forecasting method to incorporate hourly day-ahead energy market prices into the integrated model to obtain a minimal and accurate cost. We attempt to discover an answer to the following research questions: *What benefits do the static and EV batteries have in this integrated model? What type of forecasting model can accurately represent day-ahead market energy prices that are highly affected by renewable energy sources? And how does an accurate hourly day ahead market energy price forecast effect the cost of the consumer and consumer patterns?*

We found that consumers may lower their net power expenses or even profit by using an energy management system based on the recommended models. The static battery is very profitable since it stores energy during low-cost periods to be used in the household during high-cost periods. Integrating the Ridge model for hourly forecasting of the day-ahead energy market price significantly reduces electricity bills even further.

The remainder of this paper is organized as follows: First, Section 2 explains the literature review. Next, Section 3 describes the data used, followed by a description and the notation of the methods used for this research in Section 4. We then present our results in Section 5. We finally provide a conclusion in Section 6 and discuss the limitations of our research and interesting directions for future research in Section 7.

2 Literature

This section looks further into the existing literature on DR programs and (A)HEMS to identify commonly used methods and directions for future research.

The study conducted by Siano (2014) performs a thorough investigation in the benefits of DR models in the smart grid. The analysis covers various aspects, including enabling technologies

[1] Link to Day Ahead market Prices
[2] Link to renewable energy data

and systems, energy controllers, and communication systems. DR programs can help the utility companies to shift the power consuming workload from peak hours to off-peak hours with the aim of electricity price reduction or load balancing between the generation side and the demand side (Shakeri et al., 2018). Existing literature shows that DR programs are not able to change the behaviour of the consumers in terms of the electricity consumption. After all, consumers are neither power analysts nor economists to monitor and schedule their energy consumption pattern (Faruqui and Sergici, 2010), (Haider et al., 2016). This is where HEMS come in in the smart grid infrastructure. HEMS are smart technologies that can respond to signals or directions from utility companies to shift or reduce the electricity load of electrical appliances and optimize demand during peak hours. Using optimization models and algorithms capable of making 24/7 decisions in HEMS on the consumer's behalf makes the development of AHEMS installed behind the meter necessary.

Gupta et al. (2016) developed a genetic algorithm that can handle non-linearity's. In their model, they reduce energy costs while simultaneously maximizing consumer happiness. This is a very important factor to take into account; Consumers aspire for the best level of comfort and happiness when operating their energy equipment (Shafie-Khah and Siano, 2017). On the other hand, the relationship between curtailed energy volume and consumer satisfaction is debatable as it should focus on the quality of energy service given by loads rather than the volume of energy (Althaher et al., 2015).

The paper by Soares et al. (2016) presents an evolutionary algorithm to optimize the integrated usage of multiple residential energy resources considering a large set of management strategies, considering cost and discomfort objective functions. The results have shown significant savings mainly through consider thermostatically controlled loads in the multi objective model. Furthermore, they observe that savings are also dependent on the end-user's preferences and degree of willingness to accept automated control.

So, in this paper, it is useful to go more into the optimization of AHEMS. However, one important factor to consider is the long computing time and complexity of multi objective models. Merdanoğlu et al. (2020) emphasizes the importance of reducing the computational complexity of the MILP models. The goal is to enable models to be solved on low-cost hardware hence, HEMS implementations will be further investigated. Furthermore, the study by Diao et al. (2012) discusses including domestic electric water heaters (EWH) in their models and testing the reactions of consumers with various control schemes that apply DR. Their research shows that EWH responses are effective and that they depend on hot water usage. They provide new insight and highlight the importance of control measures to improve the effectiveness of energy management systems.

The article by Mamun et al. (2016) examines the use of electrochemical batteries for datacenter DR. They focuses on minimizing a Pareto combination of total electricity cost and battery, by optimizing the control policy used for DR. This work is motivated by the rapid growth of both the datacenter industry and its electricity needs. In this case, the datacenter becomes the prosumer. Meaning that the consumer both produces and consumes energy, producing it by means of generating energy using solar panels at the households. It is also useful to look at what happens when households become prosumers. According to Zhou and Cao (2019), the

role of buildings in managing on-site renewable energy will also shift from traditional consumers to prosumers. A hybrid building energy management system consists of an on-site renewable system, a building energy demand system, hybrid energy storage systems, and a utility grid.

AHEM consumers proactively control, monitor and manage their energy production and consumption using the home energy management system, so they work according to schedules, and manage energy production through photovoltaic panels for self-consumption. Consumption and production data are deemed valuable, allowing customers to adjust their schedules and habits, for example washing or cooking, to reap the benefits of solar production (Gonçalves and Patrício, 2022). Advancements in renewable energy technologies have led the electricity sector to prioritize power generation from renewable energy resources (RER) as an alternate approach to satisfy future demand (Arun and Selvan, 2017).

Energy price forecasting is essential for smart grid operations because it improves energy demand management and utility load planning. The work by Zhang et al. (2020) estimates short term electricity pricing using seasonal auto regressive integrated moving average (SARIMA) models. According to Clements and Hendry (2008), economic forecasting approaches range from basic regression models to complicated stochastic models, with ARIMA models being among the most prevalent. In Smeekes and Wijler (2018), penalized regression approaches are also used for macroeconomic forecasting and include Ridge and the Lasso selection operator. According to Emmert-Streib and Dehmer (2019), Lasso has received substantial attention, notably in applications of macroeconomic forecasting, and displays promising results, giving cause to investigate these approaches in our research. Random Forest is another machine learning approach that has been examined in macroeconomics literature. The Breiman (2001) approach combines tree predictors by drawing values from a randomly chosen vector from the same distribution for each tree in the forest.

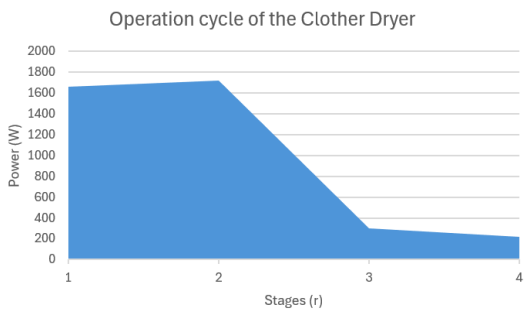
This literature review shows that further investigation in storage of batteries, different and more time periods and accurate pricing data are important and are worth studying. The work of Zhang et al. (2020) could be extended by investigating other forecasting methods proposed by Smeekes and Wijler (2018) and Breiman (2001), to obtain accurate forecasts for smart grid operations.

3 Data

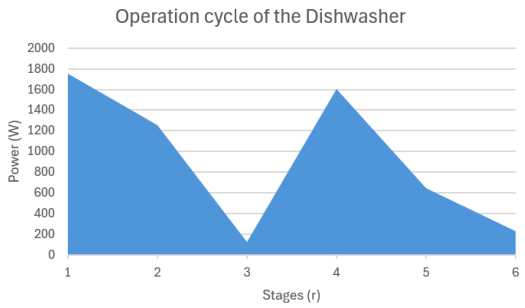
In this section, we describe the datasets used in this paper. The research is conducted with multiple datasets to optimize the complete model. First of all, the energy prices in €/kWh and the power level prices in € charged to the consumers by the retailer, shown in Tables 10 and 11 respectively, are used to determine the low and high price periods we want to buy and sell. The power requested to the grid in W and the local PV energy generation in Wh in Tables 12 and 13, this is incorporated in the energy amount available for the consumer. To determine when the shiftable loads are going to operate, the comfort time slots and operation cycles of the shiftable load j (Clother Dryer, Dishwasher and Laundry Machine) consisting of d_j stages with power g_{jr} required at each stage r can be found in Tables 14 and 15, which are also shown in Figure 1. From Figures 1a, 1b and 1c we observe that the three different loads have all a different amount of stages. Each stage consists of 15 minutes, so the total operation cycle of the Clother Dryer is

1 hour, for the Dishwasher this is 1 hour and 30 minutes, and lastly for the Laundry Machine this will be 1 hour and 45 minutes.

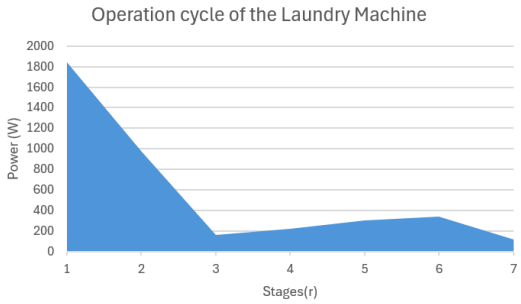
Parameters used for the EWH load, the water withdrawal for consumption in kg, ambient temperatures and parameters for the thermostatic load, can be found in Tables 16, 17, 18 and 19 respectively. Those parameter values and data are used to determine the behaviour of the EWH. Furthermore to determine the behaviour of the AC, we use the outdoor temperatures defined in Table 20. For the batteries, Static and EV, the parameters of Tables 21 and 22 are used to determine their behaviour. Lastly, the maximum allowed power for exchanges with the grid is specified in Table 23. A full description of the dataset, along with the parameter values, can be found in Appendix A, the same data is used in the research of Antunes et al. (2022) and is also available in their paper.



(a) Operation cycle of the Clother Dryer



(b) Operation cycle of the Dishwasher



(c) Operation cycle of the Laundry Machine

Figure 1: The operation cycles of the shiftable load j (Clother Dryer, Dishwasher and Laundry Machine) consisting of d_j stages with power g_{jr} required at each stage r.

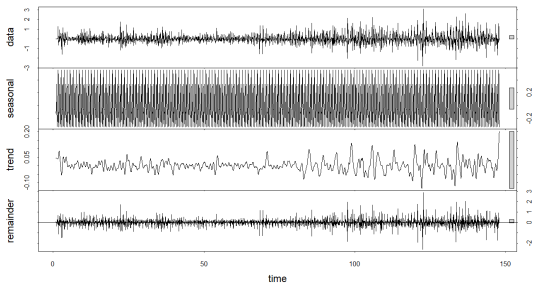
Additionally, to forecast the day-ahead market energy prices, we will use hourly historic data of the day-ahead market price obtained from ENTSO-E[3], given in MWh. In combination with this data, the renewable energy sources are included to make a more accurate forecast. The data of the production of solar energy in solar parks and wind energy that is produced on land and in the sea in the Netherlands is obtained from the "Nationale Energie Dashboard"[4], which updates the dataset hourly. Both data sets consists of hourly data in 2024, which spans the 1st of January 2024 up to the 26th May 2024. It contains data observed at an hourly frequency with the volume of the energy in kWh. To correctly prepare the data, so it can be used for forecasting, some operations are necessary. Firstly, the data of the day-ahead market

[3]Link to Day Ahead market Prices
[4]Link to renewable energy data

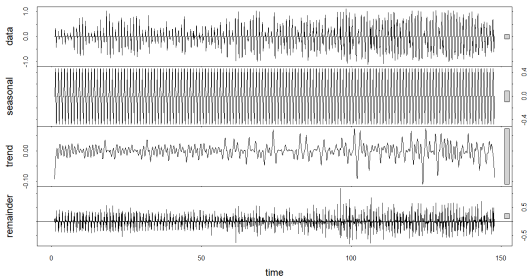
energy prices are transformed from MWh to kWh, to get comparable results with the replicated outcomes from Antunes et al. (2022). After this, we took the first differences of both data sets; this is essential making the time series stationary. Doing this reduces the trend components and the seasonality. Thereafter, we had some missing observations. We dealt with this by taking the average of the previous and upcoming observations to obtain a complete data set. Furthermore, we standardize the data to make sure we get accurate and reliable forecasts.

