

Stochastic optimal control of a domestic microgrid equipped with solar panel and battery

François Pacaud, Pierre Carpentier, Jean-Philippe Chancelier, Michel de Lara

▶ **To cite this version:**

François Pacaud, Pierre Carpentier, Jean-Philippe Chancelier, Michel de Lara. Stochastic optimal control of a domestic microgrid equipped with solar panel and battery. 2018. hal-01688666

HAL Id: hal-01688666
https://enpc.hal.science/hal-01688666

Preprint submitted on 19 Jan 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Stochastic optimal control of a domestic microgrid equipped with solar panel and battery

François Pacaud*, Pierre Carpentier†, Jean-Philippe Chancelier‡, Michel De Lara‡
* Efficacity — † ENSTA — ‡ CERMICS-ENPC

***Abstract*—Microgrids are integrated systems that gather and operate energy production units to satisfy consumers demands. This paper details different mathematical methods to design the Energy Management System (EMS) of domestic microgrids. We consider different stocks coupled together — a battery, a domestic hot water tank — and decentralized energy production with solar panel. The main challenge of the EMS is to ensure, at least cost, that supply matches demand for all time, while considering the inherent uncertainties of such systems. We benchmark two optimization algorithms to manage the EMS, and compare them with a heuristic. The Model Predictive Control (MPC) is a well known algorithm which models the future uncertainties with a deterministic forecast. By contrast, Stochastic Dual Dynamic Programming (SDDP) models the future uncertainties as probability distributions to compute optimal policies. We present a fair comparison of these two algorithms to control microgrid. A comprehensive numerical study shows that i) optimization algorithms achieve significant gains compared to the heuristic, ii) SDDP outperforms MPC by a few percents, with a reasonable computational overhead.**

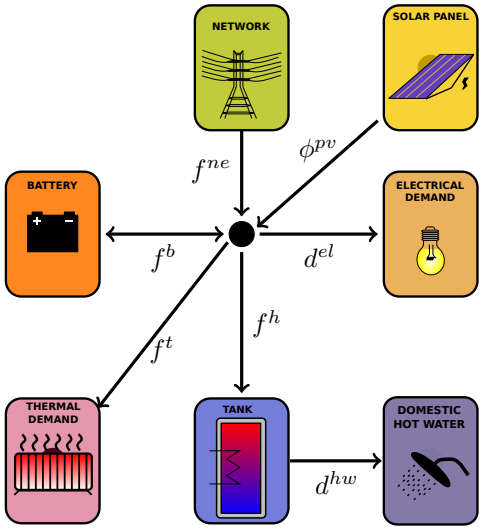


Figure 1. Electrical microgrid

I. Introduction

A. Context

A microgrid is a local energy network that produces part of its energy and controls its own demand. Such systems are complex to control, because of the different stocks and interconnections. Furthermore, at local scale, electrical demands and weather conditions (heat demand and renewable energy production) are highly variable and hard to predict; their stochastic nature adds uncertainty to the system.

We consider here a domestic microgrid (see Figure 1), equipped with a battery, an electrical hot water tank and a solar panel. We use the battery to store energy when prices are low or when the production of the solar panel is above the electrical demand. The microgrid is connected to an external grid to import electricity when needed. Furthermore, we model the building's envelope to take advantage of the thermal inertia of the building. Hence, the system has four stocks to store energy: a battery, a hot water tank, and two passive stocks being the building's walls and inner rooms. Two kind of uncertainties affect the system. First, the electrical and domestic hot water demands are not known in advance. Second, the production of the solar panel is heavily perturbed because of the varying nebulosity affecting their production.

We aim to compare two classes of algorithms to tackle the uncertainty in microgrid Energy Management Systems (EMS). The renowned Model Predictive Control (MPC) algorithm views the future uncertainties with a deterministic forecast. Then, MPC relies on deterministic optimization algorithms to compute optimal decisions. The contender — Stochastic Dual Dynamic Programming (SDDP) — is an algorithm based on the Dynamic Programming principle. Such algorithm computes offline a set of value functions by backward induction; optimal decisions are computed online as time goes on, using the value functions. We present a balanced comparison of these two algorithms, and highlight the advantages and drawbacks of both methods.

B. Litterature

1) Optimization and energy management systems: Energy Management Systems (EMS) are integrated automated tools used to monitor and control energy systems. The design of EMS for buildings has raised interest in recent years. In [1], the authors give an overview concerning the application of optimization methods in designing EMS.

The well-known Model Predictive Control (MPC) [2] has been widely used to control EMS. We refer notably to [3], [4], [5] for applications of MPC in buildings. Different solutions are investigated to tackle uncertainties, such as Stochastic MPC [3] or robust optimization [6].

2) Stochastic Optimization: as we said, at local scale, electrical demand and production are highly variable, especially as microgrids are expected to absorb renewable energies. This leads to pay attention to stochastic optimization approaches. Apart from microgrid management, stochastic optimization

has found some applications in energy systems (see [7] for an overview). Historically, stochastic optimization has been widely applied to hydrovalleys management [8]. Other applications have arisen recently, such as integration of wind energy and storage [9] or insulated microgrids management [10].

Stochastic Dynamic Programming (SDP) [11] is a general method to solve stochastic optimal control problems. In energy applications, a variant of SDP, Stochastic Dual Dynamic Programming (SDDP), has proved its adequacy for large scale applications. SDDP was first described in the seminal paper [8]. We refer to [12] for a generic description of the algorithm and its application to the management of hydrovalleys. A proof of convergence in the linear case is given in [13], and in the convex case in [14].

With the growing adoption of stochastic optimization methods, new researches aim to compare algorithms such as SDP and SDDP with MPC. We refer to the recent paper [15].

C. Structure of the paper

We detail a modelling of the microgrid in Sect. II, then formulate an optimization problem in Sect. III. We outline the different optimization algorithms in Sect. IV. Finally, we provide in Sect. V numerical results concerning the management of the microgrid.

II. Energy system model

In this section, we depict the physical equations of the energy system model described in Figure 1. These equations write naturally in continuous time t. We model the battery and the hot water tank with stock dynamics, and describe the dynamics of the building's temperatures with an electrical analogy. Such physical model fulfills two purposes: it will be used to assess different control policies; it will be the basis of a discrete time model used to design optimal control policies.

