A Comparative Study of Optimization Approaches for Battery EMS in Commercial Buildings

Daniel Mimouni[1,2], Paul Malisani[2], Jiamin Zhu[2], Welington de Oliveira[1]

1 **Centre de Mathématiques Appliquées - CMA, Mines Paris PSL, France**

2 **IFP Energies nouvelles, France**

Projet EMS
Stochastic Optimization

Outline

I. The Energy Management System (EMS) Problem

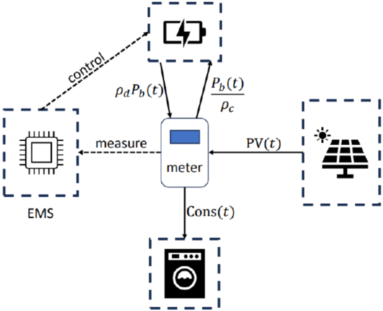
II. Computing solutions for different models

a. MPC a deterministic model
b. Scenario-based models
c. Reinforcement Learning

III. Numerical Application

Energy Management Systems: Context

- EMS aim to optimize electricity usage and minimize operational costs.
- Key challenge: Making decisions under uncertainty (consumption & production).

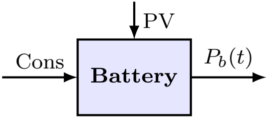


- Traditional models: stochastic programming to minimize expected cost.

Challenge

- The distribution of future data is unknown

Optimization Model for Multistage Energy Management



The effective power demand:

$$\boldsymbol{P_m}(t) = \mathbf{Cons}(t) - \mathbf{PV}(t) + \frac{1}{\rho_c}\max\{P_b(t), 0\} + \rho_d \min\{P_b(t), 0\} \tag{1}$$

Bill

The electricity bill to minimize:

$$J_{t_1:t_2}(P_b, \mathbf{Cons}, \mathbf{PV}) := \int_{t_1}^{t_2} p_r^c(t)\max\{\boldsymbol{P_m}(t), 0\} + p_r^d(t)\min\{\boldsymbol{P_m}(t), 0\}\mathrm{d}t \tag{2}$$

The battery's dynamics are governed by:

$$\dot{E}(t) = P_b(t), \tag{3}$$

and the operational constraints of the battery's charging/discharging process are

$$0 < E < 13 \text{ kWh}, \tag{4}$$

$$-8 < P_b < 8\rho_c \text{ kW} \tag{5}$$

$$E(t_1) = E(t_2) = E^0. \tag{6}$$

Model Predictive Control (MPC)

MPC solves a deterministic optimization problem

- Currently used

MPC

One of the most used model.

- MPC predicts a realization ($\widehat{\text{Cons}}$ and $\widehat{\text{PV}}$) of the random variables **Cons** and **PV**

Solves:

$$\min_{P_b \in \mathrm{L}^2} J_{t:t+\Delta T}(P_b, \widehat{\text{Cons}}, \widehat{\text{PV}}) \tag{MPC}$$

- Easy to implement
- Solution can be far from optimal depending on the accuracy of $\widehat{\text{Cons}}$ and $\widehat{\text{PV}}$

Scenario-based modelization

4 scenario based models are studied to manage the multistage EMS problem.

Scenarios

The following models rely on a finite set of scenarios:

$$\Xi_{t_1:t_2}^S := \left\{\xi^s := \mathrm{Cons}^s - \mathrm{PV}^s : s = 1, \ldots, S\right\} \quad (7)$$

- Each scenario, ξ^s, is associated to a probability $p_s > 0$, satisfying $\sum_{s=1}^S p_s = 1$.
- The more scenarios the better representativity but the harder to compute

Stochastic Programming (SP)

SP optimizes expected cost over known scenarios.

- We look for an optimal control P_b^s for each scenario.
- the controls undergo a non-anticipativity constraint $\mathcal{N}$

SP

$$\inf_{\boldsymbol{P_b} \in \mathbb{X}} \left[\sum_{s=1}^{S} p_s J(\boldsymbol{P_b}(s), \mathrm{Cons}^s, \mathrm{PV}^s) + \mathrm{i}_{\mathcal{N}}(\boldsymbol{P_b}) \right] \tag{SP}$$

- p_s is the probability of the scenario s

- $\mathrm{i}_{\mathcal{N}}(\boldsymbol{P_b}) = 0$ if the non-anticipativity is respected
- $\mathrm{i}_{\mathcal{N}}(\boldsymbol{P_b}) = \infty$ if the non-anticipativity is not respected

- Requires accurate estimation of probability distribution.
- Algorithm : Progressive Hedging Algorithm (PHA).

Robust Optimization (RO)

RO assumes the worst-case scenario among all the possible outcomes.

RO

Minimization of the cost under worst-case scenario:

$$\inf_{\boldsymbol{P_b} \in \mathbb{X}} \max_{s \in \{1,\dots,S\}} J(\boldsymbol{P_b}(s), \mathrm{Cons}^s, \mathrm{PV}^s) + \mathrm{i}_{\mathcal{N}}(\boldsymbol{P_b}) \tag{RO}$$

- Overly conservative.