Figure 2 shows the decomposition of the time series for the day-ahead market energy price in Figure 2a, for the solar energy in Figure 2b, for the wind energy on land in Figure 2c and for the wind energy in the sea in Figure 2d. All those four figures are split in the raw data, seasonality pattern, a trend component and the residuals. The trend component shows significant fluctuations, which indicates that there is no consistent increase or decrease for the data in Figure 2. There is no linearly trend present in the data, the wavy patterns indicate more complex underlying patterns. On the other hand, shows the seasonality components clearly a regular pattern, which indicate strong seasonality variations. Those signs of strong seasonality, give us strong confirmation that it is necessary to delete the seasonality, thus deseasonalize the data. This makes it possible to better understand underlying trends and random fluctuations and make better predictions.

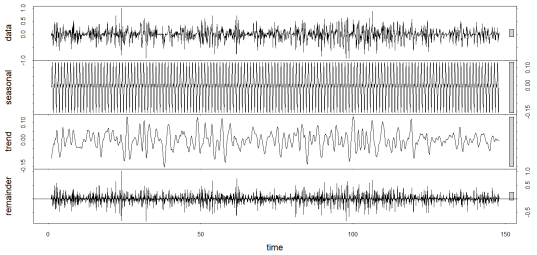
With the data specified in Appendix A there are seven price periods used with in each period different energy prices. The forecasted hourly day-ahead market energy prices will be used for a new defined 24 price periods.



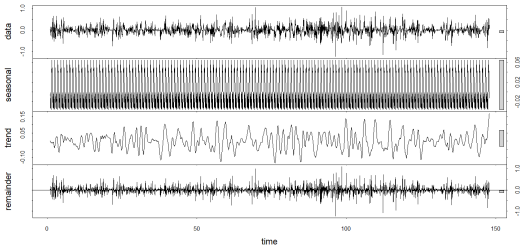
(a) Decomposition of the time series of the day-ahead market energy price.



(b) Decomposition of the time series of solar energy.



(c) Decomposition of the time series of the land wind energy



(d) Decomposition of the time series of the sea wind energy

Figure 2: Decomposition of the time series, showing the raw data, seasonality, trend and the residuals

4 Methodology

In this Section, the Methodology is explained. Section 4.1 describes the different methods developed for the shiftable loads, electric water heater (EWH), air conditioner system (AC), static battery and electric vehicle (EV) battery, and eventually the integrated model as described in Antunes et al. (2022). After we have described the integrated model, we will describe a construction of forecasts for the market energy price and how to integrate this into the integrated model, which is shown in Section 4.2.

4.1 Methods

In this study, the objective is to develop detailed optimization models that are key to retailers who are interested in offering time-of-use (ToU) tariff schemes to anticipate consumers' reactions. In this paper, this is done using shiftable loads, EWH, AC, static and EV batteries to charge with energy from the grid, sell the energy back to the grid, and include power components. In Sections 4.1.1 up to 4.1.5 the model parameters, auxiliary variables, decision variables, and constraints are well defined for those sub models. The integrated model to be optimized, taking all the sub models into account, is explained in Section 4.1.6. Throughout this paper, a planning period of one day is considered, which consists of T time intervals of length Δt, $t = 1, ..., T$. Time t is denoted as the time interval from $t-1$ to t. The length of Δt depends on the discretization adopted; typically, values of 1, 5 or 15 min are used. In this paper, intervals of 1 minute are used, so $\Delta t = 1/60$h; there is $T = 1440$ for a planning period of 24 hours. The ToU tariff is used with six different pricing periods, which are input variables denoted by C_t^{buy}. The overall model's aim is to minimize the net cost while also taking into account the power cost component related to the grid's peak power requirements.

4.1.1 Model for shiftable loads

Shiftable loads, like laundry machines, dishwashers and dryers, follow different cycles that cannot be interrupted. Unlike simplified models, our approach recognizes the different energy demands at each stage of these cycles. For example, the cycle of a laundry machine includes steps as boiling water, spinning, and drying, where each step varies in the amount of power.

In this proposed approach by Antunes et al. (2022), we take into account the duration and power requirements of each stage, as well as consumer preferences for discontinuous operational time windows. Optimization comprises identifying the best starting time for each load to ensure completion within the stated time frame while remaining consistent with the overall planning period.

This approach optimizes energy use while catering to client preferences, creating a balance between efficiency and convenience.

This model takes several technical parameters into account for each shiftable load $J \in \{1, ..., J\}$. The allowed time slots by the consumer for operation of load j are defined as $[T_{Lj}, ..., T_{Uj}]$. The parameter d_j is defined as the duration of the operation cycle of the load j and g_{jr} is defined as the power requested by load j at one stage. Both parameters hold for the time slot r of its operation cycle, with $r = 1, ..., d_j$ and $R(j) = \{1, ..., d_j\}$. The actual operation

cycle will be discretized based on Δt.

Furthermore, for the model, decision and auxiliary variables are needed for $t = 1, ..., T$. The auxiliary variable P_{jt}^{Sh} is defined as the power requested to the grid by shiftable load j in time t of the planning period in kW. P_{jt}^{Sh} will be set to zero when t falls outside the allowed time slot. The binary variable s_{jt} will be set to 1 if load j starts operating in time t. This is done for the starting times of each load. To model this, the constraints below are used for the Shiftable Loads:

$$\sum_{t=T_{L_j}}^{T_{U_j}-d_j+1} s_{jt} = 1, \quad j = 1, ..., J \tag{1}$$

$$P_{jt}^{\text{Sh}} = 0, \quad j = 1, ..., J, \quad t < T_{L_j} \vee t > T_{U_j} \tag{2}$$

$$P_{jt}^{\text{Sh}} = \sum_{r \in R(j): r \leq t \wedge r \leq t+1-T_{L_j}} g_{jr} \cdot s_{j(t-r+1)}, \quad j = 1, ..., J, \quad t' = T_{L_j}, ..., T_{U_j} \tag{3}$$

$$s_{jt} \in \{0, 1\}, \quad j = 1, ..., J, \quad t = T_{L_j}, ..., T_{U_j} - d_j + 1 \tag{4}$$

In the model for the Shiftable Loads, constraints (1) and (4) force the starting and end time of a load operation within the allowed time slot. Furthermore, constraint (2) sets the requested power to the grid to zero when time t falls outside the allowed time slot. Lastly, constraint (3) states that the power requested to the grid by load j at each time t should be within the allowed time slot. The constraint states that for every r must hold that $r \leq t \wedge r \leq t+1-T_{Lj}$, to ensure that only existing variables are considered in $s_{j(t-r+1)}$

4.1.2 Model for the Electric Water Heater

The Electric Water Heater (EWH) consumes a large share of the energy in the household, with power ranging from 1500W to 6000W, depending on household size and usage. The EWH is controlled by a thermostat and is critical for providing hot water. Simultaneously it takes the water withdrawal, supply and sanitary limits into account to avoid the growth of the Legionella bacteria.

To model the constraints of the EWH, we take several technical parameters into account, such as a heating element power, ambient and inlet water temperatures, tank characteristics, and the time required to remove germs. The allowed temperature ranges and the water withdrawal specifications are a consumer input. In the optimization is determined whether the heating element is turned on or off in order to efficiently regulate the hot water temperature within the tank. The model assumes a water temperature of at least 60 °C for 11 minutes or 75 °C for 3 minutes per day.

The EWH model requires a number of parameters. These include the power of the resistive heating element, defined as P^R in kW. The ambient temperature defined as τ_t^{amb} in time t (°C). τ^{net} is defined as the inlet water temperature in °C. The water withdrawal rate is denoted as m_t for consumption in time t in kg. The hot water tank capacity in kg is defined as M and tank

envelope area is defined as A in m^2. Furthermore, the heat transfer coefficient of the tank in W/m^2 °C is defined as U, the specific heat of water is defined as c_p in J/kg °C. The minimum and maximum allowable temperatures are defined as $\tau_{\min}$ and $\tau_{\max}$ in °C, t^{req} is the required time to eliminate the bacteria on a certain temperature and τ^{req} is defined as the required temperature for the time t^{req} to be kept to eliminate the bacteria in °C.

Furthermore, for the model, decision and auxiliary variables are needed for $t = 1, ..., T$. Firstly, the binary variable v_t is denoted as the *on/off* control variable of the heating element in time t, where v_0 is a constant. The binary variable n_t is set to 1 in the first time period t where $\tau_t > \tau^{\text{req}}$ for t^{req}. Furthermore, the variable τ_t is denoted as the hot water temperature inside the tank in time t in °C, where τ_0 is a constant. Lastly, the power losses through the envelope in time t in kW is defined as P_t^{losses}, where P_0^{losses} is a constant.

To model this, the constraints below are used for the operation of the EWH.

$$P_t^{\text{losses}} = A \cdot U(\tau_t - \tau_t^{\text{amb}}), \quad t = 1, ..., T \tag{5}$$

$$\tau_{t+1} = (\frac{M - m_t}{M} \cdot \tau_t + \frac{m_t}{M} \cdot \tau^{\text{net}}) + \frac{P^R v_t - P_t^{\text{losses}}}{M c^p} \cdot \Delta t, \quad t = 0, ..., T-1 \tag{6}$$

$$\tau_t \geq \tau^{\min} - \mathcal{M} v_t, \quad t = 1, ..., T \tag{7}$$

$$\tau_t \leq \tau^{\max} + \mathcal{M}(1 - v_t), \quad t = 1, ..., T \tag{8}$$

$$\sum_{t=1}^{T - t^{\text{req}} + 1} n_t = 1 \tag{9}$$

$$\tau_t \geq \sum_{t'=1 (t' \leq t}^{t^{\text{req}}} \tau^{\text{req}} \cdot n_{t-t'+1}, \quad t = 1, ..., T \tag{10}$$

$$v_t \in \{0, 1\}, \quad n_t \in \{0, 1\}, \quad t = 1, ..., T \tag{11}$$

$$P_t^{\text{losses}} \geq 0, \quad \tau_t \geq 0, \quad t = 1, .., T \tag{12}$$

In the model for the EWH, constraints (5) and (6) represent the temperature inside the tank. Furthermore, constraints (7) and (8) allow for a deviation from the minimum temperature when heating water. v_t becomes equal to 1 to accommodate scenarios where the initial temperature $\tau_0 < \tau_t^{\min}$ and allow the scenarios with initial temperature $\tau_t > \tau^{\max}$ when $v_t = 0$ to avoid excessive withdrawal. $\mathcal{M}$ represents a large positive number. When the EWH is turned *off* or *on*, the constraints $\tau_t > \tau^{\min}$ and $\tau_t < \tau^{\max}$ become difficult to satisfy. Lastly, to avoid the growth of the legionella bacteria constraints (9) and (10) ensure that the temperature has a minimum of τ^{req} for at least t^{req} consecutive times.

4.1.3 Model for the Air Conditioner

Residential HVAC systems significantly increase electrical energy usage. These systems can be successfully controlled to optimize energy use, particularly by utilizing time-differentiated tariffs and consumer flexibility in accommodating temporary deviations from desired indoor temperatures. This optimization approach is based on the behavior of a thermostat with hysteresis that controls an air conditioning system. We ensure to include the nominal power of the AC, the coefficient of performance, the external temperature, and the thermal properties of the building envelope into the optimization model. Furthermore, consumers can specify their preferred minimum and maximum comfort temperatures, which is incorporated into the model. The AC's *on/off* state will be determined through optimization, and the inside temperature is regulated accordingly.

