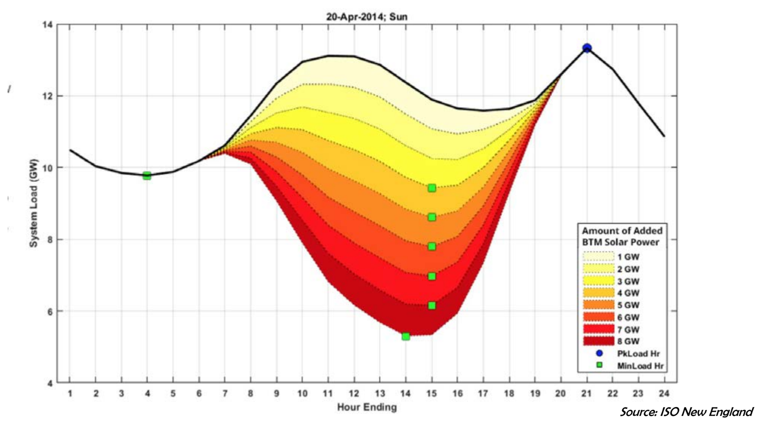
Energy Management for Microgrids: a Reinforcement Learning Approach

Tanguy Levent*, Yvan Bonnassieux, Philippe Preux, Jordi Badosa, Gonzague Henri, Erwan Le Pennec
*tanguy.levent@polytechnique.edu

Motivation

Large penetration of renewable energy could
- Weaken the grid
- Causing blacktout

because of their intermittent and unpredictable nature. Power imbalance between peak demand and renewable production is a challenge (duck curve)!

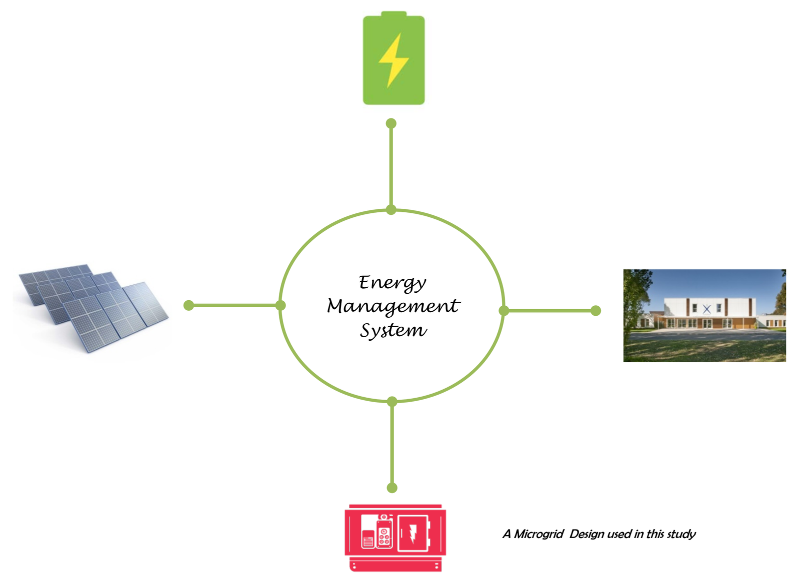


Source: ISO New England

One solution : **Microgrid** could
- Keeping the balance with the utility grid
- Reducing the peak
- Reducing periods of load variability
- Enhancing the power quality and service
- Decrease the feeder looses

Goals

An **Energy Management System** (EMS) is a collection of computer aided-tools used by power operator to monitor, control and optimize the grid performance. A **Microgrid** is a single small scale power system.



A Microgrid Design used in this study

The study goal is to create a **smart** EMS to manage the power dispatch of a microgrid to minimize the operation costs, while maintening the grid stability: the **Economic Dispatch** is an optimisation problem

Modeling

Photovoltaic (PV) power output data comes from the Site Instrumental de Recherche par Télédétection Atmosphérique (SIRTA)