Distributional Robust Optimization (DRO)

DRO optimizes against worst-case distribution in an ambiguity set.

DRO

$$\inf_{\boldsymbol{P_b} \in \mathbb{X}} \sup_{q \in \mathcal{P}_\theta} \left[\sum_{s=1}^{S} q_s J(\boldsymbol{P_b}(s), \mathrm{Cons}^s, \mathrm{PV}^s) + \mathrm{i}_{\mathcal{N}}(\boldsymbol{P_b}) \right] \tag{DRO}$$

- The scenarios are untouched but DRO optimize over the worst distribution of weights of the scenarios within an ambiguity set.

Wasserstein-based ambiguity sets.

$$\mathcal{P}_\theta \;:=\; \left\{ q \;\in\; \mathbb{R}_+^S \;:\; \sum_s q_s \;=\; 1, \mathrm{W}_2 \left(\sum_{s=1}^{S} q_s \delta_{\xi^s}, \sum_{l=1}^{L} p_l \delta_{\xi^l} \right) \;\leq\; \theta \right\} \tag{8}$$

where W_2 is the 2-Wasserstein distance: a distance between probability measures

- θ controls the size of $\mathcal{P}_\theta$,
- large values of θ can lean to RO
- small values of θ can lead to SP

- Algorithm for multistage DRO: SDAP[1].

[1] van Ackooij, W. S., and de Oliveira, W. L. (2025). Scenario Decomposition with Alternating Projections. In Methods of Nonsmooth Optimization in Stochastic Programming

Stochastic Programming with Variance Penalization (VSP)

VSP introduces a variance regularization into SP.

VSP

$$\inf_{\boldsymbol{P_b} \in \mathbb{X}} \Bigg[\sum_{s=1}^{S} p_s J_{t_1:t_2}(\boldsymbol{P_b}(s), \mathrm{Cons}^s, \mathrm{PV}^s) + \mathrm{i}_{\mathcal{N}}(\boldsymbol{P_b}) + \frac{\alpha}{2} \sum_{s=1}^{S} p_s \| \boldsymbol{P_b}(s) - \textstyle\sum_{s'=1}^{S} p_{s'} \boldsymbol{P_b}(s') \|_{\mathrm{L}^2}^2 \Bigg] \quad \text{(VSP)}$$

- $\alpha = 0$: VSP is equivalent to SP
- $\alpha = \infty$: a unique control is found for all scenarios that minimize the expectancy

- Trade-off parameter α controls robustness.
- Solved via Regularized Progressive Hedging (RPHA)[2].

[2] Malisani, P., Spagnol, A., and Smis-Michel, V. (2024). Robust stochastic optimization via regularized PHA: application to Energy Management Systems.

Reinforcement Learning (RL)

RL learns an optimal behavior by interacting with an environment and receiving costs from these interactions.

- We define a Markov Decision Process as $(\mathcal{T}, \mathcal{S}, \mathcal{A}, \mathbb{P}, c)$

- $\mathcal{T} = \{t_1, t_1 + \Delta, \dots, t_2\}$ the finite time horizon,
- $\mathcal{S}$ is the state space:
  - $E \in \{0, dE, \cdots, 13 \text{ kWh}\}$ is the discretized energy state of the battery,
  - $\bar{\xi} = \text{Cons} - \text{PV} \in \{-8 \times 2, -16 + d\bar{\xi}, \cdots, 16\rho_c \text{ kW}\}$ is the difference between the electricity demand and the solar production
- $\mathcal{A}$ is the action space where the action P_b is the battery charging power,
- $\mathbb{P}$ is the transition probability of passing from state s to state s' given action a,
- c is the cost function : $c^\tau(s, P_b) = p_r^c(\tau) \max\{P_m^\tau, 0\} + p_r^d(\tau) \min\{P_m^\tau, 0\}$.

Bellman Equation

RL[3] minimizes the Q-function with respect to π

- π is the politic : $P_b^\tau = \pi(\tau, s)$

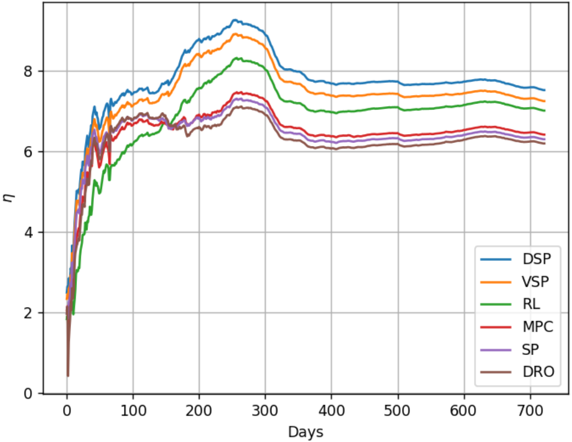
$$Q^\pi(t, s, P_b) = \mathbb{E}\left[\sum_{\tau=t}^{t_2} c^\tau(\boldsymbol{s}^\tau, \boldsymbol{P}_b^\tau) \mid \boldsymbol{s}^t = s, \boldsymbol{P}_b^t = P_b\right]. \tag{9}$$

- Learns control policy through experience.
- Scenario-free method.