The AC model requires a number of parameters. These include the outdoor temperature in time t in °C defined as θ_t^{ext}. The minimum and maximum indoor temperature allowed in °C are respectively defined as $\theta^{\min}$ and $\theta^{\max}$. The nominal power of the AC appliance in kW is defined as P^{AC}. Furthermore, the parameter β is equal to $(U \cdot A/C)\Delta t$, where U the weighted average overall heat transfer coefficient of the building unit envelope in kW/m^2°C, A the surface area of the envelope in m^2. So $U \cdot A$ is the overall thermal conductance of the unit envelope in kW/°C, and C is the overall thermal capacity in kJ/°C. Lastly, the parameter γ is equal to $\chi \cdot \Delta t/C$, where χ is the coefficient of performance of the AC appliance. Both parameters, β and γ, are calculated in the thermal model.

Furthermore, for the model, we need decision and auxiliary variables for $t = 1, ..., T$. First, the binary variable s_t^{AC} is defined as the *on/off* control variable, where s_0^{AC} is a constant. The variable for the indoor temperature in °C is defined as θ_t^{in}, where θ_0^{in} is a constant. Lastly, the binary variables y_t and z_t are used to enforce the logical conditions of the thermostat switching. Where $y_t = 1$ in case $\theta_t^{\text{in}} < \theta^{\max}$ and $z_t = 1$ if $\theta_t^{\text{in}} > \theta^{\min}$.

To model this, the constraints below are used for the operation of the AC.

$$\theta_t^{\text{in}} = (1-\beta)\theta_{t-1}^{\text{in}} + \beta\theta_{t-1}^{\text{ext}} + \gamma P^{\text{AC}} s_{t-1}^{\text{AC}}, \quad t = 1, ..., T \tag{13}$$

$$\theta_t^{\text{in}} \geq \theta^{\min} - \mathcal{M} s_t^{\text{AC}}, \quad t = 1, ..., T \tag{14}$$

$$\theta_t^{\text{in}} \leq \theta^{\min} + \mathcal{M} z_t, \quad t = 1, ..., T \tag{15}$$

$$\theta_t^{\text{in}} \geq \theta^{\max} - \mathcal{M} y_t, \quad t = 1, ..., T \tag{16}$$

$$z_t + y_t - s_{t-1}^{\text{AC}} + s_t^{\text{AC}} \leq 2, \quad t = 1, ..., T \tag{17}$$

$$z_t + y_t + s_{t-1}^{\text{AC}} - s_t^{\text{AC}} \leq 2, \quad t = 1, ..., T \tag{18}$$

$$\theta_t^{\text{in}} \leq \theta^{\max} + \mathcal{M}(1 - s_t^{\text{AC}}), \quad t = 1, ..., T \tag{19}$$

$$s_t^{\text{AC}}, z_t, y_t \in \{0, 1\}, \quad t = 1, ..., T \tag{20}$$

In the model for the AC, constraint (13) indicates the indoor temperature in each time t. The indoor temperature is determined as a function of the current indoor temperature, the external temperature, and the AC operation period in time $t-1$. Furthermore, constraint (14) guarantees that the AC is turned *on* when the indoor temperature becomes below the minimum allowed temperature. Constraints (15) - (18) ensure the system remains *on/off* while the indoor temperature is between the lower and upper limits of the thermostat dead-band. Lastly, constraint (19) forces the variable s_t^{AC} to zero when the indoor temperature is above the maximum temperature exceeds the maximum allowable temperature and switches the AC *off*.

4.1.4 Model for static and electric vehicle battery

The EV battery deals with energy exchanges and is incorporated into the overall optimization of all energy resources. The static battery is always available, but the EV battery is only available at home, resulting in additional limits for the EV battery.

Application of the vehicle-to-grid (V2G) mode is based on the expected energy service (mobility) to meet the customer needs. To model this we include the charging and discharging efficiency, a minimum and maximum amount for the battery to charges. In the model, we will optimize the charging and discharging patterns for both the static and EV batteries based on the consumer's requested battery charge in the time of departure of the EV.

We denote $\boldsymbol{x}$ as the type of battery; $\boldsymbol{x} \in \{\mathcal{B}, \mathcal{V}\}$, where $\mathcal{B}$ refers to the static battery and $\mathcal{V}$ refers to the EV battery as described in Antunes et al. (2022) . Furthermore, t_a is denoted as the first time unit after the EV arrives at home and t_d is the time of departure. The time domain of each battery is denoted as $T_{\boldsymbol{x}}$, where $T_B = T$ and $T_V = [t_a, t_d]$. The battery status will be given in the unit of energy (kWh).

The parameters and data needed to simulate battery systems are as follows: $\eta_{\boldsymbol{x}}^{ch}$ and $\eta_{\boldsymbol{x}^{dch}}$ represent the battery's charging and discharging efficiency. The parameters $E_{\boldsymbol{x}}^{\min}$ and $E_{\boldsymbol{x}}^{\max}$ represent the minimum and maximum allowed battery charge in kWh, respectively. We make use of an initial battery charge in kWh, which is defined as $E_{\boldsymbol{x}}^0$. For the battery of type $\boldsymbol{x} = B$ this is in time $t = 0$ and this is in time $t = t_a - 1$ for the battery of type $\boldsymbol{x} = V$. $E_{\boldsymbol{x}}^{\text{req}}$ is defined as the battery charge requested at the end of the planning period for $\boldsymbol{x} = B$, we consider $E_B^{\text{req}} = E_B^0$. For $\boldsymbol{x} = V$ this is defined at the time of departure t_d, for both cases given in kWh. $P_{\boldsymbol{x}}^{\text{ch\_max}}$ and $P_{\boldsymbol{x}}^{\text{dch\_max}}$ are the maximum charge and discharge power allowed for the battery in kW.

Furthermore, for the model, we need auxiliary variables for $t \in T_{\boldsymbol{x}}$. First, the power in kW is defined as $P_t^{\boldsymbol{x}2\text{H}}$, such that $P_t^{\boldsymbol{x}2\text{H}}\Delta t$ is defined as the energy transferred from the battery type $\boldsymbol{x}$ to home in time t, the battery discharge. Then the charge of the battery is defined as the energy transferred from the home to the battery of type $\boldsymbol{x}$ in time t ($P_t^{\text{H}2\boldsymbol{x}}\Delta t$), where $P_t^{\text{H}2\boldsymbol{x}}$ is the power in kW. Lastly, the energy in the battery type $\boldsymbol{x}$ in time t in kWh is defined as $E_{\boldsymbol{x},t}$.

When $t = 0$ then $E_{\boldsymbol{x},0}$ is equal to the constant $E^0_{\boldsymbol{x}}$, as defined above.

We make use of two binary decision variables, $S_t^{\mathrm{H2}\boldsymbol{x}}$ and $S_t^{\boldsymbol{x}\mathrm{2H}}$, setting the binary variables to 1 when the battery type $\boldsymbol{x}$ is charging and discharging in time t, respectively.

To model this, the constraints below are used for the operation of the static and the EV batteries.

$$E_{\boldsymbol{x},t} = E_{\boldsymbol{x},t-1} + (\eta^{\mathrm{ch}}_{\boldsymbol{x}} P_t^{\mathrm{H2}\boldsymbol{x}} \Delta t) - (\frac{P_t^{\boldsymbol{x}\mathrm{2H}} \Delta t}{\eta^{\mathrm{dch}}_{\boldsymbol{x}}}), \quad t \in T_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \{B, V\} \tag{21}$$

$$E_{\boldsymbol{x}^{\min}} \leq E_{\boldsymbol{x},t} \leq E^{\max}_{\boldsymbol{x}}, \quad t \in T_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \{B, V\} \tag{22}$$

$$0 \leq P_t^{\mathrm{H2}\boldsymbol{x}} \leq P^{\mathrm{ch\_max}}_{\boldsymbol{x}} S_t^{\mathrm{H2}\boldsymbol{x}}, \quad t \in T_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \{B, V\} \tag{23}$$

$$0 \leq P_t^{\boldsymbol{x}\mathrm{2H}} \leq P^{\mathrm{dch\_max}}_{\boldsymbol{x}} S_t^{\boldsymbol{x}\mathrm{2H}}, \quad t \in T_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \{B, V\} \tag{24}$$

$$S_t^{\mathrm{H2}\boldsymbol{x}} + S_t^{\boldsymbol{x}\mathrm{2H}} \leq 1, \quad t \in T_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \{B, V\} \tag{25}$$

$$E_{B,T} \geq E^{\mathrm{req}}_B; E_{V,t_d} \geq E^{\mathrm{req}}_V \tag{26}$$

$$S_t^{\mathrm{H2}\boldsymbol{x}}, S_t^{\boldsymbol{x}\mathrm{2H}} \in \{0, 1\}, \quad t \in T_{\boldsymbol{x}}, \quad \boldsymbol{x} \in \{B, V\} \tag{27}$$

In this model, constraint (21) represents the energy for both types of batteries considering the charging and discharging processes.Constraint (22) imposes a minimum and maximum level of energy in the battery for both types of batteries in each time unit. The maximum charging and discharging rates are limited, respectively, by the constraints (23) and (24). Constraint (25) states that the batteries can be in one mode only in each time unit; charging or discharging. Constraint (26) enforces that the energy in the batteries can not be lower than the required energy at the end of the planning periods for $\boldsymbol{x} = B$ and at the time of departure for $\boldsymbol{x} = V$. Lastly, constraint (27) states that $S_t^{\mathrm{H2}\boldsymbol{x}}$ and $S_t^{\boldsymbol{x}\mathrm{2H}}$ are binary variables.

4.1.5 Power component

In addition to the energy costs in €/kWh, it might be useful to consider the power costs by establishing a power tariff structure in €/kW per day. This consists of L price levels, and the consumer is charged based on the peak power used. To implement this in the overall model, some additional parameters and decision variables are defined. First, the auxiliary variable P_t^{G2H} is defined as the power in kW such that $P_t^{\mathrm{G2H}} \Delta t$ is the energy in kWh which is transferred from the grid to the home (G2H) in time t at unit cost C_t^{buy} in €/kWh. Furthermore, the following additional parameters and decision variables are defined for $l = 1, ...,$L. The power level prices in €/kWh are defined as C_l^{Cont}, where the consumer will be charged for the peak power used. The power levels in kW are then defined as P_l^{Cont}, and the binary variable u_l^{Cont} is equal to 1 when the maximum power l is used during the planning period. Constraint (28) states that the power

from the grid to the home can not exceed the power of level l, where the level has maximum power. Furthermore, constraint (29) states that exactly one price level uses the maximum power. The constraints to determine the power levels in the overall model are defined below:

$$P_t^{\text{G2H}} \leq \sum_{l=1}^{\text{L}} P_l^{\text{Cont}} \cdot u_l^{\text{Cont}}, \quad t = 1, ..., T \tag{28}$$

$$\sum_{l=1}^{\text{L}} u_l^{\text{Cont}} = 1 \tag{29}$$

$$u_l^{\text{Cont}} \in \{0, 1\}, \quad l = 1, ..., \text{L} \tag{30}$$

4.1.6 Complete model

The MILP model presented in the Sections above can be combined in an overall model to make an integrated optimization of all energy resources. In the model the possibility is considered of selling back to the grid at a stipulated price. There are some additional parameters and decision variables required for $t = 1, ..., T$.

Firstly, B is defined as the base load, which is non-controllable, in kW. The power in kW is defined as P_t^{PV}, such that $P_t^{\text{PV}}\Delta t$ is the local PV energy generation in time t in kWh. The energy remuneration in €/kWh, so selling to the grid, in time t is denoted by C_t^{sell}, and the maximum power allowed for exchanges with the grid in kW is defined as $P^{\text{G\_max}}$.

