POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Optimisation sous incertitude de la demande d'un réseau électrique à l'aide d'une batterie virtuelle de chauffe-eau intelligents

LOUIS-PHILIPPE PROULX

Département de mathématiques et de génie industriel

Mémoire présenté en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

Mathématiques appliquées

Décembre 2022

© Louis-Philippe Proulx, 2022.

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Ce mémoire intitulé :

Optimisation sous incertitude de la demande d'un réseau électrique à l'aide d'une batterie virtuelle de chauffe-eau intelligents

présenté par **Louis-Philippe PROULX**

en vue de l'obtention du diplôme de *Maîtrise ès sciences appliquées*

a été dûment accepté par le jury d'examen constitué de :

Issmaïl EL HALLAOUI, président

Michel GENDREAU, membre et directeur de recherche

Miguel F. ANJOS, membre et codirecteur de recherche

Roland MALHAMÉ, membre

REMERCIEMENTS

Je tiens sincèrement à remercier l'équipe de CanmetÉNERGIE à Varennes qui m'a encouragé tout au long de ma maîtrise. Mes directeur et codirecteur de recherche, Michel Gendreau et Miguel Anjos, avec qui j'ai eu le plaisir de discuter tout au long de ma maîtrise et qui ont pris à cœur ma réussite. À ma famille et mes amis qui m'ont encouragé et supporté.

RÉSUMÉ

Les chauffe-eau intelligents peuvent jouer un rôle important pour ajouter de la flexibilité au réseau électrique et intégrer davantage d'énergie renouvelable variable. Nous proposons des modèles d'optimisation qui permettent le contrôle d'une population de chauffe-eau sous forme de batterie virtuelle. Afin de modéliser cette batterie virtuelle, on utilise des données de chauffe-eau réel. Cette batterie virtuelle peut être utilisée afin de fournir des services au réseau électrique tel que la minimisation de la pointe électrique. Cette batterie virtuelle a été testée à l'aide de données historiques de la demande d'un réseau électrique. De plus, sachant que l'opérateur d'un réseau électrique devra utiliser une prévision de la demande, plusieurs scénarios d'utilisation ont été créés afin de valider la performance de celle-ci. Deux techniques ont permis d'améliorer son opération : une formulation de la fonction objectif qui tient compte d'éléments connus dans la prévision de la demande et une utilisation en horizon fuyant avec mise à jour régulière de la prévision. La formulation stochastique n'a pas permis d'améliorer la performance de la batterie virtuelle dans le contexte des données réelles utilisées.

ABSTRACT

Smart electric water heaters can play an important role in the electric grid when controlling their consumption. They increase the demand flexibility and can help to integrate more renewable energy into the grid. We propose optimization models for controlling a large population of smart electric water heaters modelled as a virtual battery. Real electric water heater historical data is used to parametrize the virtual battery. We then control this virtual battery for different grid services, e.g., to minimize the peak demand. For the different use cases, we used real historical demand from a local distribution company. To simulate the context in which the optimization models will be used, we also construct forecast models for the grid demand. We were able to validate the performance of optimization models through different realistic uses cases. We identify two modelling approaches that help improve the control. Firstly, the formulation of the objective function needs to be carefully devised by focusing on information that is more reliable or certain. Secondly, reapplying the optimization model in a rolling horizon framework with the updated forecast also proved to be a good strategy. Finally, our stochastic formulation of the optimization model could not improve further the performance for the control of the virtual battery.

TABLE DES MATIÈRES

REMERCIEMENTS III

RÉSUMÉ IV

ABSTRACT V

TABLE DES MATIÈRES VI

LISTE DES TABLEAUX VIII

LISTE DES FIGURES IX

LISTE DES SIGLES ET ABRÉVIATIONS X

CHAPITRE 1 INTRODUCTION 1

CHAPITRE 2 REVUE DE LA LITTÉRATURE 3

2.1 Le rôle des charges flexibles dans le réseau électrique 3

2.2 Modélisation d'une population de chauffe-eau 4

2.3 Prévision de la demande électrique et outils d'apprentissage supervisé 5

2.4 Gestion de l'incertitude 6

CHAPITRE 3 ORGANISATION DU MÉMOIRE 7

CHAPITRE 4 ARTICLE 1: OPTIMIZATION MODEL FOR THE OPTIMAL OPERATION OF A DATA-DRIVEN VIRTUAL BATTERY OF SMART ELECTRIC WATER HEATERS .. 8

4.1 Introduction 8

4.2 Literature Review 9

4.3 Methodology 11

4.3.1 Estimating the EWH energy losses 11

4.3.2 Deterministic Optimization Model 12

4.3.3 Probabilistic Load Forecast 16

4.3.4 Rolling Horizon and Stochastic Optimization Model 17

4.3.5 Strengths and Limitations 18
4.4 Calibration of the Optimization Model 19
4.5 Computational Results 20
4.5.1 Results of the Deterministic Model with Perfect Forecast 21
4.5.2 Results of the Deterministic Model with a Realistic Forecast 23
4.5.3 Impact of the Rolling Horizon and the Updated Forecast 24
4.5.4 Adding the Stochastic Formulation to the Rolling Horizon 25
4.5.5 With ToU Rates 26
4.6 Conclusion 26
CHAPITRE 5 DISCUSSION GÉNÉRALE 30
CHAPITRE 6 CONCLUSION ET RECOMMANDATIONS 32
RÉFÉRENCES 34

LISTE DES TABLEAUX

Table 4.1 Dataset Information 20

Table 4.2 Parameters Used in this Paper 20

Table 4.3 Key Performance Indicators Description 21

Table 4.4 KPIs with Perfect Information 22

Table 4.5 KPIs for Deterministic Models 24

Table 4.6 KPIs for Deterministic Models with Rolling Horizon 25

Table 4.7 KPIs for Stochastic Models with Rolling Horizon 26

LISTE DES FIGURES

Figure 4.1 Example of Generating EWH Losses from Power Consumption.....12

Figure 4.2 24-hour Forecast 17

Figure 4.3 Validation of calibration with objective 3 19

Figure 4.4 VB SoC under Net Demand Variation 22

Figure 4.5 VB SoC under Peak Minimization 22

Figure 4.6 Net Demand Variation 23

Figure 4.7 Peak minimization 23

Figure 4.8 VB Charging under Net Demand Variation 23

Figure 4.9 VB Charging under Peak Minimization 23

Figure 4.10 Net Demand Variation Using Forecasted Grid Demand 24

Figure 4.11 Peak Minimization Using Forecasted Grid Demand 24

Figure 4.12 Net Demand Variation with Rolling Horizon.....25

Figure 4.13 Peak Minimization with Rolling Horizon.....25

Figure 4.14 Charging of VB under ToU rates.....26

Figure 4.15 Impact on the Grid Demand for ToU Rates.....26

Figure 5.1 Répartition de l’état de charge de la population de chauffe-eau.....30

LISTE DES SIGLES ET ABRÉVIATIONS

AS Apprentissage supervisé

BV Batterie virtuelle

CEE Chauffe-eau électrique

EdC État de charge

GD Gestion de la demande

CHAPITRE 1 INTRODUCTION

Les chauffe-eau électriques sont utilisés depuis plusieurs années par les compagnies d'électricité pour diminuer la demande lors de période de pointe. Habituellement, le contrôle de ceux-ci est assez simple : lors d'événement de pointe, un signal est envoyé à un interrupteur communicant qui déconnecte alors le chauffe-eau pendant quelques heures. Plus récemment, des contrôleurs intelligents permettent une communication bidirectionnelle entre l'opérateur du réseau électrique et le chauffe-eau dans le cadre de programme en gestion de la demande. Comment optimiser la consommation d'une population de chauffe-eau en tenant compte des besoins du réseau électrique et en maintenant suffisamment d'eau chaude dans les chauffe-eau? Ce mémoire propose une solution pour le contrôle d'une population de chauffe-eau électriques et permet de valider la performance de celui-ci sous différents scénarios.

Dans les laboratoires de CanmetÉNERGIE à Varennes, plusieurs chauffe-eau de nouvelle génération sont testés. Ces chauffe-eau sont équipés de contrôleurs intelligents. De plus, ces contrôleurs envoient de l'information sur la température de l'eau dans les réservoirs ainsi que leurs consommations électriques toutes les cinq minutes. Certains réservoirs peuvent contenir de l'eau à une température supérieure à 70°C et peuvent rester sans chauffage sur une longue période sans affecter la disponibilité de l'eau chaude pour les utilisateurs. Plusieurs centaines de ces chauffe-eau ont été installés dans le cadre d'un projet pilote chez des participants. Dans les prochaines années, il est prévu de contrôler ces chauffe-eau afin d'établir le potentiel de ceux-ci à offrir de la flexibilité aux réseaux électriques. Les modèles proposés dans ce mémoire pourront être adaptés afin de mener des essais en condition réelle. Ces chauffe-eau pourraient également participer à d'autres initiatives en lien avec les réseaux de distribution électriques intelligents. Par exemple, les tarifs d'électricité pourraient être dynamiques et un agrégateur pourrait contrôler une population de chauffe-eau en fonction de ceux-ci.

Chaque population de chauffe-eau a des caractéristiques uniques en termes d'utilisation de l'eau chaude et de profil de consommation électrique. Cette consommation varie également en fonction d'autres facteurs, par exemple du temps de l'année. Un effort important a été apporté à la modélisation d'une population de chauffe-eau afin de tenir compte des données historiques

produites. Cela permet d’adapter le modèle pour le contrôle de chauffe-eau en fonction des différentes populations de chauffe-eau.

CHAPITRE 2 REVUE DE LA LITTÉRATURE

Cette revue de littérature sera divisée en quatre sections. La première section discutera du rôle des charges flexibles dans le réseau électrique, et plus particulièrement des chauffe-eau électriques. Dans la deuxième section, la modélisation des chauffe-eau sera abordée. Les avantages et inconvénients de certains modèles de chauffe-eau seront discutés. Dans la troisième section, on abordera la prévision de la demande électrique. C'est un élément important pour maximiser la flexibilité des charges. Finalement, dans la quatrième section, on fera un survol de la gestion de l'incertitude lors de la modélisation mathématique.