PV panels at SIRTA

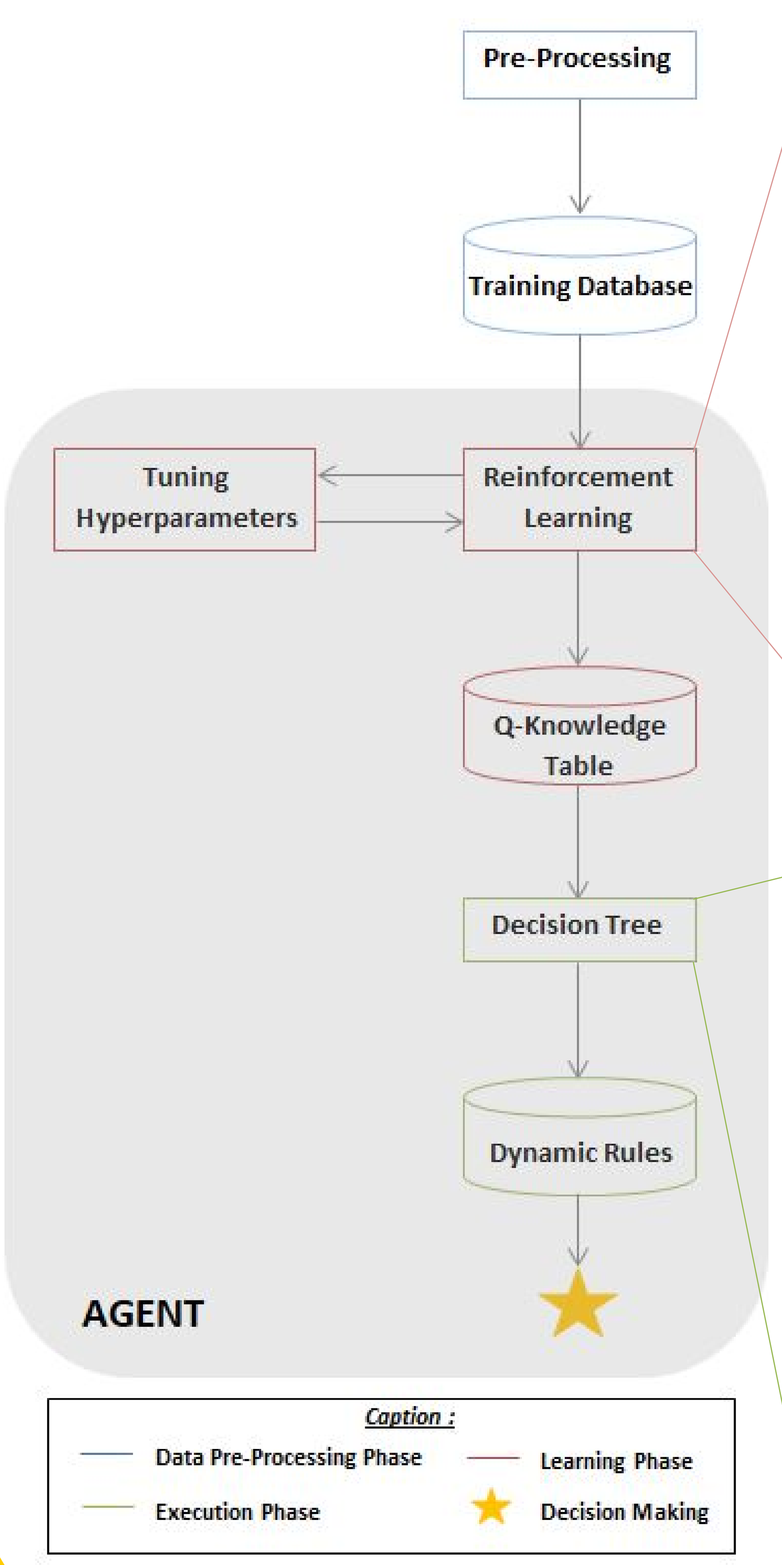
Consumption data measurements come from a tertiary building: the Drahi X Novation Center

Drahi X Novation Center

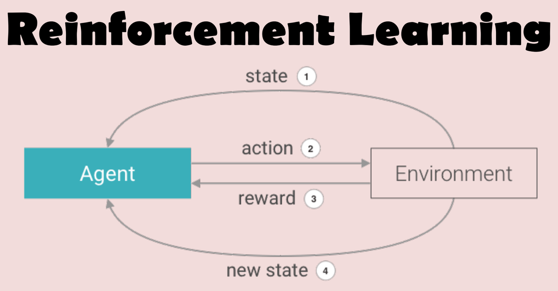
The model of this study: Microgrid **islanded** (not connected to the main grid) with the PV panels, set of batteries, a diesel generator (genset), and the building loads.

Methods

Idea: We propose a novel combinaison of two algorithms to solve the Economic Dispotch Problem of a microgrid: a learning phase with a Reinforcement Learning (RL) on a small dataset and an execution phase based on a Decision Tree (DT) induced from the trained RL



Reinforcement Learning



Goal: Obtain a function *Q(s,a)* that predicts the best action *a* in a state *s* in order to maximize a cumulative reward

Using **Q-Learning**, which iteratively updates *Q(s,a)* using Bellman Equation:

$$Q(s,a) = Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q(s',a') - Q(s,a) \right]$$