A. Load balance

Based on Figure 1, the *load balance* equation of the microgrid writes, at each time t:

$$\phi^{pv}(t) + f^{ne}(t) = f^b(t) + f^t(t) + f^h(t) + d^{el}(t) \ . \tag{1}$$

We now comment the different terms. In the left hand side of Equation (1), the load produced consists of

- the production of the solar panel $\phi^{pv}(t)$,
- the importation from the network $f^{ne}(t)$, supposed non-negative (we do not export electricity to the network).

In the right hand side of Equation (1), the electrical demand is the sum of

- the power exchanged with the battery $f^b(t)$,
- the power injected in the electrical heater $f^t(t)$,
- the power injected in the electrical hot water tank $f^h(t)$,
- the inflexible demands (lightning, cooking...), aggregated in a single demand $d^{el}(t)$.

B. Energy storage

We consider a lithium-ion battery, whose state of charge at time t is denoted by $b(t)$. The state of charge is bounded:

$$\underline{b} \leq b(t) \leq \overline{b} \ . \tag{2}$$

Usually, we set $\underline{b} = 30\% \times \overline{b}$ so as to avoid empty state of charge, which proves to be stressful for the battery. The battery dynamics is given by the differential equation

$$\frac{db}{dt} = \rho_c (f^b(t))^+ - \frac{1}{\rho_d}(f^b(t))^- \ , \tag{3}$$

with ρ_c and ρ_d being the charge and discharge efficiency and $f^b(t)$ denoting the power exchange with the battery. We use the convention $f^+ = \max(0, f)$ and $f^- = \max(0, -f)$.

As we cannot withdraw an infinite power from the battery at time t, we bound the power exchanged with the battery:

$$-\underline{f}^b \leq f^b(t) \leq \overline{f}^b \ . \tag{4}$$

C. Electrical hot water tank

We use a simple linear model for the electrical hot water tank dynamics. At time t, we denote by $T^h(t)$ the temperature inside the hot water tank. We suppose that this temperature is homogeneous, that is, that no stratification occurs inside the tank.

At time t, we define the energy $h(t)$ stored inside the tank as the difference between the tank's temperature $T^h(t)$ and a reference temperature T^{ref}

$$h(t) = \rho V_h c_p \big(T^h(t) - T^{ref}\big) \ , \tag{5}$$

where V_h is the tank's volume, c_p the calorific capacity of water and ρ the density of water. The energy $h(t)$ is bounded:

$$0 \leq h(t) \leq \overline{h} \ . \tag{6}$$

The enthalpy balance equation writes

$$\frac{dh}{dt} = \beta_h f^h(t) - d^{hw}(t) \ , \tag{7}$$

where

- $f^h(t)$ is the electrical power used to heat the tank, satisfying

$$0 \leq f^h(t) \leq \overline{f}^h \ , \tag{8}$$

- $d^{hw}(t)$ is the domestic hot water demand,
- β_h is a conversion yield.

A more accurate representation would model the stratification inside the hot water tank. However, this would greatly increase the number of states in the system, rendering the numerical resolution more cumbersome. We refer to [16] and [17] for discussions about the impact of the tank's modeling on the performance of the control algorithms.

D. Thermal envelope

We model the evolution of the temperatures inside the building with an electrical analogy: we view temperatures as voltages, walls as capacitors, and thermal flows as currents. A model with 6 resistances and 2 capacitors (R6C2) proves to be accurate to describe small buildings [18]. The model takes into account two temperatures:

- the wall's temperature $\theta^w(t)$,
- the inner temperature $\theta^i(t)$.

Their evolution is governed by the two following differential equations

$$c_m \frac{d\theta^w}{dt} = \underbrace{\frac{\theta^i(t) - \theta^w(t)}{R_i + R_s}}_{\substack{\text{Exchange} \\ \text{Indoor/Wall}}} + \underbrace{\frac{\theta^e(t) - \theta^w(t)}{R_m + R_e}}_{\substack{\text{Exchange} \\ \text{Outdoor/Wall}}} + \underbrace{\gamma f^t(t)}_{\text{Heater}} + \underbrace{\frac{R_i}{R_i + R_s}\Phi^{int}(t)}_{\substack{\text{Radiation} \\ \text{through windows}}} + \underbrace{\frac{R_e}{R_e + R_m}\Phi^{ext}(t)}_{\substack{\text{Radiation} \\ \text{through wall}}} , \quad (9a)$$

$$c_i \frac{d\theta^i}{dt} = \underbrace{\frac{\theta^w(t) - \theta^i(t)}{R_i + R_s}}_{\substack{\text{Exchange} \\ \text{Indoor/Wall}}} + \underbrace{\frac{\theta^e(t) - \theta^i(t)}{R_v}}_{\text{Ventilation}} + \underbrace{\frac{\theta^e(t) - \theta^i(t)}{R_f}}_{\text{Windows}} + \underbrace{(1-\gamma) f^t(t)}_{\text{Heater}} + \underbrace{\frac{R_s}{R_i + R_s}\Phi^{int}(t)}_{\substack{\text{Radiation} \\ \text{through windows}}} , \quad (9b)$$

where we denote

- the power injected in the heater by $f^t(t)$,
- the external temperature by $\theta^e(t)$,
- the radiation on the wall by $\Phi^{ext}(t)$,
- the radiation through the windows by $\Phi^{int}(t)$.

The time-varying quantities $\theta^e(t)$, $\Phi^{int}(t)$ and $\Phi^{ext}(t)$ are exogeneous. We denote by $R_i, R_s, R_m, R_e, R_v, R_f$ the different resistances of the R6C2 model, and by c_i, c_m the capacities of the inner rooms and the walls. We denote by γ the proportion of heating dissipated in the wall through conduction, and by $(1-\gamma)$ the proportion of heating dissipated in the inner room through convection.

E. Continuous time state equation

We denote by $x = (b, h, \theta^w, \theta^i)$ the state, $u = (f^b, f^t, f^h)$ the control, and $w = (d^{el}, d^{hw}, \phi^{pv})$ the uncertainties. The continuous state equation writes

$$\dot{x} = F(t, x, u, w) , \quad (10)$$

where the function F is defined by Equations (3)-(7)-(9).

III. Optimization problem statement

Now that we have described the physical model, we turn to the formulation of a decision problem. We aim to compute optimal decisions that minimize a daily operational cost, by solving a stochastic optimization problem.

A. Decisions are taken at discrete times

The EMS takes decisions every 15 minutes to control the system. Thus, we have to provide decisions in discrete time.