In the overall model one additional auxiliary variable and two binary variables are needed. First, the auxiliary variable P_t^{H2G} is defined as the power in kW, such that $P_t^{\text{H2G}}\Delta t$ is the energy in kWh transferred from home to the grid (H2G) in time t at unit cost C_t^{sell} in €/kWh. The calculations of the C_t^{sell} values where not given in the paper of Antunes et al. (2022). Looking at the Literature of Dutch government rules for selling back to the grid, it this seems an appropriate to take 60 % of the value of C_t^{buy}. The binary variables s_tG2H and s_t^{H2G} are equal to 1 when the energy is flowing from the grid to home and from home to the grid in time t, respectively.

The overall model is defined below:

$$\min_{P^{\text{G2H}}, P^{\text{H2G}}} \sum_{t=1}^{T} [(C_t^{\text{buy}} P_t^{\text{G2H}} \Delta t) - (C_t^{\text{sell}} P_t^{\text{H2G}} \Delta t)] + \sum_{l=1}^{L} (C_l^{\text{Cont}} u_l^{\text{Cont}}) \tag{31}$$

$$\text{s.t.} \quad 0 \leq P_t^{\text{G2H}} \leq P^{\text{G\_max}} s_t^{\text{G2H}}, \quad t = 1, ..., T \tag{32}$$

$$0 \leq P_t^{\text{H2G}} \leq P^{\text{G\_max}} s_t^{\text{H2G}}, \quad t = 1, ..., T \tag{33}$$

$$s_t^{\text{G2H}} + s_t^{\text{H2G}} \leq 1, \quad t = 1, ..., T \tag{34}$$

$$P_t^{\text{G2H}} - P_t^{\text{H2G}} + P_t^{\text{PV}} = B_t + \sum_{j=1}^{J} P_{j,t}^{\text{Sh}} + P^{\text{AC}} s_t^{\text{AC}} + P^R v_t + (P_t^{\text{H2B}} - P_t^{\text{B2H}}), \quad \forall t \in T \setminus T_V \tag{35}$$

$$P_t^{\text{G2H}} - P_t^{\text{H2G}} + P_t^{\text{PV}} = B_t + \sum_{j=1}^{J} P_{j,t}^{\text{Sh}} + P^{\text{AC}} s_t^{\text{AC}} + P^R v_t + (P_t^{\text{H2B}} - P_t^{\text{B2H}}) + (P_t^{\text{H2V}} - P_t^{\text{V2H}}), \quad \forall t \in T_V \tag{36}$$

$$\begin{aligned} &\text{Constraints (1)-(4) for the shiftable loads.} \\ &\text{Constraints (5)-(12) for the EWH.} \\ &\text{Constraints (13)-(20) for the AC.} \\ &\text{Constraints (21)-(27) for the static and EV batteries.} \\ &\text{Constraints (28)-(30) for the power component.} \end{aligned} \tag{37}$$

$$s_t^{\text{G2H}}, s_t^{\text{H2G}} \in \{0, 1\}, \quad t = 1, ..., T \tag{38}$$

In the model, constraints (32)-(34) limit exchanges between the home and the grid to a maximum and impose that the flow occurs only in one direction, from the grid to the home or vice versa. Furthermore, constraints (35) and (36) model the power balance and differ in the operation slot of the EV battery.

4.2 Forecasting Methods

In this section, five different forecasting methods are described. The objective is to forecast the energy price of the day ahead market that captures renewable energy productions, using solar and wind parks.

The observation of the dependent variable at time t, y_t, can be described by its set of explanatory variables $\boldsymbol{x}_t$ according to the following formulation:

$$y_t = \alpha + \boldsymbol{x}_t' \boldsymbol{\beta}_x + \varepsilon_t \tag{39}$$

where α is the intercept, $\boldsymbol{x}_t$ contains of the set explanatory variables described in Section 3 and lagged versions of the dependent variable y_t. With $p = 1, \ldots, 6$ is the number of lags, and ε_t is a idiosyncratic error term

First, we explain the auto regressive method (AR) and the Seasonal Auto Regressive Integrated Moving Average (SARIMA) method, followed by two penalized regression methods; Lasso and Ridge, and the machine learning method; the Random Forest (RF).

4.2.1 The Auto Regressive (AR)

The Auto Regressive (AR) model represents a form of random process in which the output variable relies linearly on its own prior values. The AR model of order p, abbreviated as AR(p), may be expressed as:

$$y_t = \alpha + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t, \tag{40}$$

α is the intercept term, ϕ_i are model parameters, p is the number of delayed observations, and ε_t is the error term at time t.

4.2.2 SARIMA

The Seasonal Auto Regressive Integrated Moving Average (SARIMA) model is an extension of the ARIMA model that takes into account univariate time series data with seasonal components, which can be described by the following formulation:

$$\hat{y}_{t+h} = c + \phi_1 y_{t+h-1} + \phi_2 y_{t+h-2} + ... + \phi_p y_{t+h-p} + \theta_1 \epsilon_{t+h-1} + \theta_2 \epsilon_{t+h-2} + ... + \theta_q \epsilon_{t+h-q} + \Phi_1 y_{t+h-s} + \Phi_2 y_{t+h-2s} + ... + \Phi_P y_{t+h-Ps} + \Theta_1 \epsilon_{t+h-s} + \Theta_2 \epsilon_{t+h-2s} + ... + \Theta_Q \epsilon_{t+h-Qs} + \beta_1 X_{t+h}^{Solar} + \beta_2 X_{t+h}^{Wind_{Sea}} + \beta_3 X_{t+h}^{Wind_{Land}} + \epsilon_{t+h} \tag{41}$$

where p is the order of the autoregressive part, q the order of the moving average part, P the order of the seasonal autoregressive part, Q the order of the seasonal moving average part and s the length of the seasonal pattern. Furthermore, the parameters ϕ_i with $i = 1, .., p$ and Φ_j with $j = 1, ..., P$ are the autoregressive parameters for the market energy prices and their seasonal counterparts. Additionally, the parameters for the market energy prices and their counterparts for the moving average part are θ_m with $m = 1, ..., q$ and Θ_n with $n = 1, ..., Q$. In the equation ϵ_t is the white noise in time t and β_1 and β_2 are the parameters indicating the impact of the solar, sea wind and land wind energy generation on the energy market prices.

4.2.3 Penalized regression

Penalized regressions reduce the dimensionality of the input variables by shrinking the parameters of the explanatory variables. A well-know method described by Tibshirani (1996) is the Lasso method, where the parameters can be shrunk to zero exactly when the variables are insignificant. The Lasso method can be described by the following formulation:

$$\min_{\boldsymbol{\beta}_x} \left\{ \sum_{t=1}^{T} (y_t - (\alpha + \boldsymbol{x}_t' \boldsymbol{\beta}_x))^2 + \lambda \left[\gamma \sum_{j=1}^{N} \frac{|\beta_{x,j}|}{\omega_j} + (1-\gamma) \sum_{j=1}^{N} \frac{(\beta_{x,j})^2}{\omega_j} \right] \right\}, \tag{42}$$

where ω_j is the weight given to parameter $\beta_{x,j}$. Different settings for λ, γ and ω_j, give the models:

1. Lasso: $\lambda > 0$, $\alpha = 1$, $\omega_j = 1$)
2. Ridge regression: $\lambda > 0$, $\alpha = 0$, $\omega_j = 1 \quad \forall j$)

Another well-known method is Ridge, which is described by Hoerl and Kennard (1970) and given in equation (42). The main difference between Lasso and Ridge, is that Lasso has a absolute penalty term and Ridge has a squared penalty term. Meaning Ridge can not shrink the parameter values exactly to zero.

4.2.4 Random Forest

A powerful machine learning technique often used for economic forecasting is the Random Forest. In this method multiple decision trees are trained on different subsets of the data, where random selection of features is used to create splits. The Random Forest method can be described by the following formulation:

$$y_t = \mathcal{F}(\boldsymbol{x}_t) + \varepsilon_t, \tag{43}$$

where the set explanatory variables $\boldsymbol{x}_t$ with corresponding dependent variable y_t are used to construct the forest $\mathcal{F}$. According to Breiman (2001), a Random Forest is a classifier that consists of a group of tree-structured classifiers h($\boldsymbol{x}_t$, Θ_k), k = 1,...,Θ_k, where Θ_k are independent identically distributed random vectors. Each tree supports the dominant class given an input of $\boldsymbol{x}_t$.

Training the Random Forest builds on bagging. Meaning we first use the bootstrap method and then aggregate to tree learners. With the bagging technique we iterative choose a random sample with replacement from the training set $\boldsymbol{x_t}$ with corresponding outcomes y_t, then fitting the trees to these samples. We excecute the bagging procedure B times, where each time a random subset of n training examples is selected from $\boldsymbol{x_t}$, y_t with replacement, designated as $x_{t,b}$ and $y_{t,b}$.

4.2.5 Constructing the forecasts

In this section we describe how the forecasts are constructed. We create an h-step ahead forecast variable $\hat{y}^h_{t+h}$ for $h = 1, ..., 24$. Two different prepared data sets are used for the forecasting. First, the standardized data will be used for all five forecasting methods. We will also forecast on the deseasonalized data for the AR method, Lasso, Ridge and the Random Forest. For the different methods different forecasting formulations are needed. Tuning is necessary to obtain better forecasting results. We have determined the number of lags used in the forecasting methods $(1, \ldots, 6)$ using the Bayesian information criterion (BIC), to criticize the model performance. Furthermore, for Lasso and Ridge we made a priori selection of the tuning parameter λ. We employ a (100×1) grid of λ-values to obtain the tuning parameter, the optimal value of λ is selected based on the BIC. For the auto regressive model, we create forecasts using the following formulation:

$$\hat{y}^h_{t+h} = \hat{\alpha} + \sum_{j=1}^{p} \hat{\phi}_j y_{t-j+h} + \varepsilon_{t+h}, \tag{44}$$

where $p \in (1, \ldots, 6)$ is the number of lags. For the SARIMA model, the forecast is obtained using the following formulation:

$$\begin{aligned} \hat{y}_{t+h} = \hat{c} + \sum_{i=1}^{p} \hat{\phi}_i y_{t+h-i} + \sum_{j=1}^{q} \hat{\theta}_j \epsilon_{t+h-j} + \sum_{k=1}^{P} \hat{\Phi}_k y_{t+h-ks} + \sum_{l=1}^{Q} \hat{\Theta}_l \epsilon_{t+h-ls} + \\ \hat{\beta}_1 X^{Solar}_{t+h} + \hat{\beta}_2 X^{Wind_{Sea}}_{t+h} + \hat{\beta}_3 X^{Wind_{Land}}_{t+h} + \hat{\varepsilon}_{t+h}, \end{aligned} \tag{45}$$

For the penalized regression methods, Lasso and Ridge, the forecast is obtained using the following formulation:

$$\hat{y}_{t+h}^{h} = \hat{\alpha} + \boldsymbol{z}_t' \hat{\boldsymbol{\beta}}_z + \sum_{j=1}^{p} \hat{\beta}_j y_{t-j+1} + \varepsilon_{t+h}. \tag{46}$$

Finally, forecast of the Random Forest method is obtained by the following formulation:

$$\hat{y}_{t+h}^{h} = \frac{1}{B} \sum_{b=1}^{B} f_b(\boldsymbol{x}_t), \tag{47}$$

where B is the number of trees in the random forest, f_b is the prediction for tree b, and $\boldsymbol{x}_{i,t}$ is the set for which the forecast is required.