2.1 Le rôle des charges flexibles dans le réseau électrique

Le gouvernement du Canada s'est donné comme objectif que le réseau électrique soit carboneutre d'ici 2035 (Canada 2022). Un réseau électrique propre contribuera également de façon importante à l'atteinte d'une carboneutralité pour l'ensemble des secteurs de l'économie canadienne à l'horizon de 2050. En effet, l'électrification permettra de diminuer l'empreinte carbone. Par exemple, les automobiles passeront de l'utilisation de l'énergie fossile à l'énergie électrique. Donc, le réseau électrique devra non seulement éliminer les sources de génération polluantes, mais également faire face à une importante augmentation de la demande. Les réseaux électriques ajouteront donc une grande quantité d'énergie renouvelable variable telle que l'énergie solaire et éolienne. Dans ce contexte, il y aura une pression additionnelle afin de gérer intelligemment la demande du réseau électrique autant au niveau de la demande industrielle et commerciale que résidentielle. En matière de demande résidentielle, il existe un certain nombre d'avenues intéressantes pour contrôler la demande, notamment, le chauffage électrique des maisons, les voitures électriques et les chauffe-eau électriques. Ces charges résidentielles offrent un potentiel de flexibilité intéressante (Wong 2014). Pour les compagnies d'électricité, chacune de ces charges représente un défi afin d'exploiter leur potentiel. En effet, le contrôle de ces charges doit non seulement se faire avec l'accord des utilisateurs, mais est également soumis à un ensemble de contraintes qui sont propres à chaque type de charges. L'intégration de ces charges à l'opération du réseau électrique pose donc plusieurs défis. De plus, d'un point de vue de l'opération du réseau, c'est la réponse globale des charges qui est importante. Il est donc nécessaire de modéliser la dynamique d'une population de charges et non, la charge elle-même. L'agrégation d'un grand

nombre de charges permet de minimiser la variabilité des charges individuelles. Cette agrégation des charges peut être représentée conceptuellement sous la forme d'une batterie virtuelle. L'agrégateur contrôle alors la batterie virtuelle et offre ainsi de la flexibilité au réseau électrique. Un autre mode utilisé pour exploiter la flexibilité des charges est par la tarification dynamique. Les tarifs reflètent ainsi la valeur pour produire et acheminer l'électricité en fonction de l'état du réseau. Ces deux approches peuvent se compléter, par exemple, en émettant un signal de prix à l'agrégateur qui contrôlera la batterie virtuelle en fonction du coût de l'électricité. Dans le cas d'un réseau de distribution d'électricité intelligent, plusieurs de ces concepts sont présents (Ding et Liu 2022)

2.2 Modélisation d'une population de chauffe-eau

Dans la littérature, il existe plusieurs modèles permettant de simuler l'opération d'un chauffe-eau. Certaines simulations permettent d'estimer la température à l'intérieur du chauffe-eau ainsi que la puissance consommée (Moreau 2011; Térache, Zmeureanu, et Laurencelle 2014; Xu et al. 2014). Ces études permettent de comprendre la distribution de la température à l'intérieur d'un chauffe-eau et de valider qu'ils contiennent suffisamment d'eau chaude lors d'une journée de gestion de la demande (GD). Cela est également utile pour comprendre la variation de la température dans le fond du chauffe-eau. En effet, la température dans le fond du chauffe-eau a tendance à être plus froide durant les journées en GD. Cette température plus froide pourrait favoriser la croissance de la bactérie légionelle (Lacroix 1999). Des recommandations ont été faites par l'Institut national de santé publique du Québec afin de prévenir la croissance de cette bactérie. Les chauffe-eau haute température auraient une température suffisamment haute pour empêcher la prolifération de la bactérie. De plus, un chauffe-eau haute température permet un plus grand stockage d'énergie. Les chauffe-eau haute température sont donc des équipements idéaux dans le cadre de programme de gestion de la demande.

Il existe également une littérature concernant le contrôle ou la simulation d'une population de chauffe-eau. Dans Wong et al. (2013), les modèles de chauffe-eau sont simplifiés afin de simuler le comportement d'un grand nombre de ceux-ci. La température des chauffe-eau est représentée à l'aide d'une seule température moyenne. La simulation des tirages d'eau se fait à l'aide d'une chaîne de Markov (Laurent et Malhame 1994). Les paramètres de cette chaîne de Markov peuvent être obtenus à l'aide d'algorithme d'identification (Khurram et al. 2020). Cette même chaîne de Markov régit l'ensemble des chauffe-eau composant la population. On obtient à l'aide de

simulations une distribution des températures des chauffe-eau pour différent moment de la journée. Une limitation de ce modèle est que la distribution de la population de chauffe-eau obtenue sous-représente les cas extrêmes. Effectivement, la probabilité de tirage d'eau est la même pour tous les chauffe-eau. Ceci est important, car la distribution des températures permet de déterminer l'état énergétique des chauffe-eau et les chauffe-eau dont l'état énergétique est trop faible devront nécessairement chauffer et cela modifie donc la séquence de contrôle de la population de chauffe-eau. Les contraintes liées à la distribution de l'état énergétique des chauffe-eau sont établies suite à une simulation utilisant une chaîne de Markov pour plusieurs solutions de contrôle proposés dans la littérature (Nazir, Galiana, et Prieur 2016; Tammam, Anjos, et Gendreau 2020; Wong et al. 2013). L'approche méthodologique utilisée dans ce mémoire se base sur l'extraction de la distribution des états énergétiques d'une population de chauffe-eau directement des données historiques de consommation électrique. La distribution des états énergétiques est donc plus évasée que celle obtenue par simulation à l'aide de chaîne de Markov. Cela est fidèle à la dynamique d'une population de chauffe-eau.

2.3 Prévision de la demande électrique et outils d'apprentissage supervisé

Les outils d'apprentissage supervisé (AS) peuvent jouer un rôle dans la décarbonisation du réseau électrique. Dans Rolnick et al (2019), on mentionne que les outils d'AS sont utilisés pour obtenir des modèles de prévision de la demande électrique. La prévision de la demande électrique joue un rôle important dans l'opération du réseau électrique. Une prévision précise de la demande permet, par exemple, d'utiliser la batterie virtuelle de manière optimale. Dans Hong & Fan (2016), les auteurs mentionnent que la recherche d'un modèle parfait de prévision n'existe pas. Les modèles de prévision sont par nature imparfaits. La communauté se concentre de plus en plus sur les prévisions probabilistes. Elles font maintenant partie des compétitions telles que le « *Global Energy Forecasting Competition* ». Ces prévisions incluent un intervalle qui tente de mesurer l'incertitude de la prédiction. Ces intervalles permettent de concevoir des algorithmes de contrôle qui tiennent compte de l'incertitude. Une prévision probabiliste a été utilisée dans ce mémoire afin d'améliorer les algorithmes de contrôle de la batterie virtuelle.

2.4 Gestion de l'incertitude

L'utilisation de la batterie virtuelle est dépendante de la qualité de la prévision de la demande dans plusieurs scénarios d'utilisation. Il est donc important de considérer l'impact des erreurs de prévision de la demande dans le modèle. Dans King & Wallace (2012), les auteurs donnent un tour d'horizon des algorithmes permettant de modéliser l'incertitude. Les modèles stochastiques à deux étapes ont été choisis afin d'anticiper la variation de la demande. Pour implémenter les modèles, les librairies Pyomo et PyPS ont été utilisées (Bynum et al. 2021; Watson, Woodruff, et Hart 2012). Le cadre général de modélisation de l'incertitude est semblable aux travaux de Tammam et al. (2020).

CHAPITRE 3 ORGANISATION DU MÉMOIRE

Ce mémoire par article est constitué d'un seul article qui constitue le chapitre 4. L'essentiel du travail de cette maîtrise se trouve donc dans cet article. L'article a été soumis le 14 novembre 2022 au « Journal of the Operational Research Society ». L'article apparaît dans ce mémoire tel que soumis à l'exception de corrections mineures de style. Il est possible que certaines modifications soient apportées plus tard en fonction des commentaires des réviseurs. La liste des auteurs de cet article est, dans l'ordre : Louis-Philippe Proulx, Miguel. F. Anjos et Michel Gendreau.

Il y aura une discussion générale (Chapitre 5) qui suit l'article et permet d'ajouter des éléments liés à l'implémentation du modèle et qui ne sont pas discutés dans l'article.

CHAPITRE 4 ARTICLE 1: OPTIMIZATION MODEL FOR THE OPTIMAL OPERATION OF A DATA-DRIVEN VIRTUAL BATTERY OF SMART ELECTRIC WATER HEATERS

Journal of the Operational Research Society

Louis-Philippe Proulx, Miguel F. Anjos and Michel Gendreau - November 14, 2022

Abstract

We propose optimization models for the optimal data-driven operation of a virtual battery of smart electric water heaters. The flexibility of this virtual battery could be used by electric utilities or aggregators to provide grid services such as minimizing the peak demand. Using forecasted grid demand and its uncertainty, we evaluate the capability of this virtual battery in a realistic setting, and compare how different algorithms improve key performance indicators. Two main techniques help alleviate the uncertainty: a formulation of the objective function that focuses on what is known and the use of the virtual battery with a rolling horizon and an updated forecast. Surprisingly, our scenario-based stochastic formulation of the virtual battery applied to real historical grid data does not improve on the performance of the deterministic model.

Keywords: Demand Response, Electric Water Heater, Optimization model, Uncertainty, Data-Driven

4.1 Introduction

Many countries are moving toward a net-zero electricity grid. For example, while Canada targets the year 2050 for net-zero GHG emissions for all sectors of the economy, it targets 2035 for a net-zero electricity grid (Canada 2022). Not only does the grid need to be net-zero regarding GHG emissions, but it also needs to support the growth in demand for clean electricity (e.g., electric cars, heat pumps, etc.). To help the expected increase in variable energy sources like solar and wind, tools are needed to mitigate the variability of such sources. Smart residential loads can be used to add flexibility on the demand side. Among these loads, electric water heaters (EWHs) are an important source of flexibility. With high temperature residential EWHs, as much as 10 kWh of flexible energy per unit can be available. When running demand response (DR) program, electric

utilities or aggregators need to control many thousands EWHs. This paper proposes models for controlling a population of EWHs or, equivalently, a virtual battery (VB).