We set $\Delta = 15$mn, and we consider an horizon $T_0 = 24$h. We adopt the following convention for discrete processes: for $t \in \{0, 1, \cdots, T = \frac{T_0}{\Delta}\}$, we set $x_t = x(t\Delta)$. That is, x_t denotes the value of the variable x at the beginning of the interval $[t\Delta, (t+1)\Delta[$. Otherwise stated, we will denote by $[t, t+1[$ the continuous time interval $[t\Delta, (t+1)\Delta[$.

B. Modeling uncertainties as random variables

Because of their unpredictable nature, we cannot anticipate the realizations of the electrical and the thermal demands. A similar reasoning applies to the production of the solar panel. We choose to model these quantities as random variables (over a sample space Ω). We adopt the following convention: a random variable will be denoted by an uppercase bold letter $\boldsymbol{Z}$ and its realization will be denoted in lowercase $z = \boldsymbol{Z}(\omega)$. For each $t = 1, \ldots, T$, we define the uncertainty vector

$$\boldsymbol{W}_t = (\mathbf{D}_t^{el}, \mathbf{D}_t^{th}, \mathbf{\Phi}_t^{pv}) , \quad (11)$$

modeled as a random variable. The uncertainty $\boldsymbol{W}_t$ takes value in the set $\mathbb{W}_t = \mathbb{R}^3$.

C. Modeling controls as random variables

As decisions depend on the previous uncertainties, the control is a random variable. We recall that, at each discrete time t, the EMS takes three decisions:

- how much energy to charge/discharge the battery $\boldsymbol{F}_t^b$,
- how much energy to store in the electrical hot water tank $\boldsymbol{F}_t^h$,
- how much energy to inject in the electrical heater $\boldsymbol{F}_t^t$.

We write the decision vector (random variable)

$$\boldsymbol{U}_t = (\boldsymbol{F}_t^b, \boldsymbol{F}_t^h, \boldsymbol{F}_t^t) , \quad (12)$$

taking values in $\mathbb{U}_t = \mathbb{R}^3$.

Then, between two discrete time indexes t and $t+1$, the EMS imports an energy $\boldsymbol{F}_{t+1}^{ne}$ from the external network. The EMS must fulfill the load balance equation (1) whatever the demand $\mathbf{D}_{t+1}^{el}$ and the production of the solar panel $\mathbf{\Phi}_{t+1}^{pv}$, unknown at time t. Hence $\boldsymbol{F}_{t+1}^{ne}$ is a recourse decision taken at time $t+1$. The load balance equation (1) now writes as

$$\boldsymbol{F}_{t+1}^{ne} = \boldsymbol{F}_t^b + \boldsymbol{F}_t^t + \boldsymbol{F}_t^h + \mathbf{D}_{t+1}^{el} - \mathbf{\Phi}_{t+1}^{pv} \quad \mathbb{P} - \text{a.s.} , \quad (13)$$

where $\mathbb{P} - $ a.s. indicates that the constraint is fulfilled in the almost sure sense. Later, we will aggregate the solar panel production $\mathbf{\Phi}_{t+1}^{pv}$ with the demands $\mathbf{D}_{t+1}^{el}$ in Equation (13), as these two quantities appear only by their sum.

The need of a recourse variable is a consequence of stochasticity in the supply-demand equation. The choice of the recourse variable depends on the modeling. Here, we choose the recourse $\boldsymbol{F}_{t+1}^{ne}$ to be provided by the external network, that is, the external network mitigates the uncertainties in the system.

D. States and dynamics

The state becomes also a random variable

$$\boldsymbol{X}_t = (\boldsymbol{B}_t, \boldsymbol{H}_t, \boldsymbol{\theta}_t^w, \boldsymbol{\theta}_t^i) \ . \tag{14}$$

It gathers the stocks in the battery $\boldsymbol{B}_t$ and in the electrical hot water tank $\boldsymbol{H}_t$, plus the two temperatures of the thermal envelope $(\boldsymbol{\theta}_t^i, \boldsymbol{\theta}_t^w)$. Thus, the state vector $\boldsymbol{X}_t$ takes values in $\mathbb{X}_t = \mathbb{R}^4$.

The discrete dynamics writes

$$x_{t+1} = f_t(x_t, u_t, w_{t+1}) \ , \tag{15}$$

where f_t corresponds to the discretization of the continuous dynamics (10) using a forward Euler scheme, that is, $x_{t+1} = x_t + \Delta \times F(t, x_t, u_t, w_{t+1})$. By doing so, we suppose that the control u_t and the uncertainty w_{t+1} are constant over the interval $[t, t+\Delta[$.

The dynamics (15) rewrites as an almost-sure constraint:

$$\boldsymbol{X}_{t+1} = f_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\big) \qquad \mathbb{P}-\text{a.s.} \ . \tag{16}$$

We suppose that we start from a given position x_0, thus adding a new initial constraint: $\boldsymbol{X}_0 = x_0$.

E. Non-anticipativity constraints

The future realizations of uncertainties are unpredictable. Thus, decisions are functions of previous history only, that is, the information collected between time 0 and time t. Such a constraint is encoded as an algebraic constraint, using the tools of Probability theory [19]. The so-called non-anticipativity constraint writes

$$\sigma(\boldsymbol{U}_t) \subset \mathcal{F}_t \ , \tag{17}$$

where $\sigma(\boldsymbol{U}_t)$ is the σ-algebra generated by $\boldsymbol{U}_t$ and $\mathcal{F}_t = \sigma(\boldsymbol{W}_1, \cdots, \boldsymbol{W}_t)$ the σ-algebra associated to the previous history $(\boldsymbol{W}_1, \ldots, \boldsymbol{W}_t)$. If Constraint (17) holds true, the Doob lemma [19] ensures that there exists a function π_t such that

$$\boldsymbol{U}_t = \pi_t(\boldsymbol{X}_0, \boldsymbol{W}_1, \ldots, \boldsymbol{W}_t) \ . \tag{18}$$

This is how we turn an (abstract) algebraic constraint into a more practical functional constraint. The function π_t is an example of policy.