An out-of-sample forecast for $h = 1$ is computed using a rolling window technique based on one week, this is 168 observations (hourly). The root-mean-square error (RMSE) and mean absolute error (MAE) are calculated and compared to the benchmark AR model to evaluate the forecasts. Additionally, Diebold-Mariano tests are conducted at a 5% significance level, with the null hypothesis of a similar prediction performance compared to the AR benchmark. If the null hypothesis is rejected, it indicates that the two models have substantial differences in predicting accuracy. This assessment technique is based on the study of Smeekes and Wijler (2018).

Furthermore, for the out-of-sample forecasts $h = 1, ..., 24$ an iterative approach has been used. where the forecasted value is fed back into the model to predict the next step. To evaluate the performance of the 24-step-ahead forecasts and their accuracy, a residual analysis is performed, where the patterns of the residuals are observed.

4.3 Integrating the forecasts in the Integrated model

The integrated model in Section 4.1.6 includes 7 time periods and accompanying grid-buying expenses. To evaluate power from and to the grid, as well as the development of battery charging and temperature changes for accurate hourly costs, anticipated hourly day-ahead energy prices from Section 4.2.5 are incorporated into the model of 4.1.6. The time frame will remain the same, consisting of one day (1440 minutes), but instead of seven time periods, there will be 24 time periods, each lasting exactly 60 minutes. In this method, the matching projected day-ahead energy market prices will be included into the model as C_{buy}.

The objective says the same, but the value of T is different. The objective is given below in equation (48), with $T = 24$. All variables and parameters say exactly the same, only C_{buy} and C_{sell} will thus change and become a vector of length 24, with the prices to buy from the grid and sell to the grid, respectively.

$$\min_{P^{\text{G2H}}, P^{\text{H2G}}} \sum_{t=1}^{T} [(C_t^{\text{buy}} P_t^{\text{G2H}} \Delta t) - (C_t^{\text{sell}} P_t^{\text{H2G}} \Delta t)] + \sum_{l=1}^{L} (C_l^{\text{Cont}} u_l^{\text{Cont}}) \tag{48}$$

5 Results

This section dives into the outcomes of our case study, offering interpretations and valuable insights into the results. The models are implemented with the CPLEX solver of IBM (IBM

ILOG CPLEX Optimization Studio) and in R studio. Both programming languages ran on an Intel i7 computer with 8.0 GB of RAM. Different running times have been used of 5 minutes and 1 hour for the optimization in the CPLEX solver. To compute the case study results a planning horizon discretization of $\Delta t = 1/60$h was chosen.

The remainder of this section is structured as follows. First we describe the optimization results of the 1 hour running time for the integrated model and after this we discuss different comfort parameterizations where a 5 minute runtime has been used in Section 5.1. In Section 5.1.1, we elaborate on the use of the static and EV batteries in the original model and how the optimization differs when in- or excluding the battery(s), in terms of the cost and balanced outcomes. Thereafter, in Section 5.2 we describe the forecast results of all models and compare these to the AR benchmark in terms of the (Relative) RMSE and the MAE. Finally, we will discuss the results of the 1 to 24 hourly forecasts integrated into the integrated optimization model in Section 5.2.1.

5.1 Replication

The integrated model has 15.519 binary variables, 1440 integer variables, 17.777 continuous variables and 49.806 constraints. Furthermore, it has a relative MIP gap of 0.010934285 for the 5 minute run time, and with 1 hour run time the gap is equal to 0.009982891. The corresponding cost objective function values were €4.079078836 and €4.077074078, respectively.

The evolution of the power from the grid to the home and from the home to the grid is shown in Figure 3. Figure 4 shows the evolution of the Static and EV battery charges and the evolution of the indoor temperature compared to the EWH temperature in the planning period. Those results are all obtained with a run time of 1 hour. The figures show that power from the grid is obtained in lower price periods, and selling to the grid occurs in higher price periods using PV generation and the energy stored in the batteries. This also means, as can be seen in Figure 4a, that the static battery will be charged in the lower price periods and supply loads operating in higher price periods. The EV battery is charged at low price periods to reach the desired state of charge at the time of departure. Figure 4b shows that the indoor temperature fluctuates within the comfort range, where the AC is supplied by the grid in low price periods and by the battery in times of high prices.

To simulate a consumer willing to trade-off comfort for cost, we consider relaxations of the minimum allowed indoor, hot water temperatures, and minimum battery charge requested at the time of departure of the EV. Table 1 shows the minimized cost and the cost reductions in comparison to the integrated model when there are no relaxations, for a 5 minute runtime. Cost savings can be obtained by lowering the minimum interior comfort temperature by 1°C and the hot water temperature by 2°C. The most visible gain, however, is the reduction in the needs for charging the electric car battery at the point of departure. The combined benefits of these three recommended settings result in a considerable cost reduction of around 12.8% over the original settings. Furthermore, the MIP gap is reasonably consistent at 0.01, indicating that changes in comfort parameters have no substantial influence on solution quality or optimization convergence. The most visible cost advantage is realized by minimizing the requirements for

the electric car battery during departure time, resulting in a 7.7% reduction. Reducing the charging time of the vehicle by reducing the departure time means significant savings. Overall this mean that when the customer is willing to accept some discomfort the cost can be reduced. The results show an effective replication of Antunes et al. (2022).

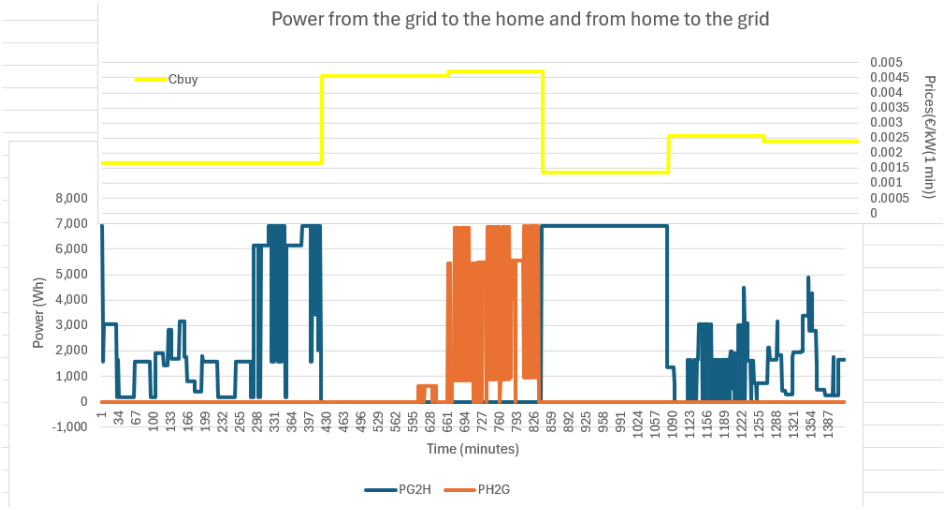
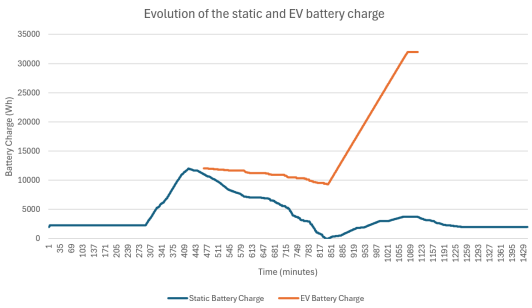
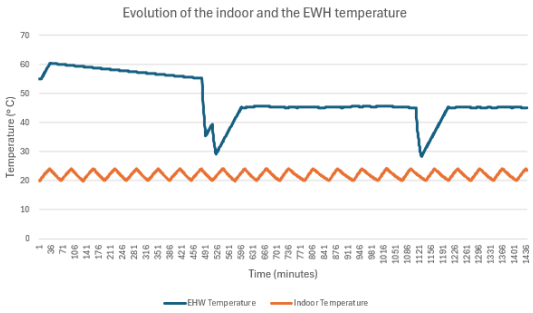


Figure 3: Power from the grid to the home and from the home to the grid.



(a) Evolution of the static and EV Battery charge.



(b) Evolution of the indoor temperature and EWH temperature.

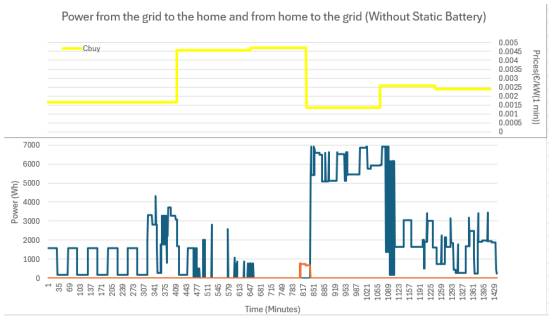
Figure 4: Comparison of the evolution of battery charge and temperature.

Table 1: Comparison of Different comfort parameterizations with 5 minute run time

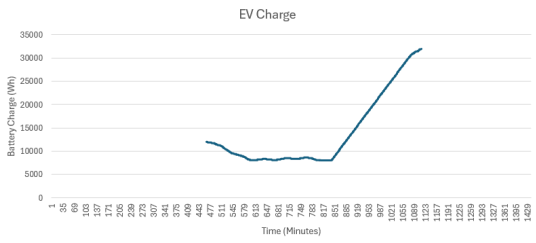
	MIP gap	**Cost**	**Cost Reduction**
Regular Integrated model	0.0109	4.0791	-
$\theta^{\min} = 19°C$	0.0107	3.9466	3.2%
$\tau^{\min} = 43°C$	0.0103	4.0016	1.9%
$E_v^{\text{req}} = 30$ **kWh**	0.0114	3.7649	7.7%
$\theta^{\min} = 19°C$ **and** $\tau^{\min} = 43°C$	0.0107	3.8712	5.1%
$\theta^{\min} = 19°C$, $\tau^{\min} = 43°C$ **and** $E_v^{\text{req}} = 30$ **kWh**	0.0116	3.5580	12.8%

5.1.1 Static and EV battery

To evaluate the performance of the integrated model without the Static, the EV, or both batteries., we performed three different models. One where the EV battery is not included in the model, where constraints with the EV battery ($\boldsymbol{x} = V$) are relaxed. Another model where the static battery is not included, for this model the constraints with the static battery ($\boldsymbol{x} = B$) are relaxed. And the last model where we do not include both batteries in the model, all the constraints of Section 4.1.4 are relaxed. The evolution of the battery charge of the EV battery and the power exchange with the grid when only using the EV battery is shown in Figure 5. The Evolution of the battery charge of the EV in Figure 5b has more or less the same pattern as the evolution of the EV battery when the static battery is involved as well (Figure 4a). The EV needs to be fully charged before the time of departure, this explains the same charging pattern. Figure 6 shows the evolution of the battery charge of the static and the power exchange with the grid when only using the static battery. According to the Figure, the static battery will be fully charged in low-price times and uncharged in high-price periods. The static battery may be charged more than the regular model when the EV, which also requires charging, is not in use and the maximum power exchange occurs over the grid. The power exchange with the grid without the use of batteries is shown in Figure 7, in this exchange there are no opportunities for storage and buying from the grid only occurs when the PV does not offer enough power and we are at the maximum of the consumer discomfort. Selling to the grid only occurs when the PV over generates relatively and the additional power is sold back to the grid. Lastly, Table 2 show the objective cost of all models and the cost reduction compared to the regular integrated model at a run time of 5 minutes. From Table 2 we can conclude that the regular model is not the most cost-efficient model. Using only the static battery appears to be more profitable, showing a 24.7 % increase in cost when the static battery is excluded, indication its significant contribution to cost savings. Conversely, the integrated model without the EV battery demonstrates the highest cost savings, with a 53.2 % improvement, suggesting the significance of the static battery. However, the regular integrated model provides a more balanced approach when considering power exchanges with the grid, as evidenced by the comparison of Figures 5, 6 and 7. This balance might be critical for maintaining grid stability and consumer comfort. The lower MIP gap in the integrated model without the static battery and the integrated model without the static and EV battery, as evidenced by Table 2 indicate more reliable solutions.