In this paper, we use real data to parametrize our models and to generate results. This includes historical grid data form a local distribution company and smart EWH data installed in real dwellings. This allows us to validate the results using a realistic setting.

4.2 Literature Review

Many approaches exist to control a population of EWHs. In Malandra et al (2020), a coordinator module sent 24-hour mean temperatures to EWH local controllers. The solution minimized the communication between the coordinator and the local controller. It also uses an algorithm to estimate the state-of-charge (SoC) of an EWH from a limited number on temperature sensors. A precise estimate of the SoC is beneficial to better prioritise which EWH should heat. More details on how to estimate the SoC from partial information are provided in Solis (2015). In Tammam et al. (2020), the authors control a population of EWHs to balance renewable energy, but there are significant differences in modelling. First, losses (water draw and heat loss) are modelled with a Markov chain. This gives probabilities for different events (e.g., shower, bath, etc.). However, because all simulated EWHs use the same Markov chain, this leads to an incorrect distribution of EWH mean temperature. This distribution is then used to constrain the charging of the VB. In this paper, we make no assumption on the EWH temperature distribution in the population, but obtain it indirectly from the available data. These authors also used a stochastic formulation to handle the wind forecast. A similar approached was followed for the stochastic formulation in this paper. In Nazir et al. (2016), histograms are used to model the distribution of the mean temperature in the population of EWHs. The tail of histogram is used to determine the constraint on how many EWHs needs to heat. Again, the authors assume that each EHW have the same probability of a water draw for different hours of the day. This is also a similar limitation of this paper. Indeed, by analysing historical data, we clearly see that some EWHs have very high probability of using hot water while other have a very low probability. The resulting distribution is different than one based on a mean value for all EWHs, which underrepresents the lower and higher case. Another limitation of this methodology is to obtain the losses. EWH population can have very different distribution base on time of the year, day of the year, location, type of population. How to calibrate the parameters for

a population of EWHs is not straightforward and is directly addressed by the data-driven approach of this paper.

Forecasts also play a key role in fully using the available flexibility. Often, decisions need to be made on forecasted variables (e.g., weather, expected demand, etc.). In Makridakis, Hyndman, and Petropoulos (2020), the authors mention that prediction intervals are becoming an important metric in forecasting competition. Probabilistic load forecasts (PLF) include prediction intervals, and this information can be used to make better decisions. However, they also explained why prediction intervals are underestimated. In Hong et Fan (2016), several techniques are presented to produce PLF. Many of them use results from past point forecast and past residuals to extrapolate future prediction intervals. Forecast models can rely on machine learning algorithms. In Rolnick et al. (2019), the authors survey possible pathways where machine learning can play a role and forecasting is one of them. However, machine learning models are not necessarily producing better forecasts with time series data than more traditional techniques (e.g., ARIMA) (Hong et Fan 2016). Our paper creates a PLF of the grid demand using available open-source machine learning library to feed the optimization model.

Optimization methods under uncertainty includes many different algorithms that have their strengths and limitations. In Ning and You (2019), the authors present the major class of problems regarding optimization under uncertainty. They show how a data-driven approach can be used to extend their capabilities. Even though data-driven stochastic optimization is not used in this paper, it could be a possible improvement in handling the distribution of the errors in the forecast by the creation of an ambiguity set. However, this involves complex modelling and implementation techniques. In Birge and Louveaux (2011) and King and Wallace (2012), a general framework is given for stochastic programming. They give insights on how modelling under uncertainty is different from sensitivity analysis and “what if” scenarios. While this paper implements a simple stochastic formulation, results are analysed carefully with regards to the uncertainty in the forecast to measure the “real-world” performance.

4.3 Methodology

4.3.1 Estimating the EWH energy losses

To model a population of EWHs, one needs information on EWH energy losses (heat losses and hot water draws). This information is usually not available or can only be estimated from smart EWHs (e.g., flow meters, temperature sensors, etc.). Our approach in this paper is to obtain the losses using the EWH energy consumption. By the energy conservation principle, the energy losses (heat loss and water draw) of an EWH are equivalent to its energy consumption over a certain period. Moreover, energy consumption is a reliable measure available through each smart local EWH controller. However, energy consumption and energy loss do not have the same profile. The energy loss from water draws happens first followed by a recovery period (heating period). For example, let us assume someone takes a bath. There will be a significant loss of energy in the EWH in a short period during that event. After a few minutes, the EWH will recover by heating with a constant power for a usually much longer period than the water draw.

To estimate the EWH energy losses, a transformation function is applied to the 24-hour power profile of each EWH. We must first choose historical power profiles when no demand response (DR) was applied. Each heating sequence is transformed into an equivalent one-time loss of the same energy. These one-time losses are located close to the start of the corresponding heating period (see Figure 4.1). The exact location of the one-time losses is adjusted once for all EWHs in the calibration phase.

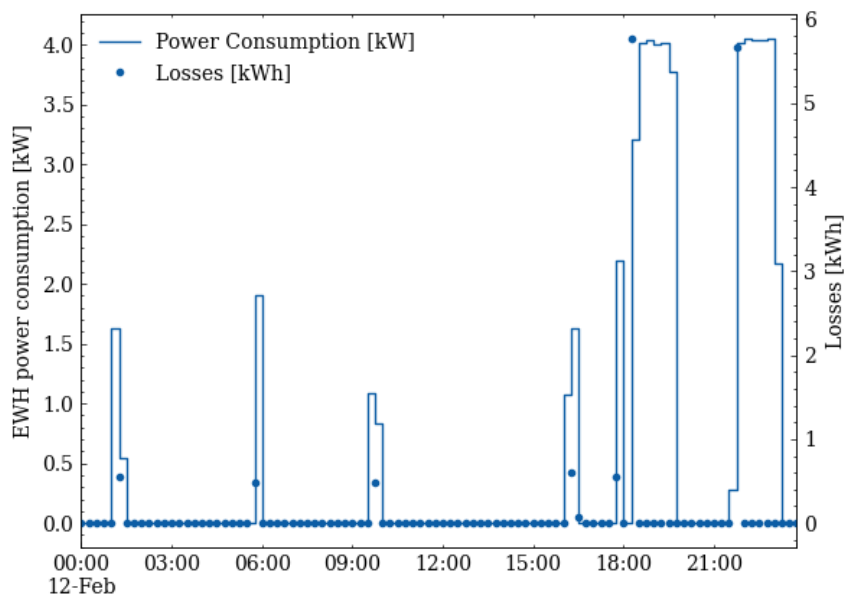

Figure 4.1 Example of Generating EWH Losses from Power Consumption

The one-time losses give an estimate of the state of charge (SoC) of each EWH. This is how the loss parameter is obtained. A possible future improvement of this transformation would be to separate heat losses from water draw losses. The two losses would be calibrated to match the total energy of the EWH over 24 hours. In Section 4.4, more details are given on the calibration of this parameter.

4.3.2 Deterministic Optimization Model

A continuous deterministic linear optimization model (LP) was used as a base case. The purpose of the model is to abstract the control of a population of EWHs as a virtual battery (VB). The model uses historical power consumption collected from real EWHs. This optimization model could be used to control millions of EWHs because only a small subset of EWHs is needed to parametrize the model. We found that a few hundred power profiles are sufficient to grasp the dynamics of the population.

The implementation was done using Pyomo (Bynum et al. 2021), an algebraic modelling language in Python that can interface with various solvers. For example, in this paper we used a sample size of 300 EWHs to control 20,000 EWHs. The decision variables represent the amount of energy delivered to each EWH in the sample at every time step (e.g., 15 minutes). Each energy input into the sample is scaled to an equivalent action on the VB (i.e., the population of EWHs). The charging

of the VB is dependent on the objective function, e.g., minimizing the peak. The sample size of EWHs must be large enough to be representative of the population while not too large that solving the model is slow. This is even more important if we reuse this base model in a stochastic formulation with multiple scenarios. The comfort of the user is ensured through a constraint that limits the number of EWHs with a negative SoC. Any time an EWH has a negative SoC, it is constrained to heat. We also assume that a certain amount of energy is available from each EWH; this can be estimated based on the physical characteristics of the EWH, the minimum temperature to prevent the growth of Legionella (Lacroix 1999), and having enough hot water available to the user. Four objective functions (OF) have been devised. The first one aims at minimizing the peak demand of the distribution grid; the second one is minimizing the variation of the grid demand; the third one is used to validate the model by maximizing the SoC of all EWHs; and finally, a fourth is minimizing the cost of charging the VB based on a ToU tariff (peak hour from 6 AM to 9 AM and 4 PM to 8 PM). Finally, one notes that, while the solution prescribes an overall amount of energy to input into the VB at each time step, in a real DR program, an intermediate step is needed to decide which EWHs will heat based on their SoC. Such an intermediate controller could allow colder EWH to heat first. We add a term in the objective function to ensure that all EWHs will heat as soon as possible as long as it is not detrimental to the main objective (e.g., peak minimization). The multiplying factor in this term is lower at the start of the time period, meaning that adding energy to the VB costs slightly less when done early. It keeps the VB SoC as high as possible. Additionally, the VB SoC needs to be maximal at the last time step.