F. Bounds constraints

By Equations (2) and (6), the stocks in the battery $\boldsymbol{B}_t$ and in the tank $\boldsymbol{H}_t$ are bounded. At time t, the control $\boldsymbol{F}_t^b$ must ensure that the next state $\boldsymbol{B}_{t+1}$ is admissible, that is, $\underline{b} \leq \boldsymbol{B}_{t+1} \leq \overline{b}$ by Equation (2), which rewrites,

$$\underline{b} \leq \boldsymbol{B}_t + \Delta\big[\rho_c(\boldsymbol{F}_t^b)^+ + \frac{1}{\rho_d}(\boldsymbol{F}_t^b)^-\big] \leq \overline{b} \ . \tag{19}$$

Thus, the constraints on $\boldsymbol{F}_t^b$ depends on the stock $\boldsymbol{B}_t$. The same reasoning applies for the tank power $\boldsymbol{F}_t^h$. Furthermore, we set bound constraints on controls, that is,

$$-\overline{f}^b \leq \boldsymbol{F}_t^b \leq \overline{f}^b \ , \ \ 0 \leq \boldsymbol{F}_t^h \leq \overline{f}_t^h \ , \ \ 0 \leq \boldsymbol{F}_t^t \leq \overline{f}_t^t \ . \tag{20}$$

Finally the load-balance equation (13) also acts as a constraint on the controls. We gather all these constraints in an admissible set on control $\boldsymbol{U}_t$ depending on the current state $\boldsymbol{X}_t$:

$$\boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t) \qquad \mathbb{P}-\text{a.s.} \ . \tag{21}$$

G. Objective

At time t, the operational cost $L_t : \mathbb{X}_t \times \mathbb{U}_t \times \mathbb{W}_{t+1} \to \mathbb{R}$ aggregates two different costs:

$$L_t(x_t, u_t, w_{t+1}) = p_t^e \times f_{t+1}^{ne} + p_t^d \times \max(0, \overline{\theta_t^i} - \theta_t^i) \ . \tag{22}$$

First, we pay a price p_t^e to import electricity from the network between time t and $t+1$. Hence, electricity cost is equal to $p_t^e \times \boldsymbol{F}_{t+1}^{ne}$. Second, if the indoor temperature is below a given threshold, we penalize the induced discomfort with a cost $p_t^d \times \max(0, \overline{\theta_t^i} - \boldsymbol{\theta}_t^i)$, where p_t^d is a virtual price of discomfort. The cost L_t is a convex piecewise linear function, which will prove important for the SDDP algorithm.

We add a final cost $K : \mathbb{X}_T \to \mathbb{R}$ to ensure that stocks are non empty at final time T

$$K(x_T) = \kappa \times \max(0, x_0 - x_T) \ , \tag{23}$$

where κ is a positive penalization coefficient calibrated by trials and errors.

As decisions $\boldsymbol{U}_t$ and states $\boldsymbol{X}_t$ are random, the costs $L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1})$ become also random variables. We choose to minimize the expected value of the daily operational cost, yielding the criterion

$$\mathbb{E}\Big[\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T)\Big] \ , \tag{24}$$

yielding an expected value of a convex piecewise linear cost.

H. Stochastic optimal control formulation

Finally, the EMS problem translates to a generic Stochastic Optimal Control (SOC) problem

$$\min_{\boldsymbol{X},\boldsymbol{U}} \ \mathbb{E}\Big[\sum_{t=0}^{T-1} L_t(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}) + K(\boldsymbol{X}_T)\Big] \ , \tag{25a}$$

$$\boldsymbol{X}_0 = x_0 \ , \tag{25b}$$

$$\boldsymbol{X}_{t+1} = f_t\big(\boldsymbol{X}_t, \boldsymbol{U}_t, \boldsymbol{W}_{t+1}\big) \quad \mathbb{P}-\text{a.s.} \ , \tag{25c}$$

$$\boldsymbol{U}_t \in \mathcal{U}_t^{ad}(\boldsymbol{X}_t) \quad \mathbb{P}-\text{a.s.} \ , \tag{25d}$$

$$\sigma(\boldsymbol{U}_t) \subset \mathcal{F}_t \ . \tag{25e}$$

Problem (25) states that we want to minimize the expected value of the costs while satisfying the dynamics, the control bounds and the non-anticipativity constraints.

IV. Resolution methods

The exact resolution of Problem (25) is out of reach in general. We propose two different algorithms that provide policies $\pi_t : \mathbb{X}_0 \times \mathbb{W}_1 \times \cdots \times \mathbb{W}_t \to \mathbb{U}_t$ that map available information $x_0, w_1, \ldots, w_t$ at time t to a decision u_t.

A. Model Predictive Control (MPC)

MPC is a classical algorithm commonly used to handle uncertainties in energy systems. At time t, it considers a

deterministic forecast $(\overline{w}_{t+1},\dots,\overline{w}_T)$ of the future uncertainties $(\boldsymbol{W}_{t+1},\dots,\boldsymbol{W}_T)$ and solves the deterministic problem

$$\min_{(u_t,\cdots,u_{T-1})} \sum_{j=t}^{T-1} \Big[L_j(x_j,u_j,\overline{w}_{j+1})\Big] + K(x_T)\ , \tag{26a}$$

$$x_{j+1} = f_j\big(x_j,u_j,\overline{w}_{j+1}\big)\ , \tag{26b}$$

$$u_j \in \mathcal{U}_j^{ad}(x_j)\ . \tag{26c}$$

At time t, we solve Problem (26), retrieve the optimal decisions $(u_t^\sharp,\dots,u_{T-1}^\sharp)$ and only keep the first decision $u_t^\sharp$ to control the system between time t and $t+1$. Then, we restart the procedure at time $t+1$.

As Problem (26) is linear and the number of time steps remains not too large, we are able to solve it exactly for every t.

B. Stochastic Dual Dynamic Programming (SDDP)

1) Dynamic Programming and Bellman principle: the Dynamic Programming method [20] looks for solutions of Problem (25) as state-feedbacks $\pi_t : \mathbb{X}_t \to \mathbb{U}_t$. Dynamic Programming computes a serie of value functions backward in time by setting $V_T(x_T) = K(x_T)$ and solving the recursive equations

$$V_t(x_t) = \min_{u\in\mathcal{U}_t^{ad}(x_t)} \int_{\mathbb{W}_{t+1}} \Big[L_t(x_t,u,w_{t+1}) + V_{t+1}\big(f_t(x_t,u,w_{t+1})\big)\Big]\mu_{t+1}^{of}(dw_{t+1})\ , \tag{27}$$

where μ_{t+1}^{of} is a (offline) probability distribution on $\mathbb{W}_{t+1}$.