[3] Weber, L., Bušić, A., and Zhu, J. (2023, December). Reinforcement learning based demand charge minimization using energy storage. In 2023 62nd IEEE CDC

Performance Comparison

- After cross-validation for θ in SDAP (DRO) and α in RPHA (VSP), the methods are tested over a 2 year period: 2022-01-22 to 2024-01-22.
- $\eta(\text{Day}) := 100 \times \frac{B-\text{Bill}}{B}$, where $B := \text{Bill}_{\text{no battery}}$



- VSP ($\alpha = \infty$) performs best overall.
- RL is competitive with less tuning.
- DRO probably limited by the number of scenarios it can deal with.

Take-away messages

- All algorithms have been tested in real conditions over a 2 year period,
- $\mathrm{VSP}_{\alpha=\infty}$ is the best model for our EMS problem. $\mathrm{RPHA}_{\alpha=\infty}$ is fast to compute and easy to implement.

D. Mimouni, P. Malisani, J. Zhu, W. de Oliveira. A Comparative Study of Optimization Approaches for Battery EMS in Commercial Buildings.
Will be submitted to an applied journal.

- Preprint available soon
  `https://dan-mim.github.io/publications`
- Python code is available at
  `https://gitlab.ifpen.fr/R1150/malisanp`

Thank you!

Contact:

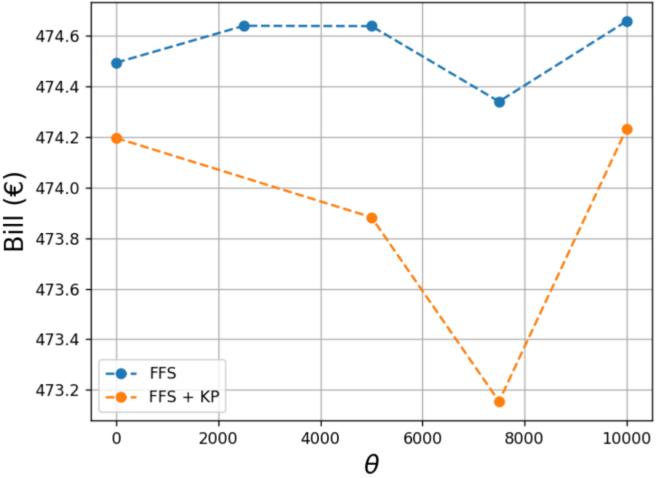
✉ daniel.mimouni@ifpen.fr, paul.malisani@ifpen.fr

https://dan-mim.github.io

Annex

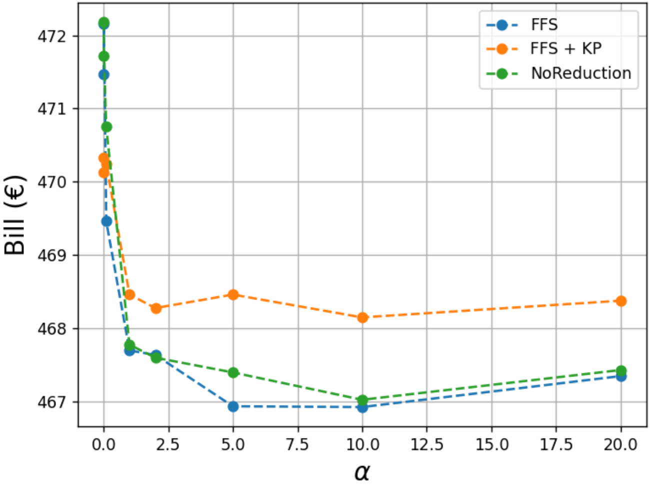
Cross-validation SDAP

- Large scenario trees created from historical data.
- Reduced using FFS and KP methods.

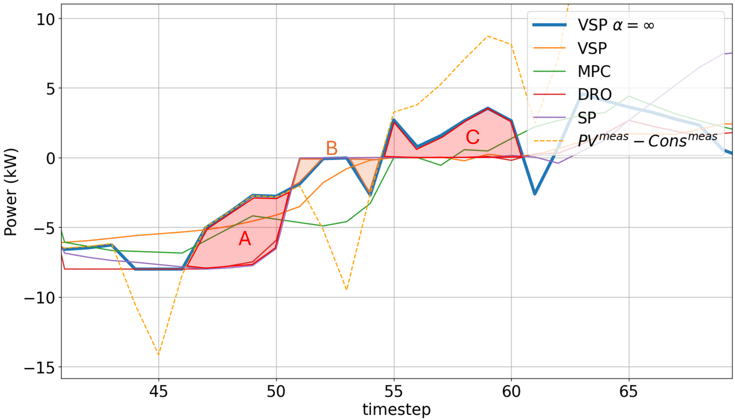


Cross-validation RPHA

- Large scenario trees created from historical data.
- Reduced using FFS and KP methods.



Battery Strategy Analysis



- Overdischarge explains poorer results,
- VSP sticks the most to the production-consumption balance.