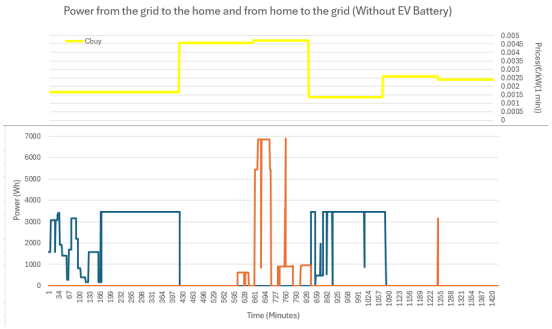


(a) Power of the grid without the static battery

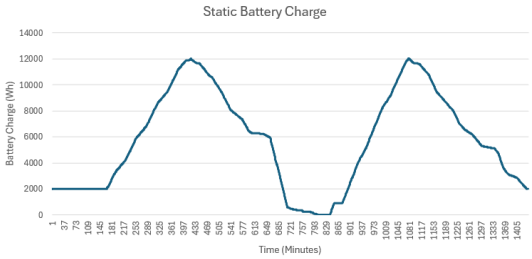


(b) Evolution of the EV Battery.

Figure 5: The evolution of the grid exchange and the EV battery charging



(a) Power of the grid without the EV battery



(b) Evolution of the Static Battery.

Figure 6: The evolution of the grid exchange and the Static battery charging

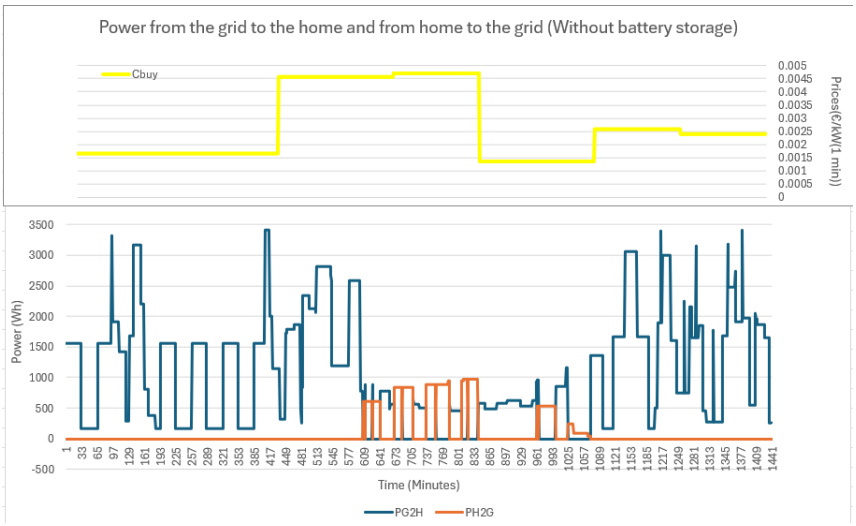


Figure 7: Power from the grid to the home and from the home to the grid without the use of batteries.

Table 2: Comparison of different models with and without static and EV battery.

	MIP gap	Cost	Cost Reduction
Regular Integrated model	0.0109	4.0791	-
Integrated model without static battery	0.0086	5.0860	-24.7%
Integrated model without EV battery	0.0218	1.9078	53.2%
Integrated model without static and EV battery	0.0064	3.8817	4.8%

5.2 Forecasting results

As mentioned in Section 3, we conducted an empirical analysis using renewable energy data and day-ahead market energy prices. We employ a rolling window with a size of 168 observation (one week), in which the in-sample period is fixed and the estimate sample is advanced by one period. The data used consists of 3525 observations, so the rolling window will be applied 3357 times. During the tuning process, the λ-values for Lasso and Ridge are selected in each rolling window, based on the lowest BIC. We chose a 100 λ-values grid evenly distributed on a

logarithmic scale between 10^{-10} and 10^{5}. To evaluate the predictions. Table 3 shows the root-mean-square errors (RMSE) and the Mean Absolute Errors (MAE) for all forecast methods used with standardized data and Table 4 shows the RMSE and the MAE for the four forecasts methods with deseasonalized data as explained in Section 4. The low RMSE and MAE values in indicate that all models have outstanding prediction performance. The low errors indicate that the models make extremely accurate predictions. Based on the RMSE and Diebold-Mariano tests shown in for the standardized data in Table 5 and for the deseasonalized data in Table 6, we can conclude that the Ridge method outperforms all other forecasting models when compared to the AR benchmark. The difference in errors is minimal, but Ridge has the best performance compared to the AR benchmark. In Section 3 we observed that it is important to delete the seasonality pattern, the same is shown in Table 7. Table 7 shows the diebold-mariano tests of the AR, Lasso, Ridge and Random Forest models with deseasonalized data compared to the same models without the deseasonalized data as a benchmark. As can be seen from Table 7, all models significantly outperform the models without the deseasonalized data, except the AR model. The AR models does not outperform the AR model without deseasonlized data, but simultaneously is not significant. Now we have seen the importance of deseasonalizing the data, we can conclude that the ridge method with the deseasonalized data gives the best forecasting performances. Furthermore, Ridge has a very low computational time and is therefore ideal to include in the real time applications. A high computational time would negatively effect the consumer satisfaction.

As a result, we utilize this forecasting approach to generate 1 to 24 step forward forecasts to acquire one day's worth of energy pricing for our integrated model. In Section 5.2.1, these results will be further discussed.

Table 3: Comparison of different forecasting models based on RMSE and MAE for standardized data.

	AR	**SARIMA**	**Lasso**	**Ridge**	**Random Forest**
RMSE	1.035647	0.8243219	0.3264832	0.3248561	0.3807453
MAE	0.007949645	0.02861189	0.0001659838	0.0001660624	0.001308491

Table 4: Comparison of different forecasting models based on RMSE and MAE for the standardized and deseasonalized data.

	AR	**Lasso**	**Ridge**	**Random Forest**
RMSE	0.7549985	0.2866007	0.2867557	0.3381442
MAE	0.009069322	1.492988e-05	1.488271e-05	0.00101846

Table 5: DM Test Statistic and P-Value for different forecasting models relative to the AR benchmark model for the standardized data.

	SARIMA	**Lasso**	**Ridge**	**Random Forest**
DM Test Statistic	-2.8186	-6.7474	-6.7476	-6.6941
P-Value	0.004851	1.766e-11	1.764e-11	2.533e-11

Table 6: DM Test Statistic and P-Value for different forecasting models relative to the AR benchmark model for the standardized and deseasonalized data.

	Lasso	**Ridge**	**Random Forest**
DM Test Statistic	-1.9144	-1.9143	-1.905
P-Value	0.05565	0.05566	0.05687

Table 7: DM Test Statistic and P-Value for different forecasting models with deseasonalized data relative to the same models without deseasonlized data as a benchmark model.

	AR	**Lasso**	**Ridge**	**Random Forest**
DM Test Statistic	1.304	-4.3688	-4.3075	-5.9599
P-Value	0.1923	1.287e-05	1.698e-05	2.786e-09

5.2.1 Integrating the forecasts in the integrated model.

We performed residual analyses on the results from the 1 to 24 step-ahead predictions. Figure 8 demonstrates that the residuals are generally randomly distributed around the null line. Furthermore, the residuals' dispersion appears steady, yet some of them appear to fall outside of the cluster. We see no apparent patterns or trends in the order of the residuals, which gives us reason to assume the forecasting method is accurate.

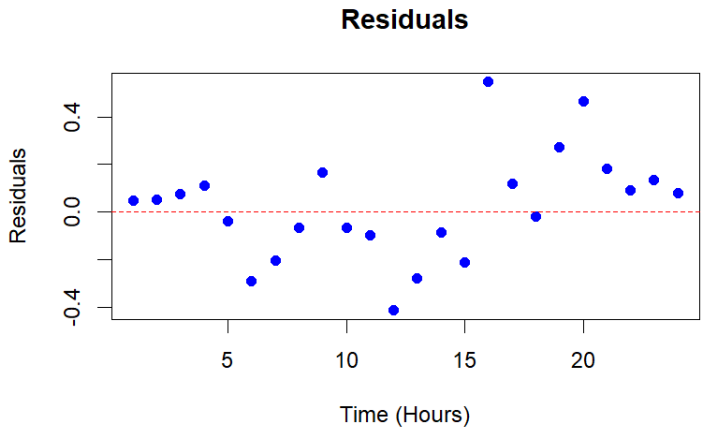


Figure 8: Residuals of the Ridge forecast 1 up to 24 steps ahead.

The integrated model with the forecasted day ahead market energy prices has a relative MIP gap of 0.000135487 for the 5 minute run time, with corresponding cost objective function value of -78.68847926. This means that there is no cost, instead the consumer makes profit

of 78.68847926. The forecast is for a day in April, where there is a lot of produced solar energy and a long period where the sun is of energy. Furthermore, this day the produced energy of wind on land and in the sea is very high, when we look at the data. This explains the expected profit made when we make 24-ahead prediction.

Figure 9 shows that the hourly forecasted prices allow for more detailed and dynamic optimization, compared to the 7-period pricing in Figure 3, which are the same outputs as the paper of Antunes et al. (2022). The hourly divided price periods reveal a more frequent shift between buying and selling power, indicating a highly responsive system. This responsiveness can maximize cost savings by capitalizing on short term price drops and avoiding high price periods. The 7-period pricing produces longer periods of consistent behavior, but may lose out on smaller price fluctuations. On the other hand leads the hourly pricing system certainly to a better cost management, but it needs more complex forecasting and real-time modifications. Overall, the decision between various pricing systems is based on the complexity of the energy management system versus the possible cost savings from more frequent price adjustments.

Figure 10a shows the charging and discharging patterns of the static and EV batteries, those patterns suggest different usage strategies. The static battery seems to be used more dynamically, possible responding to short-term price fluctuations, while the EV battery follows a more predictable charging schedule. The static battery charge shows a more fluctuating pattern compared to the EV battery, also more fluctuating pattern comparing the Figure 4a, the same as the original figure of Antunes et al. (2022). The EV battery charge shows a consistent increase until it reaches full charge and then drops sharply likely due to usage or a scheduled discharge. Furthermore, Figure 10b shows the changes of the indoor temperature and the EWH temperature. The indoor temperature has exactly the same pattern compared to the original figure in Figure 4b, the same as in Antunes et al. (2022), while the EWH reaches a higher temperature in the beginning of the time period. In the hourly pricing model the EWH temperature reaches a temperature higher than 80°C, while for the 7-period pricing the temperature touches the 60°C.