Once these functions are computed, we compute a decision at time t as a state-feedback:

$$\pi_t(x_t) \in \operatorname*{arg\,min}_{u\in\mathcal{U}_t^{ad}(x_t)} \int_{\mathbb{W}_{t+1}} \Big[L_t(x_t,u,w_{t+1}) + V_{t+1}\big(f_t(x_t,u,w_{t+1})\big)\Big]\mu_{t+1}^{on}(dw_{t+1})\ , \tag{28}$$

where μ_{t+1}^{on} is an online probability distribution on $\mathbb{W}_{t+1}$. This method proves to be optimal when the uncertainties $\boldsymbol{W}_1,\dots,\boldsymbol{W}_T$ are stagewise independent and when $\mu_t^{on} = \mu_t^{of}$ is the probability distribution of $\boldsymbol{W}_t$ in (27).

2) Description of Stochastic Dual Dynamic Programming: Dynamic Programming suffers from the well-known *curse of dimensionality* [11]: its numerical resolution fails for state dimension typically greater than 4 when value functions are computed on a discretized grid. As the state $\boldsymbol{X}_t$ in §III-D has 4 dimensions, SDP would be too slow to solve numerically Problem (25). The Stochastic Dual Dynamic Programming (SDDP) can bypass the curse of dimensionality by approximating value functions by polyhedral functions. It is optimal for solving Problem (25) when uncertainties are stagewise independent, costs L_t and K are convex and dynamics f_t are linear [14].

SDDP provides an outer approximation $\underline{V}_t^k$ of the value function V_t in (27) with a set of supporting hyperplanes $\{(\lambda_t^j,\beta_t^j)\}_{j=1,\cdots,k}$ by

$$\underline{V}_t(x_t) = \min_{\theta_t\in\mathbb{R}} \theta_t\ , \tag{29a}$$

$$\big\langle \lambda_t^j\ , x_t\big\rangle + \beta_t^j \le \theta_t\ ,\ \ \forall j = 1,\cdots,k\ . \tag{29b}$$

Each iteration k of SDDP encompasses two passes.

- During the *forward pass*, we draw a scenario $x_0,\dots,w_T^k$ of uncertainties, and compute a state trajectory $\big\{x_t^k\big\}_{t=0\cdots T}$ along this scenario. Starting from position x_0, we compute x_{t+1}^k in an iterative fashion: i) we compute the optimal control at time t using the available $\underline{V}_{t+1}^k$ function

$$u_t^k \in \operatorname*{arg\,min}_{u\in\mathcal{U}_t^{ad}(x_t)} \int_{\mathbb{W}_{t+1}} \big[L_t(x_t^k,u,w_{t+1}) + \underline{V}_{t+1}^k\big(f_t(x_t^k,u,w_{t+1})\big)\big]\mu_{t+1}^{of}(dw_{t+1})\ , \tag{30}$$

  and ii), we set $x_{t+1}^k = f_t(x_t^k,u_t^k,w_{t+1}^k)$ where f_t is given by (15).
- During the *backward pass*, we update the approximated value functions $\big\{\underline{V}_t^k\big\}_{t=0,\cdots,T}$ backward in time along the trajectory $\big\{x_t^k\big\}_{t=0,\cdots,T}$. At time t, we solve the problem

$$\theta_t^{k+1} = \min_{u\in\mathcal{U}_t^{ad}(x_t)} \int_{\mathbb{W}_{t+1}} \big[L_t(x_t^k,u,w_{t+1}) + \underline{V}_{t+1}^{k+1}\big(f_t(x_t^k,u,w_{t+1})\big)\big]\mu_{t+1}^{of}(dw_{t+1})\ , \tag{31}$$

  and we obtain a new cut $(\lambda_t^{k+1},\beta_t^{k+1})$ where λ_t^{k+1} is a subgradient of optimal cost (31) evaluated at point $x_t = x_t^k$ and $\beta_t^{k+1} = \theta_t^{k+1} - \big\langle \lambda_t^{k+1}, x_t^k\big\rangle$. This new cut allows to update the function $\underline{V}_t^{k+1}$: $V_t^{k+1} = \max\{\underline{V}_t^k, \big\langle \lambda_t^{k+1},.\big\rangle + \beta_t^{k+1}\}$.

Otherwise stated, SDDP only explores the state space around "interesting" trajectories (those computed during the forward passes) and refines the value functions only in the corresponding space regions (backward passes).

3) Obtaining online controls with SDDP: in order to compute implementable decisions, we use the following procedure.

- Approximated value functions $\big\{\underline{V}_t\big\}$ are computed with the SDDP algorithm (see §IV-B2). These computations are done offline.
- The approximated value functions $\big\{\underline{V}_t\big\}$ are then used to compute online a decision at any time t for any state x_t.

More precisely, we compute the SDDP policy π_t^{sddp} by

$$\pi_t^{sddp}(x_t) \in \operatorname*{arg\,min}_{u\in\mathcal{U}_t^{ad}(x_t)} \int_{\mathbb{W}_{t+1}} \big[L_t(x_t,u,w_{t+1}) + \underline{V}_{t+1}\big(f_t(x_t,u,w_{t+1})\big)\big]\mu_{t+1}^{on}(dw_{t+1})\ , \tag{32}$$

which corresponds to replacing the value function V_{t+1} in Equation (28) with its approximation $\underline{V}_{t+1}$. The decision $\pi_t^{sddp}(x_t)$ is used to control the system between time t and $t+1$. Then, we resolve Problem (32) at time $t+1$.

To solve numerically problems (30)-(31)-(32) at time t, we will consider distributions with finite support $w_t^1,\dots,w_t^S$. The offline distribution μ_{t+1}^{of} now writes: $\mu_{t+1}^{of} = \sum_{s=1}^S p_s\delta_{w_{t+1}^s}$ where $\delta_{w_{t+1}^s}$ is the Dirac measure at w_{t+1}^s and $(p_1,\dots,p_S)$ are probability weights. The same reasoning applies to the online distribution μ_{t+1}^{on}. For instance, Problem (32) reformulates as

$$\pi_t^{sddp}(x_t) \in \operatorname*{arg\,min}_{u\in\mathcal{U}_t^{ad}(x_t)} \sum_{s=1}^S p_s\big[L_t(x_t,u,w_{t+1}^s) + \underline{V}_{t+1}\big(f_t(x_t,u,w_{t+1}^s)\big)\big]\ . \tag{33}$$

V. Numerical results

A. Case study

1) Settings: we aim to solve the stochastic optimization problem (25) over one day, with 96 time steps. The battery's size is 3 kWh, and the hot water tank has a capacity of 120 l. We suppose that the house has a surface $A_p = 20$ m^2 of solar panel at disposal, oriented south, and with a yield of 15%. We penalize the recourse variable $\mathbf{F}_{t+1}^{ne}$ in (22) with on-peak and off-peak tariff, corresponding to Électricité de France's (EDF) individual tariffs. The building's thermal envelope corresponds to the French RT2012 specifications [21]. Meteorological data comes from Meteo France measurements corresponding to the year 2015.