The optimization model is as follows:

$$min\, p_{\max}^{AM} + p_{max}^{PM} + \frac{1}{1000 * \Delta t} \sum_{t} \frac{t}{I * T} x_t^{batt} \tag{O.1}$$

$$min \sum_{t} \mathrm{z_t} + \frac{1}{1000 * \Delta t} \sum_{t} \frac{t}{I * T} x_t^{batt} \tag{O.2}$$

$$max \sum_{t} e_t^{batt} \tag{O.3}$$

$$min \sum_{t} c_t x_t^{batt} + \min(\mathrm{c_t}) * \sum_{t} \frac{t}{I * T} x_t^{batt} \tag{O.4}$$

$s.t.$

$$p_t \leq p_{max}^{AM}, \qquad t = 1..\frac{T}{2} \tag{C.O1-A}$$

$$p_t \leq p_{max}^{PM}, \qquad t = \frac{\mathrm{T}}{2} + 1..T \tag{C.O1-B}$$

$$z_t \geq p_t - p_{t-1}, \ \ \mathrm{t} > 2 \tag{C.O2-A}$$

$$z_t \geq p_{t-1} - p_t, \ \ t > 2 \tag{C.O2-B}$$

$$p_t = d_t^{forecast} + \frac{1}{1000 * \Delta t} x_t^{batt} - d_t^{EWHs}, \qquad \forall t \tag{C.2}$$

$$e_t^i = e_t^{i^+} - e_t^{i^-}, \qquad \forall i, \forall t \tag{C.3}$$

$$e_t^i = e_{t-1}^i + x_t^i - \ l_t^i, \qquad \forall i, \forall t \tag{C.4}$$

$$e_0^i = e_{init_{SoC}}^i, \qquad \forall i \tag{C.5}$$

$$e_t^{i^-} \leq e_t^{i\bar{}_{NODR}}, \qquad \forall i, \forall t \tag{C.6}$$

$$e_t^{i^+} \leq e_{max}^i \ , \qquad \forall i, \forall t \tag{C.7}$$

$$x_t^i \leq H^i * \Delta t, \qquad \forall i, \forall t \tag{C.8}$$

$$x_t^{batt} = \frac{N^{EWH}}{I} \sum_{i=1}^{I} x_t^i \ , \qquad \forall t \tag{C.9}$$

$$e_t^{batt} = \frac{N^{EWH}}{I} \sum_{i=1}^{I} e_t^i, \qquad \forall t \tag{C.10}$$

$$e_t^{batt} \geq 0, \qquad \mathrm{t} = 1..\mathrm{T} - 1 \tag{C.11}$$

$$e_t^{batt} \geq \frac{N^{EWH}}{I} \sum_{i=1}^{I} e_t^{i_{NODR}}, \qquad \mathrm{t} = \mathrm{T} \tag{C.12}$$

All variables are positive and continuous, except for the SoC of each EWH, e_t^i, which is replaced by positive and continuous variables using the standard change of variables $e_t^i = e_t^{i^+} - e_t^{i^-}$.

Our model has four objectives. Objective 1 together with constraints (C.O1-A) and (C.O1-B) minimizes the maximum AM and PM peaks; (C.O1-A) and (C.O1-B) define AM and PM peaks from 00:00 to 12:00 and 12:00 to 24:00 respectively. Objective 2 together with constraints (C.O2-A) and (C.O2-B) minimizes the absolute value of the variability of the expected grid demand. Objective 3 maximizes the SoC of the VB. Under objective 3, EWHs will heat as soon as possible, a behaviour that is similar to the normal operation of a population of EWHs. It is used to validate the parameters of the optimization model. More details are given in Section 4.4. Objective 4 minimizes the cost of charging the VB based on ToU tariff. This is different than objectives 1 and 2, because the forecasted grid demand has no influence on the operation of the VB. Objective 4 is an interesting comparison to objectives 1 and 2, because it is a realistic setting for many aggregators that operate under a cost signal. Constraint C.2 states that the expected grid demand is the sum of forecasted grid demand and power input to the virtual battery minus the power consumption of the population of EWHs when there is no DR. The constant term $\frac{1}{1000*\Delta t}$ transforms the energy input (kWh) in the VB to an equivalent power input (MW). In constraint 3, the SoC is replaced by the difference of two positive variables. Constraint C.4 represents the SoC of each EWH. The SoC is the result of the previous SoC, the energy added with heating minus energy losses. In constraint C.5, we initialize the SoC of each EWH in the sample. As for constraint C.6, it states that the negative SoC during DR cannot exceed the negative SoC when there is no DR. A negative SoC means the SoC is below zero. If the SoC is positive, the negative SoC is zero. The negative SoC when there is no DR parameter is obtained for all EWHs in the sample assuming they can always heat. In constraint C.7, each EWH SoC cannot exceed his maximum SoC at any time. In constraint

C.8, the energy added for each EWH is a function of the power of the heating element and the duration of the time step (e.g., with a 15-minute step and a heating power of 4 kW, this would constrain the energy added between 0 kWh and 1 kWh). In constraint C.9, we scale the energy added into the EWH sample to the size of the EWH population, i.e., the VB. Similarly, in constraint C.10, we obtain the SoC of the VB from the mean of all EWH SoC in the sample. Finally, for constraints C.10 and C.11, the SoC of the VB must always be kept above zero and, at the last time step, the SoC of the VB must be back at full charge (like when there is no DR).

4.3.3 Probabilistic Load Forecast

The main source of uncertainty for the optimization model is the grid demand. Objectives 1 and 2 base their decisions to charge the VB on the forecast of the grid demand. Forecasted peak and actual peak are most often not happening at the same time and are not of the same size. Even if our EWH VB model is accurate, inaccurate forecasts would lead to bad decisions on when to charge it. However, perfect point forecast is not realistically achievable. The input data available does not contain all the information needed to create a perfect forecast. Instead, a probabilistic load forecast (PLF) is used to estimate a prediction interval of the forecasted grid demand for each hour. In Section 4.3.4, more information will be given on how this interval is used to improve the optimization model.

To forecast the grid demand, we use a machine learning model that relies on multiple inputs (or features). These inputs are the past 24-hour grid demand, the forecasted weather information (temperature, relative humidity), time of the day and day of the week. This machine learning model implements a tree-based learning algorithm (Ke et al. 2017). Real historical data was used to train the model and hyper-parameters were optimized by grid searching. The forecast model outputs the next 24-hour grid demand with its prediction interval by recursively applying the model 24 times. The implementation uses the python library developed by Amat Rodrigo & Escobar Ortiz (2022). The hourly forecast is also interpolated to a 15-minute time step (96 values). Prediction intervals are estimated by bootstrapping (Hyndman and Athanasopoulos 2021).This provides a prediction interval with a 95% confidence. Even though this 95% prediction interval estimate is too narrow for many reasons, see (Makridakis et al. 2020). It is still useful information for the optimization model. In Figure 4.2, we give an example of a 24-hour forecast. In this example, we see that the prediction interval is no longer valid in the afternoon. A rolling horizon where the forecast is

updated every hour based on the latest data available was used to mitigate that problem and will be discussed in the next section.

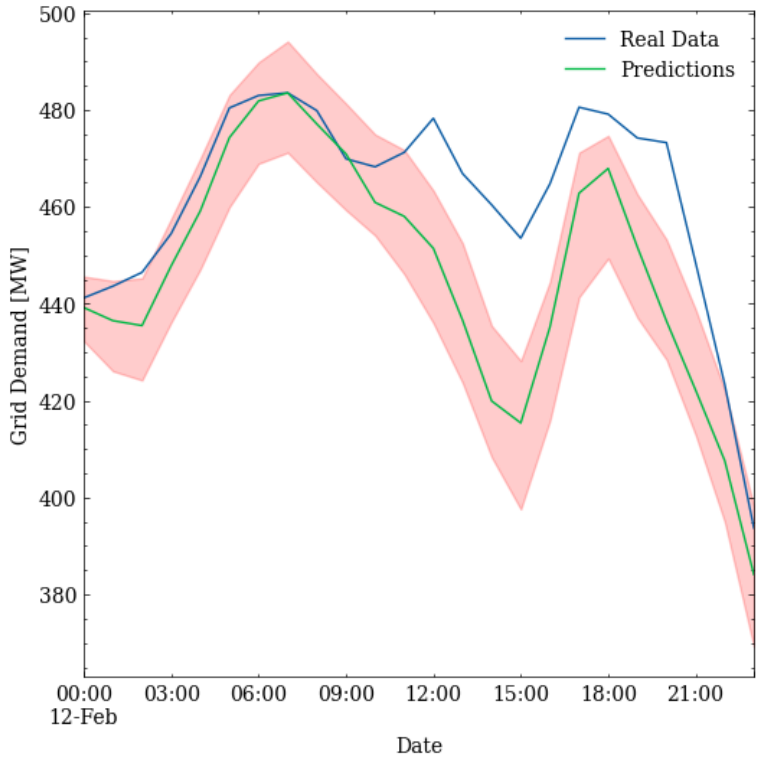

Figure 4.2 24-hour Forecast

4.3.4 Rolling Horizon and Stochastic Optimization Model

To mitigate the forecasting error of the grid demand, we experiment with two methods: a rolling horizon algorithm and a rolling horizon with a stochastic formulation of the optimization model. The rolling horizon consists of updating the forecast every hour and reapplying the deterministic model by keeping past decisions. For the stochastic formulation, we used a similar approach, but with a 2-stage stochastic formulation. The first-stage decision is to choose the charging of the VB for the next hour (four 15-minute time steps). The second-stage decision is for the remaining hours and is based on three scenarios from the forecasted prediction intervals: high, average and low forecast. These scenarios are created based on the information of the PLF. For example, the demand variation objective becomes:

$$min \sum_{t_h}^{t_h+3} \mathrm{Z_t} + \frac{1}{1000 * \Delta t} \sum_{t_h}^{t_h+3} \frac{t}{I * T} x_t^{batt} + \sum_{\omega \in \Omega} E\left[Q(x, \xi)\right]$$

The parameter t_h represents the current time step; the first-stage decision variables are $\{x_h, x_{h+1}, x_{h+2}, x_{h+3}\}$ and $\{x_{h+4}^{\omega}, \dots, x_T^{\omega}\}$ are the second-stage decision variables. Stochastic parameters represent the forecasted demand for the three scenarios, $\xi = d_{h+4}^{Forecast^{\omega}}, \dots, d_T^{Forecast^{\omega}}$. Only the first-stage decisions are implemented.

$$OF\ of\ Q(x,\xi) = min \sum_{t_{h+4}}^{T} \mathrm{Z}_{\mathrm{t}}^{\omega} + \frac{1}{1000 * \Delta t} \sum_{t_{h+4}}^{T} \frac{t}{I * T} x_t^{batt^{\omega}}$$

Constraints in $Q(x,\xi)$ are modified similarly, for example:

$$p_t = d_t^{all} + \frac{1}{1000 * \Delta t} x_t^{batt} - d_t^{EWHs}, \qquad \mathrm{t} = 1, \dots, h+3$$

$$p_t = d_t^{Forecast^{\omega}} + \frac{1}{1000 * \Delta t} x_t^{batt^{\omega}} - d_t^{EWHs}, \qquad \mathrm{t} = \mathrm{h}+4, .., \mathrm{T}, \omega = 1,2,3$$

The extended form of the stochastic model is solved every hour. The probability of the low and high scenarios is 25% and 50% for the average scenario.

