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**

ICCOPT 2025
Optimization in Power Systems and Energy

Outline

I. The Energy Management System (EMS) Problem

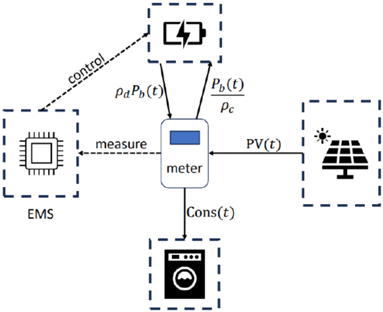
II. Computing solutions for different models

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).

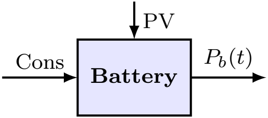


- Traditional models: stochastic programming to minimize expected cost.

Challenge

- How to propose a control robust to unpredictable events ?

Optimization Model for Multistage Energy Management



The effective power demand:

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

Bill

The electricity bill to minimize:

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

which, in our setting, is a convex function of P_b.

The battery's dynamics are governed by the constraints C:

$$\dot{E}(t) = P_b(t)$$

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

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

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

$$E(t_1) = E(t_2) = E^0$$

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:

$$\min_{P_b \in C} J_{t:t+\Delta T}(P_b, \widehat{\text{Cons}}, \widehat{\text{PV}}) \qquad \text{(MPC)}$$

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

Scenario-based modelization

Scenario based models are studied to manage the multistage EMS problem.

Scenarios

The following models rely on a finite set of scenarios:

$$\Xi^S_{t_1:t_2} := \left\{ \xi^s := \mathrm{Cons}^s - \mathrm{PV}^s : s = 1, \dots, S \right\} \tag{1}$$

- 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

Deterministic Control for SP

DSP computes a single control policy that minimizes the expected cost across all scenarios

- We look for an optimal control P_b for all scenarios

DSP

$$\inf_{P_b \in \mathrm{C}} \left[\sum_{s=1}^{S} p_s J_{t_1:t_2}(P_b, \mathrm{Cons}^s, \mathrm{PV}^s) \right], \tag{DSP}$$

- Can be solved using standard control methods such as [1]
- DSP is numerically tractable even when using a large number of scenarios

[1] Malisani, P. (2024). Interior point methods in optimal control. ESAIM: Control, Optimisation and Calculus of Variations, 30, 59.

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

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

- p_s is the probability of the scenario s

- Requires estimation of probability distribution.
- A possible 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:

$$\inf_{\substack{\boldsymbol{P_b} \in \mathbb{C} \\ \boldsymbol{P_b} \in \mathcal{N}}} \max_{s \in \{1, \ldots, S\}} J(\boldsymbol{P_b}(s), \mathrm{Cons}^s, \mathrm{PV}^s) \tag{RO}$$

- Overly conservative in general.

Distributional Robust Optimization (DRO)

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

DRO

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

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

Wasserstein-based ambiguity sets.

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

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 lead to RO
- small values of θ can lead to SP

- A possible algorithm for multistage DRO: SDAP[2].

[2] 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

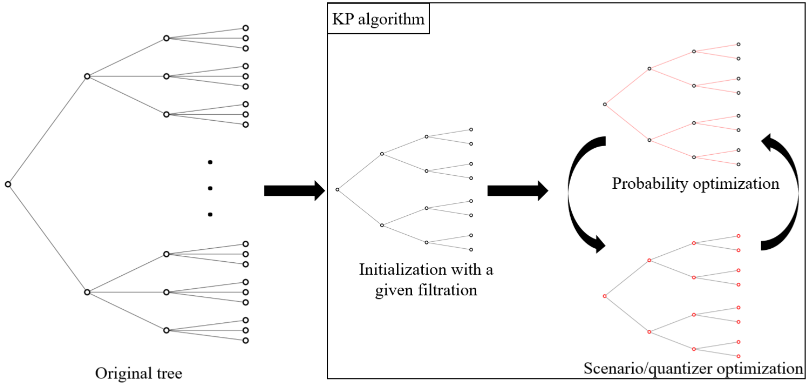
$$\inf_{\substack{\boldsymbol{P_b} \in \mathbb{C} \\ \boldsymbol{P_b} \in \mathcal{N}}} \Bigg[\sum_{s=1}^{S} p_s J_{t_1:t_2}(\boldsymbol{P_b}(s), \mathrm{Cons}^s, \mathrm{PV}^s) + \frac{\alpha}{2} \sum_{s=1}^{S} p_s \| \boldsymbol{P_b}(s) - \textstyle\sum_{s'=1}^{S} p_{s'} \boldsymbol{P_b}(s') \|_{\mathrm{L}^2}^2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