2) Demands scenarios: we have scenarios of electrical and domestic hot water demands at 15 minutes time steps, obtained with StRoBe [22]. Figure 2 displays 100 scenarios of electrical and hot water demands. We observe that the shape of these scenarios is consistent: demands are almost null during night, and there are peaks around midday and 8 pm. Peaks in hot water demands corresponds to showers. We aggregate the production of the solar panel $\mathbf{\Phi}^{pv}$ and the electrical demands $\mathbf{D}^{el}$ in a single variable $\mathbf{D}^{el}$ to consider only two uncertainties $(\mathbf{D}_t^{el}, \boldsymbol{D}_t^{hw})$.

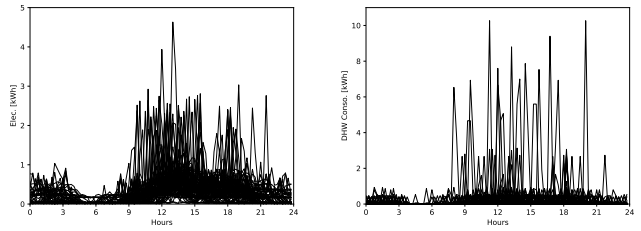


Figure 2. Electrical (left) and domestic hot water (right) demand scenarios.

3) Out of sample assessment of strategies: to obtain a fair comparison between SDDP and MPC, we use an out-of-sample validation. We generate 2,000 scenarios of electrical and hot water demands, and we split these scenarios in two distinct parts: the first N_{opt} scenarios are called *optimization scenarios*, and the remaining N_{sim} scenarios are called *assessment scenarios*. We made the choise $N_{opt} = N_{sim} = 1,000$.

First, during the offline phase, we use the optimization scenarios to build models for the uncertainties, under the mathematical form required by each algorithm (see Sect. IV). Second, during the online phase, we use the assessment scenarios to compare the strategies produced by these algorithms. At time t during the assessment, the algorithms cannot use the future values of the assessment scenarios, but can take advantage of the observed values up to t to update their statistical models of future uncertainties.

B. Numerical implementation

1) Implementing the algorithms: we implement MPC and SDDP in Julia 0.6, using JuMP [23] as a modeler, `StochDynamicProgramming.jl` as a SDDP solver, and Gurobi 7.02 [24] as a LP solver. All computations run on a Core i7 2.5 GHz processor, with 16Go RAM.

2) MPC procedure: Electrical and thermal demands are naturally correlated in time [25]. To take into account such a dependence across the different time-steps, we chose to model the process $\boldsymbol{W}_1, \ldots, \boldsymbol{W}_T$ with an auto-regressive (AR) process.

a) Building offline an AR model for MPC: we fit an AR(1) model upon the optimization scenarios (we do not consider higher order lag for the sake of simplicity). For $i \in \{el, hw\}$, the AR model writes

$$d_{t+1}^i = \alpha_t^i d_t^i + \beta_t^i + \varepsilon_t^i \,, \tag{34a}$$

where the non-stationary coefficients (α_t^i, β_t^i) are, for all time t, solutions of the least-square problem

$$(\alpha_t^i, \beta_t^i) = \arg\min_{a,b} \sum_{s=1}^{N_{opt}} \left\| d_{t+1}^{i,s} - a d_t^{i,s} - b \right\|_2^2 \,. \tag{34b}$$

The points $(d_t^{i,1}, \ldots d_t^{i,N_{opt}})$ correspond to the optimization scenarios. The AR residuals $(\varepsilon_t^{el}, \varepsilon_t^{hw})$ are a white noise process.

b) Updating the forecast online: once the AR model is calibrated, we use it to update the forecast during assessment (see §IV-A). The update procedure is threefold:

i) we observe the demands $w_t = (d_t^{el}, d_t^{hw})$ between time $t-1$ and t,

ii) we update the forecast $\overline{w}_{t+1}$ at time $t+1$ with the AR model

$$\overline{w}_{t+1} = \left(\overline{d}_{t+1}^{el}, \overline{d}_{t+1}^{hw}\right) = \left(\alpha_t^{el} d_t^{el} + \beta_t^{el},\ \alpha_t^{hw} d_t^{hw} + \beta_t^{hw}\right) \,,$$

iii) we set the forecast between time $t+2$ and T by using the mean values of the optimization scenarios:

$$\overline{w}_\tau = \frac{1}{N^{opt}} \sum_{i=1}^{N^{opt}} w_\tau^i \quad \forall \tau = t+2, \cdots, T \,.$$

Once the forecast $(\overline{w}_{t+1}, \ldots, \overline{w}_T)$ is available, it is fed into the MPC algorithm that solves Problem (26).

3) SDDP procedure: even if electrical and thermal demands are naturally correlated in time [25], the SDDP algorithm only relies upon marginal distributions.

a) Building offline probability distributions for SDDP: rather than fitting an AR model like done with MPC, we use the optimization scenarios to build marginal probability distributions μ_t^{of} that will feed the SDDP procedure in (30)-(31).

We cannot directly consider the discrete empirical marginal probability derived from all N_{opt} scenarios, because the support size would be too large for SDDP. This is why we use optimal quantization to map the N_{opt} optimization scenarios to S representative points. We use a Loyd-Max quantization scheme [26] to obtain a discrete probability distribution: at each time t, we use the N_{opt} optimization scenarios to build a partition $\Xi = (\Xi_1, \cdots, \Xi_S)$, where Ξ is the solution of the optimal quantization problem

$$\min_{\Xi} \quad \sum_{s=1}^{S} \left(\sum_{w_t^i \in \Xi_s} \left\| w_t^i - \widetilde{w}_t^s \right\|_2^2 \right) \tag{35}$$

where $\widetilde{w}_t^s = \frac{1}{card(\Xi_s)} \sum_{w_t^i \in \Xi_s} w_t^i$ is the so-called centroid of Ξ_s. Then, we set for all time $t = 0, \cdots, T$ the discrete offline distributions $\mu_t^{of} = \sum_{s=1}^{S} p_s \delta_{\widetilde{w}_t^s}$, where $\delta_{\widetilde{w}_t^s}$ is the Dirac measure at point $\widetilde{w}_t^s$ and $p_s = card(\Xi_s)/N^{opt}$ is the associated probability weight. We have chosen $S = 20$ to have enough precision.

b) Computing value functions offline: then, we use these probability distributions as an input to compute a set of value functions with the procedure described in §IV-B2.

c) Using the value functions online: once the value functions have been computed by SDDP, we are able to compute online decisions with Equation (33)[1]. SDDP, on the contrary of MPC, does not update the online probability distribution μ_t^{on} during assessment to consider the information brought by the previous observations.