4.3.5 Strengths and Limitations

By using historical data to model EWH losses, the model is highly adaptable to any population of EWHs. EWH power consumption and associated losses are dependent on many variables: hot water usage, cold water temperature, time of the year, time of the week (weekday vs weekends), holiday, etc. In this model, we used historical power consumption from a few days before the DR events, consequently the parameters reflect the current state of the EWH population. Another benefit of our approach is that we only need the power consumption, an information that is usually readily available with smart EWHs. We do not make any assumptions about water usage, but compute it directly from the power consumption data.

The main limitation is that the aggregator needs information about the power consumption, and this depends on the willingness of participants to share their data. In this paper, we use a sample size of 300 EWHs, and our results confirm that a sample of a few hundred EWHs suffices to control a population of many thousands or millions of EWHs, if the sample is representative of the population. In that case, the data needs to be collected only from the sample.

4.4 Calibration of the Optimization Model

To calibrate the parameters of the optimization model, we used objective 3 that maximizes the SoC of the VB and compared the results to the historical data. Objective 3 is equivalent to the normal operation of an EWH where it starts heating after it measures a temperature drop. The calibration helps identify two parameters: the capacity of the heating element H^i and the energy loss of each EWH l_t^i. The energy loss is obtained from a function that transforms the EWH power consumption (see Figure 4.1), but the timing needs to be slightly adjusted so energy consumption from the model is consistent with the actual data (see Figure 4.3). This 5- or 10-minute adjustment is performed only once.

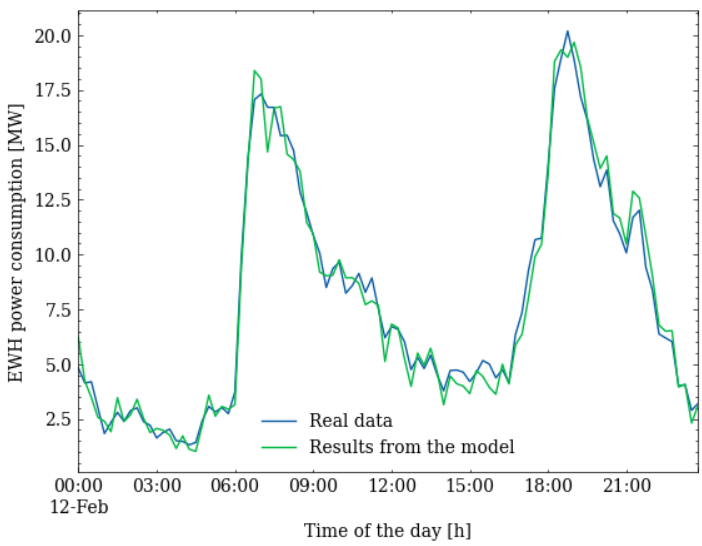

Figure 4.3 Validation of calibration with objective 3

We used multiple datasets for this paper: historical power consumption of a population of EWHs, historical grid demand from a local distribution company and historical weather information (see Table 4.1).

Table 4.1 Dataset Information

Dataset	Time interval	Type	Variables
EWH	2 years of data at 5 min time step	60 imp gal EWH	Power consumption [kW]
Grid demand	2 years of data at 15 min time step	Historical demand from the local distribution company	Demand [MW]
Weather	Hourly time step	Historical air temperature	Temperature [°C]

We have a dataset of around 65 EWHs with a capacity of 60 imperial gallons. We use a sample size of 300 EWHs for the optimization model. To collect enough 24-hour EWH profiles in the sample, we collect data from a few days prior to the test day. We always skip weekend days. Our EWH dataset only contains non-DR day. The number of EWHs in the VB is 20,000, because it is a realistic considering the size of the distribution grid. Table 4.2 gives the value of all parameters that are kept constant through this paper.

Table 4.2 Parameters Used in this Paper

Parameters	I	T	N^{EWH}	H^i	ΔT^i	V_{60G}	$e^i_{init_{SoC}}$	$e^i_{\max}$
Value	300	96	20000	3.8 kW	10°C	279 L	3.19 kWh	3.19 kWh

4.5 Computational Results

In this section, we present results going from the simplest to more complex models. We start with the deterministic model where the grid forecast is replaced by historical data. This allows us to

validate the performance of the model for different objective functions with perfect information about the grid demand. As a second step, we use deterministic models, but with a realistic (or imperfect) grid forecast to measure the impact on key performance indicators (KPIs). Thirdly, we present the result using the hourly rolling horizon and the corresponding updated grid forecast. Finally, we add the stochastic model to the rolling horizon and forecast.

Table 4.3 gives the list of KPIs. To compute realistic KPIs, we use the actual grid demand for the day (d_t^{Real}) and not the forecasted grid demand. For all KPIs, we average the results of applying the model to the same 5 days from 5 different weeks from January 22, 2019, to February 19, 2019.

Table 4.3 Key Performance Indicators Description

Key Performance Indicators	Description	Units	Formula $p_t^{Real} = d_t^{Real} + \frac{1}{1000 * \Delta t} x_t^{batt} - d_t^{EWHs}$
AM Peak	Grid peak between 00:00 and 12:00	MW	$\max(p_t^{Real}), t = 0..\frac{T}{2}$
PM Peak	Grid peak between 12:00 and 24:00	MW	$\max(p_t^{Real}), t = \frac{T}{2}..T$
Variability	Net demand variation	MW	$\sum_{t=2}^{T} \lvert p_t^{Real} - p_{t-1}^{Real} \rvert$

4.5.1 Results of the Deterministic Model with Perfect Forecast

In this section, the deterministic model is used with a perfect forecast ($d_t^{Forecast} = d_t^{Real}$). This gives the best possible KPIs and allows us to measure how close other methods are when using the forecasted grid demand. Table 4.4 presents the KPIs for the 5 DR days. The OF to minimize variability performs as well as the OF for peak minimization with respect to AM and PM Peak. This is related to the fact that peaks are associated with a large variation in the demand. So, by minimizing variation, AM and PM peaks are also minimized.

Table 4.4 KPIs with Perfect Information

	AM Peak [MW]	PM Peak [MW]	Variability Index [MW]
Historical data	470.9	473.0	302.1
OF: Peak	461.5	464.7	241.2
OF: Variability	461.5	464.7	145.9

However, the SoC of the VB is very different between both (see Figure 4.4 and Figure 4.5). With the OF to minimize variability, the SoC of the VB is almost fully discharged and charged twice through the day. For peak minimization, the VB SoC remains high, only decreasing to reduce peaks.

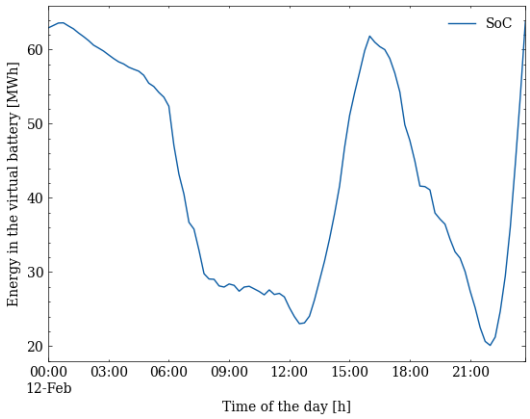

Figure 4.4 VB SoC under Net Demand Variation

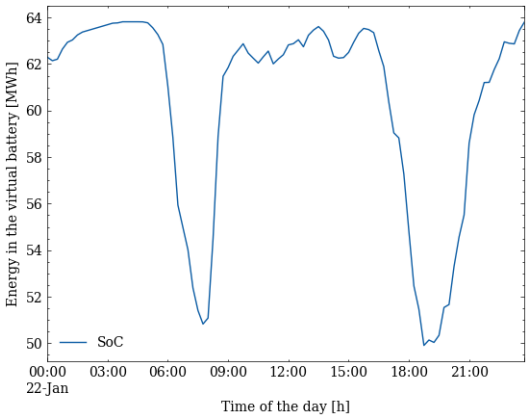

Figure 4.5 VB SoC under Peak Minimization

Figure 4.6 and Figure 4.7 show the grid demand after controlling the VB.

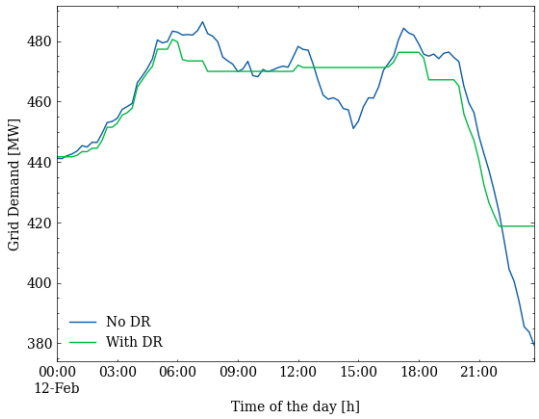

Figure 4.6 Net Demand Variation

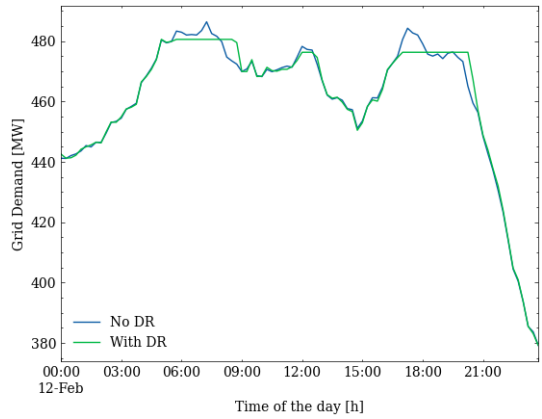

Figure 4.7 Peak minimization

Figure 4.8 and Figure 4.9 give the charging of the VB. This charging could be seen as the instruction given to an aggregator. The aggregator would use that information to implement the DR program.