- $\alpha = 0$: VSP is equivalent to SP
- $\alpha = \infty$: VSP is equivalent to DSP

- Trade-off parameter α controls robustness.
- Solved via Regularized Progressive Hedging (RPHA)[3].

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

Scenario Reduction Algorithms



Kovacevic & Pichler algorithm: to approximate a tree, a smaller tree with a given filtration is improved in order to minimize the distance with the original tree. The probabilities and the scenario values are alternatively optimized until convergence[4].

[4] More on this subect at ICSP 2025 and in Mimouni, D., Malisani, P., Zhu, J., & de Oliveira, W. (2024). Scenario Tree Reduction via Wasserstein Barycenters.

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[5] minimizes the Q-function with respect to π

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

$$Q^\pi(t, s, P_b) = \mathbb{E}\left[\sum_{\tau=t}^{t_2} c^\tau(\boldsymbol{s}^\tau, \boldsymbol{P}_b^\tau) \mid \boldsymbol{s}^t = s, \boldsymbol{P}_b^t = P_b\right]. \tag{2}$$

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

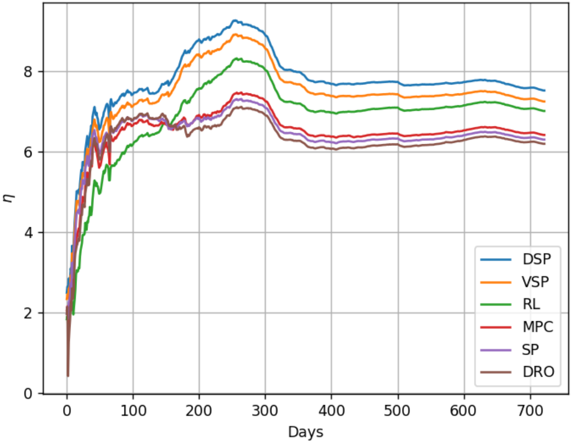
[5] Weber, L., Bušić, A., and Zhu, J. (2023). Reinforcement learning based demand charge minimization using energy storage. In 2023 62nd IEEE CDC

Models to Methods

- MPC
  One prediction
- DSP → Control optimization (interior points)
  One control for all scenarios
- SP → PHA
  Unknown distribution apriori
- DRO → SDAP
  Ambiguity set
- VSP → RPHA
  Variance penalization
- RL
  Scenario free model

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 $B := \text{Bill}_{\text{no battery}}$



- DSP 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 models have been tested in real conditions over a 2 year period,
- DSP is the best model for our EMS problem. DSP 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.

- Preprint available soon
  `https://dan-mim.github.io/publications`
- Python code is available at
  public: `https://github.com/dan-mim/EMS-RL-DRO`,
  industrial: `https://gitlab.ifpen.fr/R1150/malisanp`

Thank you!