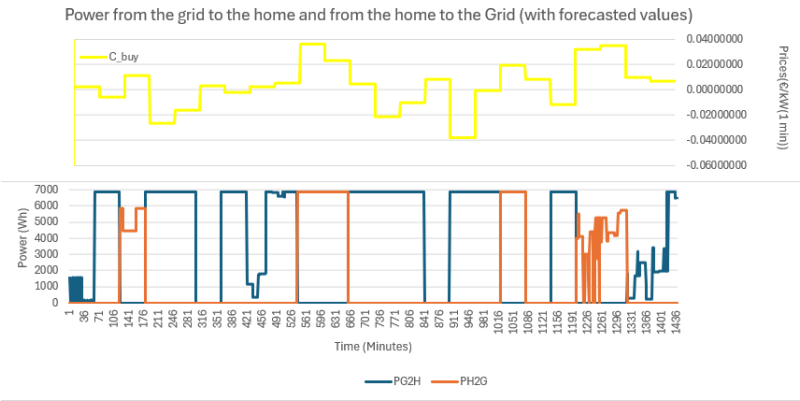
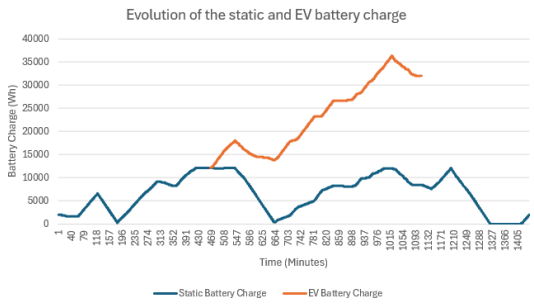
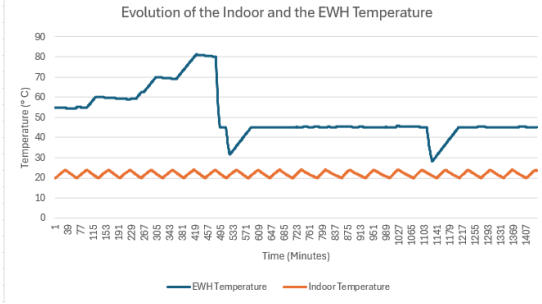


Figure 9: Power from the grid to the home and from the home to the grid.



(a) Evolution of the static and EV Battery charge.



(b) Evolution of the indoor temperature and EWH temperature.

Figure 10: Comparison of the evolution of battery charge and temperature.

Table 8 shows the minimized cost and the cost reductions in comparison to the integrated hourly model when there are no relaxations, for a 5 minute runtime. Cost reductions can be obtain by given the consumer some discomfort by lowering the minimum indoor temperature, the hot water temperature or reduction in the needs for the charging of the EV before departure. Again, as seen in Table 1 the combined discomfort parameters result in the highest cost reduction, in this case a reduction of 1.9% compared to the regular integrated model. The MIP gaps are consistent at a value of 0.0001, what again indicates that changes in comfort parameters have no substantial influence on solution quality or optimization convergence.

Furthermore, Table 9 shows the objective cost of all models and the cost reduction compared to the regular integrated hourly model at a run time of 5 minutes. The Table clearly shows that without the static battery the costs are extremely higher. From Table 2 we have already noticed this, but at an hourly frequency this is much more noticeable. Table 9 shows a 51.0 % increase in cost when the static battery is excluded and a 88.6 % increase in cost when the static and EV battery are excluded, indication its significant contribution to cost savings. Again the regular integrated model is more balanced when considering the power exchanges with the grid and the low MIP gap indicates that the solutions are reliable.

The comparisons of the Tables 1 and 8 and Tables 2 and 9 show that the objective values of the cost vary a lot between the 7-period pricing and the hourly pricing. However, the patterns in the importance of the static battery and how the cost can be reduced if the consumer is willing to accept some discomfort stay the same.

Table 8: Comparison of Different comfort parameterizations with 5 minute run time with forecasted costs

	MIP gap	**Cost**	**Cost Reduction**
Regular Integrated model	0.000135487	-78.68847926	-
$\theta^{\min} = 19°C$	0.000116739	-79.0529292	0.5%
$\tau^{\min} = 43°C$	0.000219794	-79.0129046	0.4%
$E_v^{\text{req}} = 30$ **kWh**	0.000095595	-79.49738687	1.0%
$\theta^{\min} = 19°C$ **and** $\tau^{\min} = 43°C$	0.000137254	-79.37263082	0.9%
$\theta^{\min} = 19°C$, $\tau^{\min} = 43°C$ **and** $E_v^{\text{req}} = 30$ **kWh**	0.000136083	-80.17770353	1.9%

The use of more time periods, along with accurate forecasts, demonstrates that costs may

Table 9: Comparison of different models with and without static and EV battery with forecasted costs.

	MIP gap	Cost	Cost Reduction
Regular Integrated model	0.000135487	-78.68847926	-
Integrated model without static	0.008256615	-38.58224178	-51.0%
Integrated model without EV	0.001206819	-78.75279017	0.1%
Integrated model without static and EV	0.035138156	-8.984618469	-88.6%

be reduced and there can even be profit made on days when energy prices are low (or negative). This also demonstrates that ToU tariff systems benefit customers when they have the flexibility or desire to tolerate some inconvenience related with appliance operation shifting, indoor and hot water temperatures, or battery state of charge requirements.

6 Conclusion

This paper examined a comprehensive and modular set of MILP models to balance optimization model detail with practical solution capabilities while requiring minimal computer resources. The models are created as adaptable building blocks, including economic and comfort objectives using a variety of modeling methodologies. They can be tailored to individual requirements, such as appliance specifications, comfort preferences, time-differentiated pricing systems, and power cost structures. Section 3 provides entire data sets and the paper presents a case study with these actual data. Additionally, this paper evaluated the forecasting performance of the AR, SARIMA, Lasso, Ridge, and Random Forest methods for day-ahead market energy prices using renewable energy sources. Forecast accuracy was assessed using the root-mean-square error, compared to the AR benchmark model, and the Diebold-Mariano test. The most accurate method was then used to forecast hourly energy prices for one day, integrating this into the MILP model.

The results show that when customers have the ability to adjust their consumption habits, they may reduce their net power cost or even profit by using an energy management system equipped with the models suggested in this paper to make optimal decisions on their behalf. The utilization of a static battery is extremely profitable; the more capacity available, the more the user will store power in low-cost periods and then use this energy in high-cost periods rather than purchasing from the grid. On the other hand, the EV battery will not just serve as a storage mechanism; it must also be completely charged before departure. This does not enhance the grid's purchasing strategy for consumers. Using the Ridge method as an hourly forecasting tool and integrating it into an energy management system would allow you to significantly reduce your electricity cost.

Overall, the results showed that independent of how much price periods used, the importance of the static battery and the cost reduction when the consumer is willing to accept some discomfort stay the same. On top of this, the accurate forecasted hourly price periods will better divide the energy consumption and returning energy back to the grid for the consumer. Hereby, the prediction can be used to implement real time price periods. In conclusion, consumers who

are willing to accept some annoyance due to changes in appliance performance, indoor and hot water temperatures, or battery state of charge needs are rewarded by ToU tariff systems.

7 Discussion

In this section, we discuss some limitations to our research and provide ideas for further research.

An issue we encountered, is the unknown definition of C_t^{sell}. In the paper of Antunes et al. (2022) there is no specification how C_t^{sell} is related to C_t^{buy}. We did some further research what is commonly used as a percentage of the price to sell back to the grid, and came to the conclusion this varies a lot in the literature. For further research it is useful to consider multiple relations of C_t^{sell} with C_t^{buy} or use real-time data on both costs. Another issue we encountered, is that the RMSE and the MAE seems a bit low on the scale of kWh we use. There is no further explanation what could be the cause. We have made sure that the data is standardized, stationary and deseasonalized. In addition, we performed a residual analyses on the results from the 1 to 24 step-ahead estimates to check if the forecasting method is accurate. There was no reason to suspect that the prediction was incorrect, but the low errors gives the suspicion that something is still incorrect. Therefore it is necessary to investigate this further.

In this paper, the battery lifespans has not been taken into account. Therefore, for further research it is useful to take the performance and lifespan of the batteries into account. Over time the performance will probability get worse, and this will effect the costs of the household. Furthermore, it is also of interest to investigate the effects of the volatility of the energy market prices. If all households eventually make use of the batteries and optimize their consumption, will there be leveling. A positive effect would be a higher demand of batteries, so the price of batteries will be lower. Nevertheless, even with a less expensive battery, it may take longer to earn back the investment when energy costs drop because there is a surplus supply due to the batteries.

In our research, the focus is on a single household, where we found that is very interesting to store energy at low price periods and use it in high periods for regular activities in the house or even sell it to the grid. For one single household the contribution to the CO_2 emission is probably not that much. If every household could manage their consumption patterns, it might be possible to have as little CO_2 emission as possible, and that there will ultimately are less of no coal-fired power stations needed anymore. In this way energy would be generated in a more sustainable way and will it be used to the maximum extent. Further research should thus expand the scope to include multiple households with diverse energy usage patterns to provide a more comprehensive understanding of potential benefits and challenges. The models presented in this research have limited capabilities in adaptability of price fluctuations and supply and demand. Therefore, it could be very beneficial to investigate more in dynamic models that consider the variable battery capabilities, fluctuating energy demands and real-time adaptive control systems, to lower the energy costs of the households. Besides the benefits of the cost reductions, the dynamic models can increase the reliability and sustainability of the entire network. Another direction for further research, is investigating the impact of integrating additional renewable energy sources, such as biomass, geothermal energy and hydroelectric energy. This research focused on the variability of the energy produced in solar and wind parks, while it could be

beneficial to add those other renewable energies to increase the resilience and sustainability of the energy management system. Furthermore, it is interesting to observe the behaviour of the consumers, how will the act on certain changes within the DR program. In this research some discomfort and changes in parameters is analyzed in Section 5.1, but it is usefull to look in to this in more details and other changes, for example the changes in the environment or financial changes. Lastly, since the application of machine learning (ML) and artificial intelligence (AI) gives more insights and is in full development, it could be interesting if those techniques could be more integrated within the Autonomous Home Energy Management Systems (AHEMS), to get considerably improvements en more efficient results.

References

S. Althaher, P. Mancarella, and J. Mutale. Automated demand response from home energy management system under dynamic pricing and power and comfort constraints. *IEEE Transactions on Smart Grid*, 6(4):1874–1883, 2015.

C. H. Antunes, M. J. Alves, and I. Soares. A comprehensive and modular set of appliance operation milp models for demand response optimization. *Applied Energy*, 320:119142, 2022.

S. Arun and M. Selvan. Intelligent residential energy management system for dynamic demand response in smart buildings. *IEEE Systems Journal*, 12(2):1329–1340, 2017.

M. Beaudin and H. Zareipour. Home energy management systems: A review of modelling and complexity. *Renewable and sustainable energy reviews*, 45:318–335, 2015.

K. Bhattacharya, M. H. Bollen, and J. E. Daalder. *Operation of restructured power systems*. Springer Science & Business Media, 2012.

L. Breiman. Random forests. *Machine learning*, 45:5–32, 2001.

M. P. Clements and D. F. Hendry. *A companion to economic forecasting*. John Wiley & Sons, 2008.

R. Diao, S. Lu, M. Elizondo, E. Mayhorn, Y. Zhang, and N. Samaan. Electric water heater modeling and control strategies for demand response. In *2012 IEEE power and energy society general meeting*, pages 1–8. IEEE, 2012.

F. Emmert-Streib and M. Dehmer. High-dimensional lasso-based computational regression models: regularization, shrinkage, and selection. *Machine Learning and Knowledge Extraction*, 1 (1):359–383, 2019.

A. Faruqui and S. Sergici. Household response to dynamic pricing of electricity: a survey of 15 experiments. *Journal of regulatory Economics*, 38(2):193–225, 2010.

L. Gonçalves and L. Patrício. From smart technologies to value cocreation and customer engagement with smart energy services. *Energy Policy*, 170:113249, 2022.

A. Gupta, B. P. Singh, and R. Kumar. Optimal provision for enhanced consumer satisfaction and energy savings by an intelligent household energy management system. In *2016 IEEE 6th International Conference on Power Systems (ICPS)*, pages 1–6. IEEE, 2016.