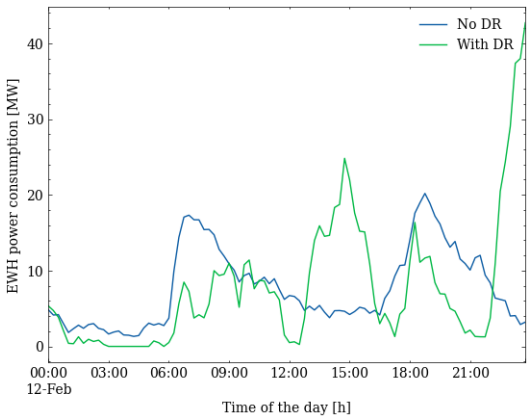

Figure 4.8 VB Charging under Net Demand Variation

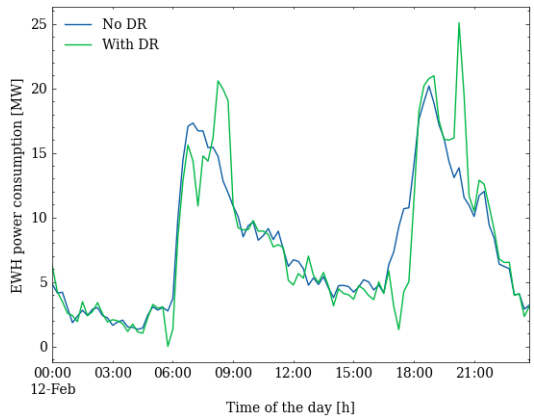

Figure 4.9 VB Charging under Peak Minimization

4.5.2 Results of the Deterministic Model with a Realistic Forecast

When using a forecast instead of the actual grid demand, we see a decline in KPIs (see Table 4.5), primarily for the peak minimization model. In fact, with the peak minimization OF, it is doing worst on average than historical data. With the variability OF, it is achieving a peak reduction of 4.2 MW and 3.1 MW for AM and PM peaks. This needs to be compared to Table 4.4, with perfect information, where peak reductions are 9.4 MW and 8.3 MW.

Table 4.5 KPIs for Deterministic Models

	AM Peak [MW]	PM Peak [MW]	Variability Index [MW]
Historical data	470.9	473.0	302.1
OF: Peak	472.2	473.8	336.2
OF: Variability	466.7	469.9	276.7

Figure 4.10 and Figure 4.11 are examples of the result of applying the deterministic model with a forecasted grid demand. In Figure 4.11, there is a rebound after the PM peak that creates a second peak higher than the first one.

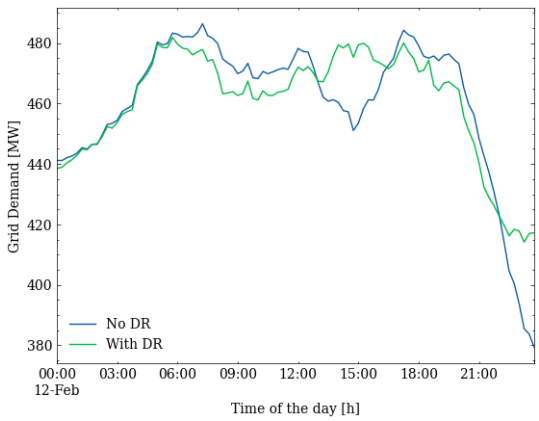

Figure 4.10 Net Demand Variation Using Forecasted Grid Demand

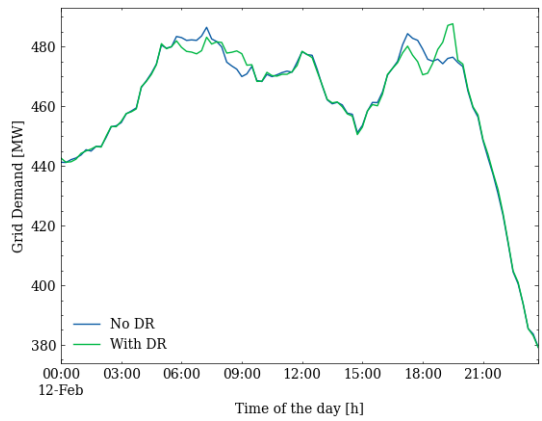

Figure 4.11 Peak Minimization Using Forecasted Grid Demand

4.5.3 Impact of the Rolling Horizon and the Updated Forecast

Every hour, the deterministic model is applied with an updated forecast. This strategy improves the KPIs of both OFs. With the variability OF, it is now achieving a peak reduction of 6.7 MW and 4.5 MW for AM and PM peaks, a substantial improvement over the deterministic model applied only once. The peak minimization OF is now doing better than without DR.

Table 4.6 KPIs for Deterministic Models with Rolling Horizon

	AM Peak [MW]	PM Peak [MW]	Variability Index [MW]
Historical data	470.9	473.0	302.1
OF: Peak	469.5	471.4	317.6
OF: Variability	464.2	468.5	282.0

The rolling horizon is not reducing the variability index as errors in the next hour forecast are driving the variability. In Figure 4.13, the peak minimization is still highly dependent of forecast errors.

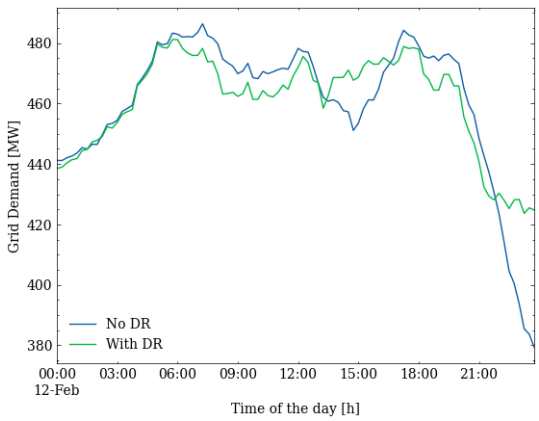

Figure 4.12 Net Demand Variation with Rolling Horizon

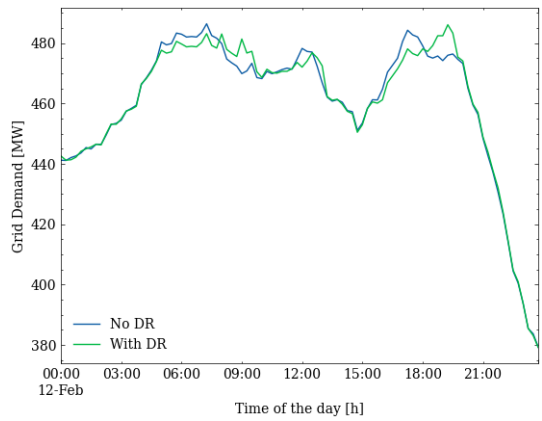

Figure 4.13 Peak Minimization with Rolling Horizon

4.5.4 Adding the Stochastic Formulation to the Rolling Horizon

Results when adding the stochastic formulation to the rolling horizon are close to the deterministic model with the rolling horizon. The first-stage decision is highly dependent on the next-hour forecast. Overall, results from the stochastic models are similar to those obtained for the rolling horizon with the deterministic models.

Table 4.7 KPIs for Stochastic Models with Rolling Horizon

	AM Peak [MW]	PM Peak [MW]	Variability Index [MW]
Historical data	470.9	473.0	302.1
OF: Peak	469.6	470.0	346.3
OF: Variability	464.9	468.0	295.8

4.5.5 With ToU Rates

DR programs are often based on ToU rates or dynamic rates. One benefits of ToU rates for the aggregator is the certainty on the forecasted cost. In conjunction with objective 4, we used a high tariff during AM and PM peak to investigate the impact on the grid.

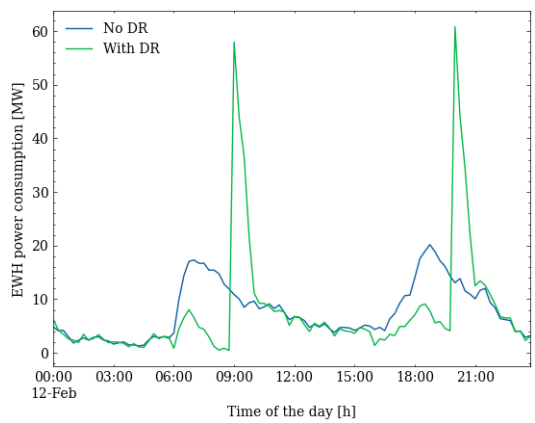

Figure 4.14 Charging of VB under ToU rates

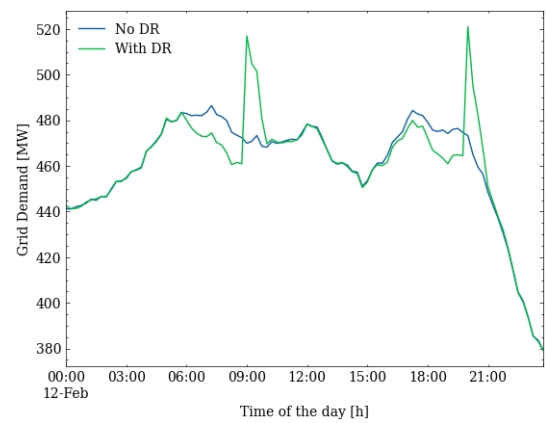

Figure 4.15 Impact on the Grid Demand for ToU Rates

ToU rates need to be devised carefully, as we can see in Figure 4.15, because it can create a higher peak at the end of the high tariff.