4) Heuristic procedure: we choose to compare the MPC and SDDP algorithms with a basic decision rule. This heuristic is as follows: the battery is charged whenever the solar production $\mathbf{\Phi}^{pv}$ is available, and discharged to fulfill the demand if there remains enough energy in the battery; the tank is charged ($\boldsymbol{F}_t^h > 0$) if the tank energy $\boldsymbol{H}_t$ is lower than $\boldsymbol{H}_0$, the heater $\boldsymbol{F}_t^t$ is switched on when the temperature is below the setpoint $\overline{\theta_t^i}$ and switched off whenever the temperature is above the setpoint plus a given margin.

C. Results

1) Assessing on different meteorological conditions: we assess the algorithms on three different days, with different meteorological conditions (see Table I). Therefore, we use three distinct sets of N_{sim} assessment scenarios of demands, one for each typical day.

	Date	Temp. ($^\circ C$)	PV Production (kWh)
Winter Day	February, 19th	3.3	8.4
Spring Day	April, 1st	10.1	14.8
Summer Day	May, 31st	14.1	23.3

Table I
DIFFERENT METEOROLOGICAL CONDITIONS

These three different days corresponds to different heating needs. During *Winter day*, the heating is maximal, whereas it is medium during *Spring day* and null during *Summer day*. The production of the solar panel varies accordingly.

2) Comparing the algorithms performances: during assessment, we use MPC (see (26)) and SDDP (see (32)) strategies to compute online decisions along N_{sim} assessment scenarios. Then, we compare the average electricity bill obtained with these two strategies and with the heuristic. The assessment results are given in Table II: means and standard deviation σ are computed by Monte Carlo with the N_{sim} assessment scenarios; the notation $\pm$ corresponds to the interval $\pm 1.96 \frac{\sigma}{\sqrt{N_{sim}}}$, which is a 95% confidence interval.

We observe that MPC and SDDP exhibit close performance, and make better than the heuristic. If we consider mean electricity bills, SDDP achieves better savings than MPC during *Summer day* and *Winter day*, but SDDP and MPC display similar performances during *Spring day*.

[1] In practice, the quantization size of μ_t^{on} is bigger than those of μ_t^{of}, to have a greater accuracy online

	SDDP	MPC	Heuristic
Offline time	50 s	0 s	0 s
Online time	1.5 ms	0.5 ms	0.005 ms
Electricity bill (€)			
Winter day	4.38 $\pm$ 0.02	4.59 $\pm$ 0.02	5.55 $\pm$ 0.02
Spring day	1.46 $\pm$ 0.01	1.45 $\pm$ 0.01	2.83 $\pm$ 0.01
Summer day	0.10 $\pm$ 0.01	0.18 $\pm$ 0.01	0.33 $\pm$ 0.02

Table II
COMPARISON OF MPC, SDDP AND HEURISTIC STRATEGIES

In addition, SDDP achieves better savings than MPC for the vast majority of scenarios. Indeed, if we compare the difference between the electricity bills scenario by scenario, we observe that SDDP is better than MPC for about 93% of the scenarios. This can be seen on Figure 3 that displays the histogram of the absolute gap savings between SDDP and MPC during *Summer day*. The distribution of the gap exhibits a heavy tail that favors SDDP on extreme scenarios. Similar analyses hold for *Winter* and *Spring day*. Thus, we claim that SDDP outperforms MPC for the electricity savings. Concerning the performance on thermal comfort, temperature trajectories are above the temperature setpoints specified in §III-G for both MPC and SDDP.

In term of numerical performance, it takes less than a minute to compute a set of cuts as in §IV-B2 with SDDP on a particular day. Then, the online computation of a single decision takes 1.5 ms on average, compared to 0.5 ms for MPC. Indeed, MPC is favored by the linearity of the optimization Problem (26), whereas, for SDDP, the higher the quantization size S, the slowest is the resolution of Problem (32), but the more information the online probability distribution μ_t^{on} carries.

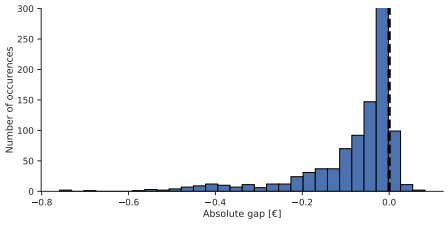


Figure 3. Absolute gap savings between MPC and SDDP during *Summer day*

3) Analyzing the trajectories: we analyze now the trajectories of stocks in assessment, during *Summer day*. The heating is off, and the production of the solar panel is nominal at midday.

Figure 4 displays the state of charge of the battery along a subset of assessment scenarios, for SDDP and MPC. We observe that SDDP charges earlier the battery at its maximum. On the contrary MPC charges the battery later, and does not use the full potential of the battery. The two algorithms discharge the battery to fulfill the evening demands. We notice that each trajectory exhibits a single cycle of charge/discharge, thus decreasing battery's aging.

Figure 5 displays the charge of the domestic hot water tank along the same subset of assessment scenarios. We observe a similar behavior as for the battery trajectories: SDDP uses more the electrical hot water tank to store the excess of PV energy, and the level of the tank is greater at the end of the day than in MPC.

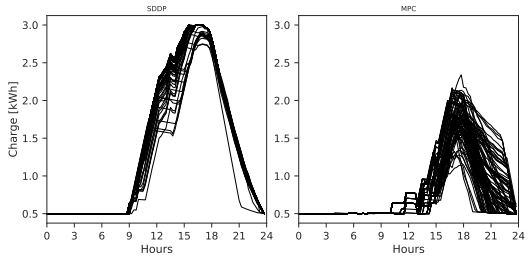


Figure 4. Battery charge trajectories for SDDP and MPC during *Summer day*

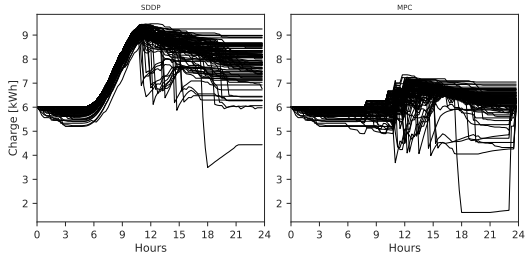


Figure 5. Hot water tank trajectories for SDDP and MPC during *Summer day*

This analysis suggests that SDDP makes a better use of storage capacities than MPC.