H. T. Haider, O. H. See, and W. Elmenreich. A review of residential demand response of smart grid. *Renewable and Sustainable Energy Reviews*, 59:166–178, 2016.

A. E. Hoerl and R. W. Kennard. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, 12(1):55–67, 1970.

A. Ipakchi and F. Albuyeh. Grid of the future. *IEEE power and energy magazine*, 7(2):52–62, 2009.

A. Mamun, I. Narayanan, D. Wang, A. Sivasubramaniam, and H. Fathy. Multi-objective optimization of demand response in a datacenter with lithium-ion battery storage. *Journal of Energy Storage*, 7:258–269, 2016.

H. Merdanoğlu, E. Yakıcı, O. T. Doğan, S. Duran, and M. Karatas. Finding optimal schedules in a home energy management system. *Electric Power Systems Research*, 182:106229, 2020.

M. Shafie-Khah and P. Siano. A stochastic home energy management system considering satisfaction cost and response fatigue. *IEEE Transactions on Industrial Informatics*, 14(2): 629–638, 2017.

M. Shakeri, M. Shayestegan, S. S. Reza, I. Yahya, B. Bais, M. Akhtaruzzaman, K. Sopian, and N. Amin. Implementation of a novel home energy management system (hems) architecture with solar photovoltaic system as supplementary source. *Renewable energy*, 125:108–120, 2018.

P. Siano. Demand response and smart grids—a survey. *Renewable and sustainable energy reviews*, 30:461–478, 2014.

S. Smeekes and E. Wijler. Macroeconomic forecasting using penalized regression methods. *International journal of forecasting*, 34(3):408–430, 2018.

A. Soares, A. Gomes, C. H. Antunes, and C. Oliveira. A customized evolutionary algorithm for multiobjective management of residential energy resources. *IEEE Transactions on Industrial Informatics*, 13(2):492–501, 2016.

R. Tibshirani. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 58(1):267–288, 1996.

J. S. Vardakas, N. Zorba, and C. V. Verikoukis. A survey on demand response programs in smart grids: Pricing methods and optimization algorithms. *IEEE Communications Surveys & Tutorials*, 17(1):152–178, 2014.

J. Winkler, M. Pudlik, M. Ragwitz, and B. Pfluger. The market value of renewable electricity–which factors really matter? *Applied energy*, 184:464–481, 2016.

J. Zhang, Z. Tan, and Y. Wei. An adaptive hybrid model for short term electricity price forecasting. *Applied Energy*, 258:114087, 2020.

Y. Zhou and S. Cao. Energy flexibility investigation of advanced grid-responsive energy control strategies with the static battery and electric vehicles: A case study of a high-rise office building in hong kong. *Energy Conversion and Management*, 199:111888, 2019.

A Data used for the case study

Table 10: Electricity prices (in €/kWh) charged to the consumers by the retailer.

	P1	**P2**	**P3**	**P4**	**P5**	**P6**
Tariff Range	[1,420]	[421,660]	[661,840]	[841,1080]	[1081,1260]	[1261,1440]
Price (€/kWh)	0.0996	0.2739	0.2828	0.0817	0.1548	0.1438

Table 11: Power level prices (in €) charged to the consumers by the retailer.

Power Level	**Prices (€/day)**	**Max Power (kW)**
1	0.2047	2.30
2	0.2206	3.45
3	0.2834	4.60
4	0.3492	5.75
5	0.4198	6.90
6	0.6280	10.35
7	0.8302	13.80
8	1.0324	17.25
9	1.2351	20.70

Table 12: Power requested to the grid (in W) in each unit of time $t \in T$ (expressed in intervals of time [initial time, final time]) by (non-controllable) base load.

Time intervals	**Power (W)**
[1,480]	165
[481,510]	700
[511,540]	170
[541,660]	85
[661,810]	160
[811,960]	130
[961,1200]	160
[1201,1215]	500
[1216,1245]	1600
[1246,1275]	750
[1276,1290]	250
[1291,1305]	450
[1306,1350]	280
[1351,1365]	1080
[1366,1440]	250

Table 13: Local PV energy generation (in Wh) in each unit of time $t \in T$ (expressed in intervals of time [initial time, final time]).

Time intervals	**Energy Generation (Wh)**
[1,465]	0
[466,495]	200
[496,525]	250
[526,585]	400
[586,645]	700
[646,705]	1000
[706,765]	1050
[766,825]	1100
[826,885]	1050
[886,945]	1000
[946,1005]	700
[1006,1020]	400
[1021,1050]	250
[1051,1080]	200
[1081,1440]	0

Table 14: Comfort time slots ($T_j = [T_{1_j}, T_{2_j}], j \in \{1, ..., J\}$) allowed for the operation of each shiftable load.

Shiftable Loads		
DW	LM	CD
[1,480]	[406,870]	[1126,1440]

Table 15: Operation cycles of the shiftable loads. Appliance Power required by the appliance at each stage of its operation cycle (W).

Appliance	**Power required by appliance at each stage of its operation cycle (W)**						
	1-15	**16-30**	**31-45**	**46-60**	**61-75**	**76-90**	**76-105**
DW	1750	1250	120	1600	640	220	
LM	1840	980	160	220	300	340	120
CD	1660	1720	300	220			

Table 16: Parameters of the EWH load.

Parameters										
P_R	M	AU	c_p	τ^{net}	τ^{min}	τ^{max}	τ^{req}	t^{req}	τ_0	v_0	P_0^{losses}
1500	100	2.06	1.1419408	18	$45°C$	$85°C$	$60°C$	11 min	$55°C$	0	0

Table 17: Water withdrawal for consumption (in kg) in each unit of time $t \in T$ (expressed in intervals of time [initial time, final time]).

Time intervals	[0,479]	[480,490]	[491,509]	[510,520]	[521,1109]	[1110,1125]	[1126,1440]
Water Withdrawal	0	7.2	0	7.2	0	7.2	0

Table 18: Ambient temperatures around the EWH for a period of 24h ($T = 1440$) (expressed in intervals of time [initial time, final time]).

Time intervals	[1,181]	[182,541]	[542,721]	[722,901]	[902,1081]	[1082,1261]	[1262,1440]
Temperature (ºC)	19.6	18.5	22.4	23.2	23.7	22.6	21

Table 19: Parameters of the thermostatic load.

Parameters							
θ^{max}	θ^{min}	θ_0^{in}	P_{AC}^{nom}	s_0	α	β	γ
24 °C	20 °C	20 °C	1400W	0	0.99046	0.00954	0.000185

Table 20: Outdoor temperatures for a period of 24h – T=1440 (expressed in intervals of time [initial time, final time]).

t	$\tau_t^{ext}(^\circ C)$	t	$\tau_t^{ext}(^\circ C)$	t	$\tau_t^{ext}(^\circ C)$
0	9.45				
[1,15]	9.45	[481,495]	8.96	[961,975]	12.92
[16,30]	9.40	[496,510]	8.92	[976,990]	12.79
[31,45]	9.35	[511,525]	8.92	[991,1005]	12.64
[46,60]	9.30	[526,540]	9.00	[1006,1020]	12.50
[61,75]	9.25	[541,555]	9.19	[1021,1035]	12.40
[76,90]	9.20	[556,570]	9.43	[1036,1050]	12.35
[91,105]	9.15	[571,585]	9.66	[1051,1065]	12.32
[106,120]	9.10	[586,600]	9.80	[1066,1080]	12.30
[121,135]	9.05	[601,615]	9.81	[1081,1095]	12.29
[136,150]	9.01	[616,630]	9.75	[1096,1110]	12.27
[151,165]	8.96	[631,645]	9.72	[1111,1125]	12.25
[166,180]	8.90	[646,660]	9.80	[1126,1140]	12.20
[181,195]	8.83	[661,675]	10.06	[1141,1155]	12.13
[196,210]	8.78	[676,690]	10.48	[1156,1170]	12.02
[211,225]	8.76	[691,705]	10.97	[1171,1185]	11.88
[226,240]	8.80	[706,720]	11.50	[1186,1200]	11.70
[241,255]	8.91	[721,735]	12.00	[1201,1215]	11.48
[256,270]	9.06	[736,750]	12.43	[1216,1230]	11.25
[271,285]	9.20	[751,765]	12.78	[1231,1245]	11.05
[286,300]	9.30	[766,780]	13.00	[1246,1260]	10.90
[301,315]	9.32	[781,795]	13.09	[1261,1275]	10.83
[316,330]	9.28	[796,810]	13.07	[1276,1290]	10.85
[331,345]	9.19	[811,825]	12.99	[1291,1305]	10.94
[346,360]	9.10	[826,840]	12.90	[1306,1320]	11.10
[361,375]	9.02	[841,855]	12.82	[1321,1335]	11.31
[376,390]	8.95	[856,870]	12.78	[1336,1350]	11.52
[391,405]	8.91	[871,885]	12.77	[1351,1365]	11.67
[406,420]	8.90	[886,900]	12.80	[1366,1380]	11.70
[421,435]	8.92	[901,915]	12.87	[1381,1395]	11.56
[436,450]	8.96	[916,930]	12.95	[1396,1410]	11.18
[451,465]	9.00	[931,945]	13.00	[1411,1425]	10.52
[466,480]	9.00	[946,960]	13.00	[1426,1440]	9.50

Table 21: Parameters of the static battery.

η_{ch}^B	η_{dch}^B	E_{min}^B	E_{max}^B	P_{max}^{Bch}	P_{max}^{Bdch}	E_0^B
0.95	0.95	0 kWh	12 kWh	6 W	6 W	2 kWh

Table 22: Parameters of the EV.

η^V_{ch}	η^V_{dch}	E^V_{min}	E^V_{max}	E^V_{req}	P^{Bch}_{max}	P^{Bdch}_{max}	t_a	t_d	$E^V_{t_a}$
0.95	0.95	8 kWh	40 kWh	32 kWh	6 kW	6 kW	466	1110	12 kWh

Table 23: Maximum power allowed for exchanges with the grid.

P^G_{max}
6.9 kW

B Code explanation

To obtain all the results two different program languages are used; the CPLEX solver of IBM (IBM ILOG CPLEX Optimization Studio) and R studio. For the the optimization problem of Section 4.1.6 the CPLEX solver is used, to obtain the same results as Section 5.1 the run configuration: Run, should be ran. To obtain the different comfort parametrerizations the data file: Thesis_data.dat should be changed with some values. The values of $\tau^{\min}$, $\theta^{\min}$ and E_v^{req} could be changed to obtain the same results.

To obtain the results of Section 5.1.1 the run configurations: Run_extension_without_EV, Run_extension_without_Static and Run_extension_without_Statis_(andEV) should be ran.

To obtain the results of Section 5.2 the thesis_main.R file should be ran. This consists of 4 parts; the data preparation, a part where the forecasting methods are constructed, a part where the Diebold-Mariano test is performed, and a part where the 1 up to 24 steps ahead are determined. In the main file 4 other scripts are called in the top; "thesis_autoregressive_models.R", "thesis_forecast_rolling_window.R", "thesis_principle_component_models.R" and "thesis_steps_ahead_forecast_RF.R".

The 1 up to 24 forecasts are then integrated into the integrated model and the results of Section 5.2.1 are obtained by running the run configuration: Run_extension_hourly_data in CPLEX. Furthermore to compare the outcomes with and without the batteries the following run configurations should be ran: "Run_extension_hourly_without_EV", "Run_extension_hourly_without_static" and "Run_extension_hourly_without_static_and_EV". To obtain the different comfort parametrerizations the data file: Thesis_extension_hourly.dat should be changed with some values. The values of $\tau^{\min}$, $\theta^{\min}$ and E_v^{req} could be changed to obtain the results.