4.6 Conclusion

We have proposed two optimization models that can be used to control a population of EWHs with the aim of reducing peak demand or demand variation in an electrical grid. Through computational

experiments based on real data, we showed that these models, if properly calibrated, can lead to an improvement of the target KPIs of the distribution grid under consideration. Two modelling/implementation techniques lead to noticeable improvement in the results. The first of these techniques is using an OF that minimizes the net demand variation instead of AM and PM peak. This OF is more robust to errors in the forecast. The forecast is more accurate in predicting if the grid demand will increase or decrease than it is in predicting the absolute value of the demand. The net demand variation can take advantage of that fact. The second technique consists in reapplying the deterministic model every hour in a rolling horizon fashion. The main advantage of this approach is a more accurate forecast. The optimization model takes less than 3 seconds to run on a PC with 300 EWHs in the sample and 96-time steps. Resorting to a stochastic formulation does not further improve the results, because the first-stage decision is still dependent on the forecast. This leads to the conclusion that, for an aggregator, good results can be obtained by applying the approach proposed in this paper without having to devise a complex stochastic model.

Acknowledgements: Financial support for this research was provided by Natural Resources Canada through the Program on Energy Research and Development. The authors wish to thank Hydro-Sherbrooke, a local utility, for its support of this research by providing data. We also thank TeeGarden for providing the electric water heater data

Appendix

List of symbols

SoC: State of charge

VB: Virtual Battery

EWH(s): Electric water heater(s)

DR: Demand response

KPI: Key performance indicators

OF: Objective function

ToU: Time-of-Use

PLF: Probabilistic Load Forecast

Indices

$t = 1..T, 15\,\text{min}\, time\ step\ over\ 24h$

$i = 1..I, for\ each\ EWH\ in\ the\ sample$

Parameters

$N^{EWH}: Number\ of\ EWHs\ in\ the\ virtual\ battery$

$I: Number\ of\ EWHs\ in\ the\ sample$

$T: Number\ of\ time\ step\ in\ the\ optimisation\ model$

$\Delta t: Time\ duration\ of\ one\ time\ step\ \ [h]$

$H^i:\ Capacity\ of\ the\ heating\ element\ of\ EWH\ i\ \ [kW]$

$C^p: Specific\ heat\ of\ hot\ water\ \left[\frac{kWh}{kg.K}\right]$

$\rho: Water\ density\ \left[\frac{kg}{l}\right]$

$V_{60G}: Volume\ of\ hot\ water\ for\ one\ 60\ gallon\ EWH[l]$

$\Delta T^i: Flexible\ temperature\ interval\ available for\ each\ EWH\ i\ [°C]$

$e_t^{i_{NODR}}: EWH's\ SoC\ when\ there\ is\ no\ DR\ at\ time\ step\ t\ \ [kWh]$

$e_t^{i_{\bar{N}ODR}}$: EWHs with a $negative\ SoC\ when\ there\ is\ no\ DR\ at\ time\ step\ t\ [kWh]$

$e^i_{\max}$: $Maximum\ usable\ energy\ of\ EWH\ i\ \ [kWh]$ $(e^i_{max} = C^p \rho V_{60G} \Delta T^i, \forall i)$

$e^i_{init_{SoC}}$: $Initial\ SoC\ of\ each\ EWHs\ \ [kWh]$

l^i_t: $Energy\ loss\ of\ EWH\ i\ at\ time\ step\ t\ after\ transformation\ of\ the\ power\ consumption\ profile\ [kWh]$

$d^{Forecast}_t$: $Forecasted\ grid\ demand\ at\ time\ step\ t\ [MW]$

d^{Real}_t: $Real\ grid\ demand\ at\ time\ step\ t\ [MW]$

d^{EWHs}_t: $Power\ consumption\ of\ the\ population\ of\ EWHs\ when\ there\ is\ no\ DR\ at\ time\ step\ t\ [MW]$

Variables

Grid variables

p^{AM}_{max}: $Highest\ peak\ of\ the\ morning$ [MW]

p^{PM}_{max}: $Highest\ peak\ of\ the\ afternoon$ [MW]

p_t: $Expected\ grid\ demand\ at\ time\ step\ t$ [MW]

c_t: $cost\ of\ electricity\ at\ time\ step\ t\ [\$/kWh]$

Electric water heater sample variables

e^i_t: $EWH\ SoC\ for\ EWH\ i\ at\ time\ step\ t\ \ [kWh]$

$e^{i^+}_t$: $Positive\ EWH\ SoC\ \ at\ time\ step\ t\ [kWh]$

$e^{i^-}_t$: $Negative\ SoC\ of\ EWH\ i\ at\ time\ step\ t\ [kWh]$

x^i_t: $Energy\ input\ in\ EWH\ i\ at\ time\ step\ t\ [kWh]$

Virtual battery variables

e^{batt}_t: $SoC\ of\ the\ EWH\ virtual\ battery\ at\ time\ step\ t\ [kWh]$

x^{batt}_t: $[Decision\ variable] Energy\ input\ in\ the\ VB\ at\ time\ step\ t\ [kWh]$

CHAPITRE 5 DISCUSSION GÉNÉRALE

Le modèle de batterie virtuelle nous permet d'obtenir une commande globale de puissance qui doit être envoyée à la population de chauffe-eau. Cette commande tient compte du comportement de la population de chauffe-eau. L'énergie totale envoyée durant l'événement correspond à l'énergie normalement consommée par cette population de chauffe-eau. Cependant, la distribution de cette commande globale à chacun des chauffe-eau n'est pas définie par le modèle. Dans la Figure 5.1, on représente l'état de charge des chauffe-eau à 15 h lors d'une simulation où l'objectif est de minimiser la variabilité de la demande. On remarque que les chauffe-eau avec un état de charge plus faible ont une plus grande probabilité de chauffer. Cependant, on remarque qu'un certain nombre de chauffe-eau avec un état de charge plus faible ne chauffent pas. En effet, le modèle d'optimisation anticipe les pertes thermiques de chaque chauffe-eau et peut ainsi prioriser des chauffe-eau qui subiront bientôt des pertes énergétiques importantes afin que leur état de charge reste positif.

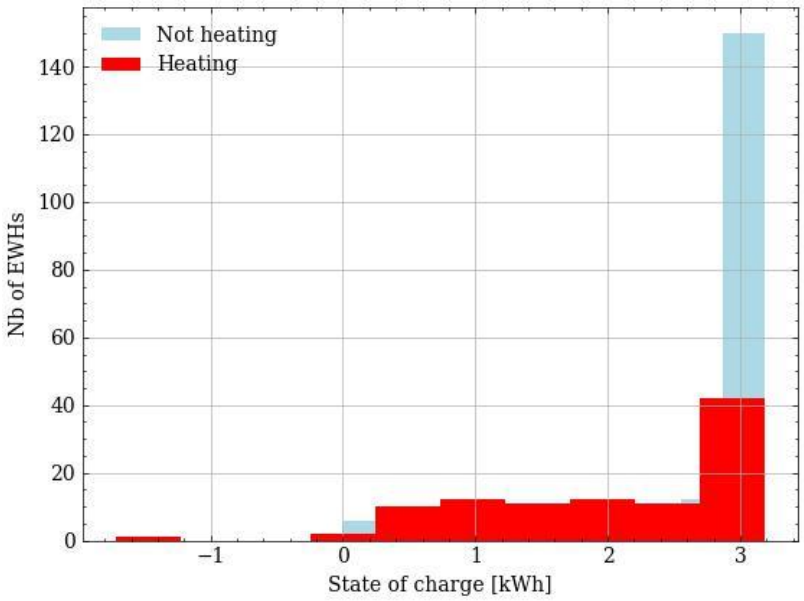

Figure 5.1 Répartition de l'état de charge de la population de chauffe-eau

Lors d'essais en condition réelle, une première approche d'implémentation serait d'utiliser la commande globale et de la distribuer aux chauffe-eau les plus froids. Cela ne reproduirait pas parfaitement le fonctionnement de la batterie virtuelle, mais la plupart des chauffe-eau se

rapprocheraient d'un état énergétique moyen qui offre un maximum de flexibilité. Pour certains chauffe-eau qui n'auraient pas été chauffés en prévision d'une perte énergétique importante, leurs états de charge se retrouveraient alors plus faibles que prédit. Une autre alternative serait d'ajouter un ensemble de contraintes à la batterie virtuelle afin que la commande globale soit redistribuée de façon prioritaire aux chauffe-eau froids. Le comportement serait alors équivalent à la façon de les contrôler lors d'essais en condition réelle. Afin d'inclure cette contrainte, il faudrait ajouter des variables entières afin de comparer l'état de charge des chauffe-eau. La complexité du modèle serait augmentée. Il s'agirait alors d'un modèle d'optimisation linéaire à nombre entier.

Il appert également important de préciser le concept d'état de charge négatif. Le modèle attribue une certaine quantité d'énergie flexible à chaque chauffe-eau selon ses caractéristiques physiques. Par exemple, pour un chauffe-eau haute température qui peut atteindre une température de 75°C, une plage de 10°C (de 75°C à 65°C) pourrait correspondre à un bloc d'énergie flexible d'environ 3 kWh selon ses caractéristiques physiques. Bien que le modèle n'utilise pas la température, il est utile d'établir une correspondance entre la température moyenne du chauffe-eau et son état énergétique. Un état énergétique maximal (par exemple, 3 kWh) correspond à 75°C dans notre exemple. Un état énergétique nul correspond à la limite inférieure de la plage de température (dans ce cas-ci 65°C) et un état énergétique négatif correspond à une température inférieure à 65°C. Lorsqu'un chauffe-eau atteint un état énergétique négatif, il est alors contraint de chauffer. Autrement dit, le contrôleur peut décider d'éteindre un chauffe-eau uniquement s'il a un état énergétique positif.

Au lieu de gérer l'incertitude à même le modèle de batterie virtuelle, il est fort possible dans le contexte des réseaux électriques intelligents que la commande reçue par la batterie virtuelle ne soit pas la prévision de la demande, mais un signal de prix. La 4e fonction objective du modèle permet de réutiliser la batterie virtuelle dans ce contexte. Une grande partie de la gestion de l'incertitude repose alors sur l'entité qui génère le signal de prix. La batterie virtuelle permet alors de minimiser les frais de fonctionnement des chauffe-eau tout en ayant un impact positif sur l'opération du réseau électrique. Cependant, si le signal de prix est mal défini, cela pourrait avoir des impacts négatifs importants sur le réseau électrique (p.ex., effet de rebond après la période de pointe).