VI. Conclusion

We have presented a domestic microgrid energy system, and compared different optimization algorithms to control the stocks with an Energy Management System.

The results show that optimization based strategies outperform the proposed heuristic procedure in term of money savings. Furthermore, SDDP outperforms MPC during *Winter* and *Summer day* — achieving up to 35% costs savings — and displays similar performance as MPC during *Spring day*. Even if SDDP and MPC exhibit close performance, a comparison scenario by scenario shows that SDDP beats MPC most of the time (more than 90% of scenarios during *Summer day*). Thus, we claim that SDDP is better than MPC to manage uncertainties in such a microgrid, although MPC gives also good performance. SDDP also makes a better use of storage capacities.

Our study can be extended in different directions. First, we could mix SDDP and MPC to recover the benefits of these two algorithms. Indeed, SDDP is designed to handles the uncertainties' variability but fails to capture the time correlation, whereas MPC ignores the uncertainties' variability, but considers time correlation by means of a multistage optimization problem. Second, we are currently investigating the optimization of larger scale microgrids — with different interconnected buildings — by decomposition methods.

References

[1] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi, C. Canizares, R. Iravani, M. Kazerani, A. H. Hajimiragha, O. Gomis-Bellmunt, M. Saeedifard, R. Palma-Behnke, *et al.*, "Trends in microgrid control," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, 2014.

[2] C. E. Garcia, D. M. Prett, and M. Morari, "Model predictive control: theory and practice—a survey," *Automatica*, vol. 25, no. 3, pp. 335–348, 1989.

[3] F. Oldewurtel, *Stochastic model predictive control for energy efficient building climate control*. PhD thesis, Eidgenössische Technische Hochschule ETH Zürich, 2011.

[4] P. Malisani, *Pilotage dynamique de l'énergie du bâtiment par commande optimale sous contraintes utilisant la pénalisation intérieure*. PhD thesis, Ecole Nationale Supérieure des Mines de Paris, 2012.

[5] M. Y. Lamoudi, *Distributed Model Predictive Control for energy management in buildings*. PhD thesis, Université de Grenoble, 2012.

[6] K. Paridari, A. Parisio, H. Sandberg, and K. H. Johansson, "Robust scheduling of smart appliances in active apartments with user behavior uncertainty," *IEEE Transactions on Automation Science and Engineering*, vol. 13, no. 1, pp. 247–259, 2016.

[7] M. De Lara, P. Carpentier, J.-P. Chancelier, and V. Leclère, *Optimization Methods for the Smart Grid.* Conseil Francais de l'Energie, 2014.

[8] M. V. Pereira and L. M. Pinto, "Multi-stage stochastic optimization applied to energy planning," *Mathematical programming*, vol. 52, no. 1-3, pp. 359–375, 1991.

[9] P. Haessig, *Dimensionnement et gestion d'un stockage d'énergie pour l'atténuation des incertitudes de production éolienne*. PhD thesis, Ecole normale supérieure de Cachan, 2014.

[10] B. Heymann, J. F. Bonnans, F. Silva, and G. Jimenez, "A stochastic continuous time model for microgrid energy management," in *European Control Conference (ECC)*, pp. 2084–2089, IEEE, 2016.

[11] D. P. Bertsekas, *Dynamic programming and optimal control*, vol. 1. Athena Scientific Belmont, MA, third ed., 2005.

[12] A. Shapiro, "Analysis of stochastic dual dynamic programming method," *European Journal of Operational Research*, vol. 209, no. 1, pp. 63–72, 2011.

[13] A. B. Philpott and Z. Guan, "On the convergence of stochastic dual dynamic programming and related methods," *Operations Research Letters*, vol. 36, no. 4, pp. 450–455, 2008.

[14] P. Girardeau, V. Leclère, and A. B. Philpott, "On the convergence of decomposition methods for multistage stochastic convex programs," *Mathematics of Operations Research*, vol. 40, no. 1, pp. 130–145, 2014.

[15] A. N. Riseth, J. N. Dewynne, and C. L. Farmer, "A comparison of control strategies applied to a pricing problem in retail," *arXiv preprint arXiv:1710.02044*, 2017.

[16] T. Schütz, R. Streblow, and D. Müller, "A comparison of thermal energy storage models for building energy system optimization," *Energy and Buildings*, vol. 93, pp. 23–31, 2015.

[17] N. Beeker, P. Malisani, and N. Petit, "Discrete-time optimal control of electric hot water tank," in *DYCOPS*, 2016.

[18] T. Berthou, *Development of building models for load curve forecast and design of energy optimization and load shedding strategies*. PhD thesis, Ecole Nationale Supérieure des Mines de Paris, Dec. 2013.

[19] O. Kallenberg, *Foundations of Modern Probability*. Springer-Verlag, New York, second ed., 2002.

[20] R. Bellman, *Dynamic Programming*. New Jersey: Princeton University Press, 1957.

[21] J. Officiel, "Arrêté du 28 décembre 2012 relatif aux caractéristiques thermiques et aux exigences de performance énergétique des bâtiments nouveaux," 2013.

[22] R. Baetens and D. Saelens, "Modelling uncertainty in district energy simulations by stochastic residential occupant behaviour," *Journal of Building Performance Simulation*, vol. 9, no. 4, pp. 431–447, 2016.

[23] I. Dunning, J. Huchette, and M. Lubin, "JuMP: A modeling language for mathematical optimization," *SIAM Review*, vol. 59, no. 2, pp. 295–320, 2017.

[24] G. O. Inc, "Gurobi Optimizer Reference Manual," 2014.

[25] J. Widén and E. Wäckelgård, "A high-resolution stochastic model of domestic activity patterns and electricity demand," *Applied Energy*, vol. 87, no. 6, pp. 1880–1892, 2010.

[26] S. Lloyd, "Least squares quantization in PCM," *IEEE transactions on information theory*, vol. 28, no. 2, pp. 129–137, 1982.