Contact:

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

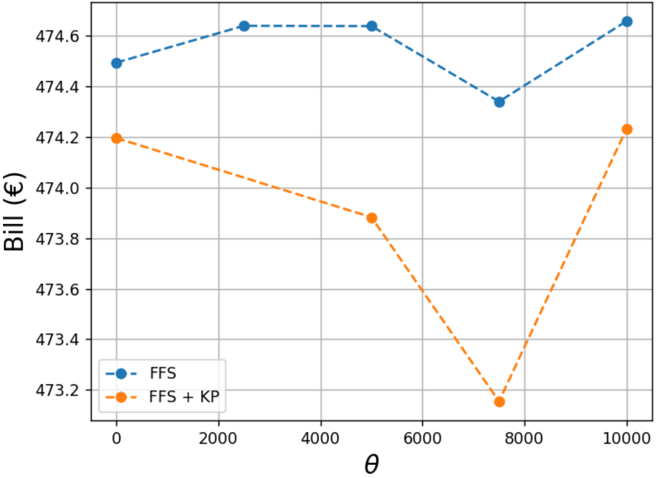
`https://dan-mim.github.io`

IFP Energies nouvelles | MINES PARIS | PSL

Annex

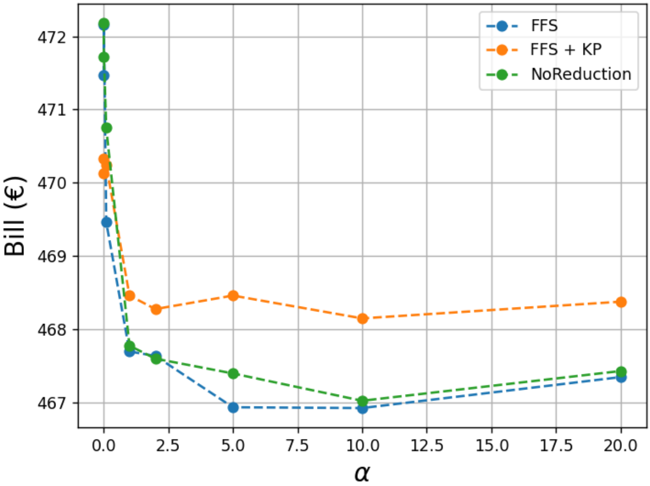
Cross-validation SDAP

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

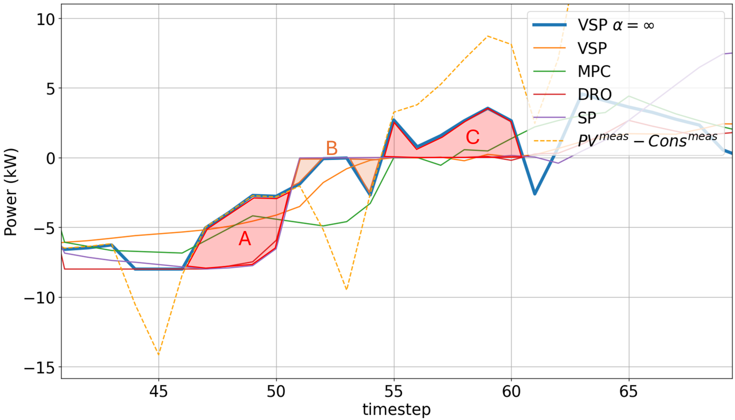


Cross-validation RPHA

- 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.

Performance Ratio

We define four criteria to help explain the differences in model performance, where $P_d := -\min(\rho_d P_b^{\text{meas}}, 0)$ and $P_c := \max(P_b^{\text{meas}}/\rho_c, 0)$:

- **Autoproduction gain ratio:**

$$\texttt{PG} = 100 \times \frac{\int \texttt{min}\{P_d(t), \text{Cons}^{\text{meas}}(t) - \text{PV}^{\text{meas}}(t)\} \mathbb{1}_{\mathcal{C}_1}(t)\text{d}t}{\int \text{Cons}^{\text{meas}}(t)dt}. \quad (3)$$

Where, $\mathcal{C}_1 : \text{PV}^{\text{meas}}(t) < \text{Cons}^{\text{meas}}(t)$.