CHAPITRE 6 CONCLUSION ET RECOMMANDATIONS

Nous espérons que le modèle de batterie virtuelle présenté dans ce mémoire contribuera à l'utilisation des chauffe-eau comme outil de gestion de la demande par les réseaux électriques. Ce modèle permet d'être configuré à l'aide des données historiques de consommation d'énergie des chauffe-eau. Il permet ainsi de modéliser une grande variété de population de chauffe-eau si les données sont disponibles. Il permet également de modéliser une même population pour des conditions d'opération changeante. Les pertes énergétiques d'une population de chauffe-eau varient selon plusieurs conditions : le moment de l'année, s'il s'agit d'un jour de semaine, de fin de semaine, d'une journée férié, etc. Avec l'avènement des chauffe-eau intelligents, les données de consommation seront disponibles en grand nombre. Avec les données d'un sous-ensemble de la population de chauffe-eau, on pourra former une batterie virtuelle adaptée à la population de chauffe-eau qu'on veut contrôler. De plus, quatre fonctions objectives sont proposées pour adapter le modèle à différent cas d'utilisation. Si la batterie virtuelle doit être intégrée à l'opération du réseau électrique, la prévision de la demande peut être traitée avec une fonction objectif qui minimise la variabilité de la demande et être appliquée sous forme d'horizon fuyant. Si l'opérateur du réseau électrique fonctionne avec un signal de coût, la 4e fonction objectif proposée rendra le modèle compatible avec ce mode d'opération. Également, les résultats ont été validés à l'aide de données historiques d'un réseau électrique.

Il faut cependant être conscient des limites et contraintes du modèle de batterie virtuelle. Premièrement, l'opérateur du réseau électrique doit avoir accès aux données de consommation d'un sous-ensemble des chauffe-eau pour créer le modèle. Le modèle peut en théorie fonctionner avec une population hétérogène. Si une grande variété de types de chauffe-eau compose la population, il faudra diviser la population en sous-groupes homogènes ou créer un échantillon représentatif tenant compte de l'importance de chacun de ces types de chauffe-eau. Il faudra évidemment adapter les paramètres physiques du modèle : taille du chauffe-eau, puissance de l'élément chauffant, plage de température flexible, etc. Le modèle n'a pas été validé pour une population hétérogène de chauffe-eau, mais il faudrait alors calibrer le modèle en s'assurant qu'il est en mesure de reproduire le comportement de la population de chauffe-eau. Finalement, une implémentation qui priorise le chauffage des chauffe-eau froids créera une distribution qui diffère de la batterie virtuelle. Certains chauffe-eau pourront se retrouver avec un état de charge plus faible qu'anticipée. Il est donc

possible que dans certains cas la commande globale ne soit pas suffisante pour le nombre de chauffe-eau devant obligatoirement chauffer. Cependant, cet effet devrait être minimal, car ce type de contrôle devrait faire converger l'ensemble des chauffe-eau vers une température moyenne.

Le modèle présenté dans ce mémoire peut être utilisé tel quel, mais un certain nombre d'avenues de recherche sont proposées. L'amélioration du modèle stochastique pourrait être entreprise. Par exemple, une modélisation utilisant le cadre du « distributionally robust optimization » pourrait permettre d'anticiper un ensemble de distribution de la prévision de la demande et se prévenir du pire cas. D'autre part, il est prévu d'utiliser le modèle de batterie virtuelle dans des essais en condition réelle. S'il s'avère que la commande globale est parfois non réalisable en raison de la distribution des températures dans la population des chauffe-eau. Le modèle d'optimisation pourrait être transformé en modèle en nombre entier pour intégrer la priorisation du chauffage des chauffe-eau froid. Finalement, la batterie virtuelle de chauffe-eau pourrait être intégrée dans l'opération d'un réseau électrique intelligent. Cette batterie virtuelle serait alors un joueur parmi un ensemble d'autres ressources énergétiques distribuées.

RÉFÉRENCES

Amat Rodrigo, Joaquin, et Javier Escobar Ortiz. 2022. « skforecast ».

Birge, John R., et François Louveaux. 2011. *Introduction to stochastic programming*. 2nd ed. New York: Springer.

Bynum, Michael L., Gabriel A. Hackebeil, William E. Hart, Carl D. Laird, Bethany L. Nicholson, John D. Siirola, Jean-Paul Watson, et David L. Woodruff. 2021. *Pyomo -- Optimization Modeling in Python*. Cham: Springer.

Canada, Environment and Climate Change. 2022. « A Clean Electricity Standard in Support of a Net-Zero Electricity Sector: Discussion Paper ». Consulté 7 novembre 2022 (https://www.canada.ca/en/environment-climate-change/services/canadian-environmental-protection-act-registry/achieving-net-zero-emissions-electricity-generation-discussion-paper.html).

Ding, Fei, et Weijia Liu. 2022. « Federated Architecture for Secure and Transactive Distributed Energy Resource Management Solutions (FAST-DERMS) ».

Hong, Tao, et Shu Fan. 2016. « Probabilistic Electric Load Forecasting: A Tutorial Review ». *International Journal of Forecasting* 32(3):914-38. doi: 10.1016/j.ijforecast.2015.11.011.

Hyndman, R. J., et G. Athanasopoulos. 2021. *Forecasting: Principles and Practice (3rd ed)*.

Ke, Guolin, Qi Meng, Thomas Finley, Taifeng Wang, Wei Chen, Weidong Ma, Qiwei Ye, et Tie-Yan Liu. 2017. « LightGBM: A Highly Efficient Gradient Boosting Decision Tree ». dans *Advances in Neural Information Processing Systems*. Vol. 30. Curran Associates, Inc.

Khurram, Adil, Roland Malhamé, Luis Duffaut Espinosa, et Mads Almassalkhi. 2020. « Identification of Hot Water End-Use Process of Electric Water Heaters from Energy Measurements ». *Electric Power Systems Research* 189:106625. doi: 10.1016/j.epsr.2020.106625.

King, Alan J., et Stein W. Wallace. 2012. *Modeling with Stochastic Programming*. New York, NY: Springer New York.

Lacroix, Marcel. 1999. « Electric water heater designs for load shifting and control of bacterial contamination ». *Energy Conversion and Management* 40(12):1313-40. doi: 10.1016/S0196-8904(99)00013-8.

Laurent, J. C., et R. P. Malhame. 1994. « A physically-based computer model of aggregate electric water heating loads ». *IEEE Transactions on Power Systems* 9(3):1209-17. doi: 10.1109/59.336078.

Makridakis, Spyros, Rob J. Hyndman, et Fotios Petropoulos. 2020. « Forecasting in Social Settings: The State of the Art ». *International Journal of Forecasting* 36(1):15-28. doi: 10.1016/j.ijforecast.2019.05.011.

Malandra, Filippo, Arman C. Kizilkale, Frédéric Sirois, Brunilde Sansò, Miguel F. Anjos, Michel Bernier, Michel Gendreau, et Roland P. Malhamé. 2020. *Smart distributed energy storage controller (smartDESC)*. G-2020-15. GERAD.

Moreau, Alain. 2011. « Control Strategy for Domestic Water Heaters during Peak Periods and its Impact on the Demand for Electricity ». *Energy Procedia* 12:1074-82. doi: 10.1016/j.egypro.2011.10.140.

Nazir, M. S., F. D. Galiana, et A. Prieur. 2016. « Unit Commitment Incorporating Histogram Control of Electric Loads With Energy Storage ». *IEEE Transactions on Power Systems* 31(4):2857-66. doi: 10.1109/TPWRS.2015.2483542.

Ning, Chao, et Fengqi You. 2019. « Optimization under Uncertainty in the Era of Big Data and Deep Learning: When Machine Learning Meets Mathematical Programming ». *Computers & Chemical Engineering* 125:434-48. doi: 10.1016/j.compchemeng.2019.03.034.

Rolnick, David, Priya L. Donti, Lynn H. Kaack, Kelly Kochanski, Alexandre Lacoste, Kris Sankaran, Andrew Slavin Ross, Nikola Milojevic-Dupont, Natasha Jaques, Anna Waldman-Brown, Alexandra Luccioni, Tegan Maharaj, Evan D. Sherwin, S. Karthik Mukkavilli, Konrad P. Kording, Carla Gomes, Andrew Y. Ng, Demis Hassabis, John C. Platt, Felix Creutzig, Jennifer Chayes, et Yoshua Bengio. 2019. « Tackling Climate Change with Machine Learning ». *arXiv:1906.05433 [cs, stat]*.

Solis, Jérôme. 2015. « Développement d'un estimateur d'état énergétique d'un chauffe-eau pour un contrôle par champ moyen ». masters, École Polytechnique de Montréal.

Tammam, Adham I., Miguel F. Anjos, et Michel Gendreau. 2020. « Balancing Supply and Demand in the Presence of Renewable Generation via Demand Response for Electric Water Heaters ». *Annals of Operations Research* 292(2):753-70. doi: 10.1007/s10479-020-03580-1.

Térache, Célia, Radu Zmeureanu, et François Laurencelle. 2014. « Modélisation sur TRNSYS d'un réservoir d'eau chaude sanitaire et comparaison avec des mesures expérimentales ».

Watson, Jean-Paul, David L. Woodruff, et William E. Hart. 2012. « PySP: Modeling and Solving Stochastic Programs in Python ». *Mathematical Programming Computation* 4(2):109-49. doi: 10.1007/s12532-012-0036-1.

Wong, Steven. 2014. *Canadian Residential Demand Response and Ancillary Service Market Opportunities*. CanmetENERGY technical report, Varennes Research Centre, Natural Resources Canada.

Wong, Steven, Wajid Muneer, Salman Nazir, et Alexandre Prieur. 2013. *Designing, Operating, and Simulating Electric Water Heater Populations for the Smart Grid*. 2013-136 (RP-TEC). CanmetENERGY, Varennes Research Centre, Natural Resources Canada. doi: 10.13140/RG.2.1.1232.9686.

Xu, Zhijie, Ruisheng Diao, Shuai Lu, Jianming Lian, et Yu Zhang. 2014. « Modeling of Electric Water Heaters for Demand Response: A Baseline PDE Model ». *IEEE Transactions on Smart Grid* 5(5):2203-10. doi: 10.1109/TSG.2014.2317149.