- **Autoconsumption gain ratio:**

$$\texttt{CG} = 100 \times \frac{\int \texttt{max}\{P_c(t), \text{PV}^{\text{meas}}(t) - \text{Cons}\} \mathbb{1}_{\mathcal{C}_2}(t)dt}{\int \text{PV}^{\text{meas}}(t)dt}.$$

Where, $\mathcal{C}_2 : \text{PV}^{\text{meas}}(t) > \text{Cons}^{\text{meas}}(t)$.

Performance Ratio

- **Discharging error ratio:**

$$\texttt{DE} = 100 \times \frac{\int \left(P_d(t) - (\text{Cons}^{\text{meas}}(t) - \text{PV}^{\text{meas}}(t))\right) \mathbb{1}_{\mathcal{C}_3}(t)dt}{\int \text{Cons}^{\text{meas}}(t)dt}.$$

Where, $\mathcal{C}_3 : \text{PV}^{\text{meas}}(t) < \text{Cons}^{\text{meas}}(t)$ and $\text{Cons}^{\text{meas}}(t) - \text{PV}^{\text{meas}}(t) < P_d(t)$.

- **Grid charging ratio:**

$$\texttt{GC} = 100 \times \frac{\int (P_c(t) - (\text{PV}^{\text{meas}}(t) - \text{Cons}^{\text{meas}}(t)))\mathbb{1}_{\mathcal{C}_4}(t)dt}{\int \text{PV}^{\text{meas}}(t)dt}.$$

Where, $\mathcal{C}_4 : \text{PV}^{\text{meas}}(t) > \text{Cons}^{\text{meas}}(t)$ and $\text{PV}^{\text{meas}}(t) - \text{Cons}^{\text{meas}}(t) < P_c(t)$.

Performance Ratio

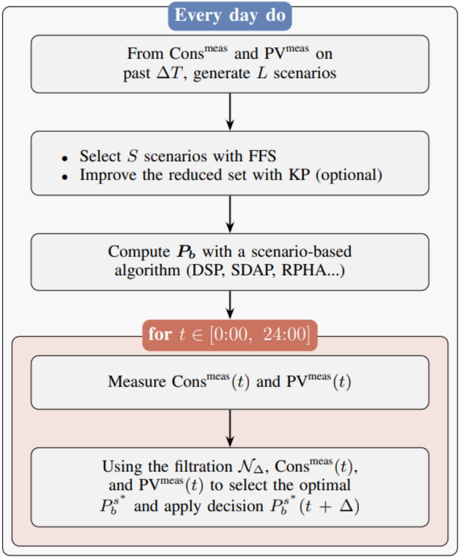
Table: Evaluation of models according to four performance criteria. Values in parentheses indicate the percentage difference relative to MPC.

Method	CG	PG
DSP	1.2900 (49.7%)	1.4437 (12.2%)
VSP	0.9248 (7.3%)	1.3086 (1.7%)
RL	0.9189 (6.6%)	1.1824 (-8.1%)
MPC	0.8620 (0.0%)	1.2867 (0.0%)
DRO	0.8850 (2.7%)	1.2715 (-1.2%)
SP	0.8811 (2.1%)	1.2883 (0.1%)

Method	DE	GC
DSP	0.0378 (-65.3%)	2.1109 (-9.8%)
VSP	0.0783 (-28.1%)	2.2602 (-3.4%)
RL	0.0649 (-40.5%)	2.0080 (-14.2%)
MPC	0.1090 (0.0%)	2.3400 (0.0%)
DRO	0.1135 (4.1%)	2.2927 (-2.0%)
SP	0.1223 (12.2%)	2.3593 (+0.8%)

Decision Process for Scenario Based Methods

Process 1 Decision process for scenario-based methods



The control algorithm, where $\text{Cons}^{\text{meas}}, \text{PV}^{\text{meas}}$ denote, respectively, the electric consumption and photovoltaic production measured at the meter.