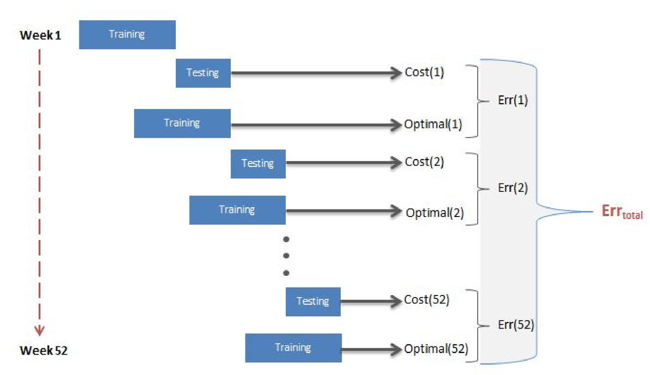
Decision Tree

Goal: Using que Q-table obtained with the RL learning phase as a supervised learning approach. The states as inputs combined with the best action at each state as output

We are using a **CART decision tree** which is a binary tree (only two branches at each node)

Results

Manage the Economic Dispatch over the microgrid model for 52 weekdays. The agent trains over the 4 previous weekdays. The performance indicator $Err(t)$ is the loss between the decision taken by the EMS during the testing day and the optimal cost calculated when the day is ended. The cumulated loss over the 52 weeks is defined by Err_{total}



$$Err_{total} = \sum_{n=1}^{52} Err(t)$$

States: Contain all the information to choose the best action:

$$s = (P_{Net}, P_{BCap})$$

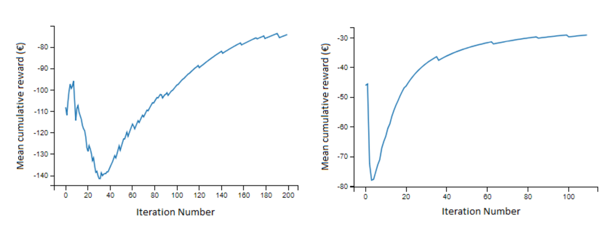
with P_{Net} is the net demand (PVs power output – Consumption in kWh) and P_{BCap} is the battery capacity

Actions: The set of actions A considered in this study is:
- Action 1 = Charge : batteries charge
- Action 2 = Discharge : batteries discharge
- Action 3 = Genset : genset produces electricity
- Action 4 = Idle

Reward Function: $r(s,a)$ is represented as a real value and is associated with the cost of the generator used to meet the net demand P_{Net}. Each generator have a distinct cost and is also affected to the violation of the constraints (included into the model)

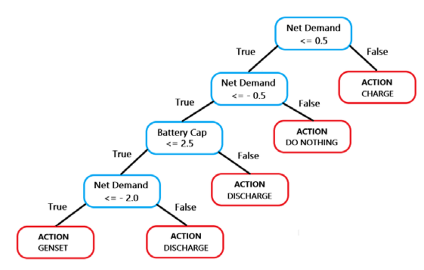
$$r(s,a) = \begin{cases} -m * P_{Net}, & \text{if charge or discharge the battery} \\ -q * P_{Net}, & \text{if power produced by the genset} \\ -c * P_{Net}, & \text{if the constraints are not respected} \\ 0, & \text{if do nothing} \end{cases}$$

Training Phase Results: For each weekday, we use a decreasing number of episodes (ranging from 200 iterations to 30 less / day)

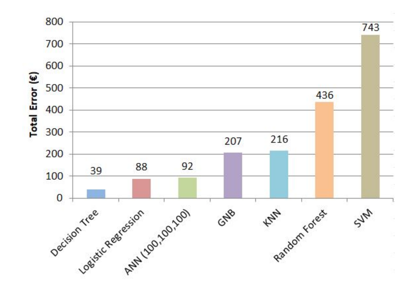


Mean cumulated reward for the 1st day (left) and the 4th day (right)

Execution Phase Results: The average Err_{total} over 10 try is equal to 39.3€ with a standard deviation of 1,90€. We have compared different methods to validate the DT.



A CART decision tree output



Comparison between different machine learning methods

	Optimal	Human	RL-DT*
Coût Total	-6 452;5 €	-6 461.5 €	-6 492.5 €
Perf/Optimal		-0.14%	-0.62%

Table of the algorithm performance
*RL-DT: Reinforcement Learning Decision Tree

References

[1]: R. Lasseter et al. Integration of distributed energy resources. the certs microgrid concept, 2002
[2]: N. Hatziargyriou et al. Microgrids. IEEE Power and Energy Magazine, 2007
[3]: Katiraei et al. Microgrids management. Power and Energy Magazine, 2008
[4]: A.G. Barto R.S. Sutton. Reinforcement Learning: An Introduction, 2018
[5]: E. Kuznetsova et al. Reinforcement learning for microgrid energy management. Energy, 2013
[6]: P. Kofinas et al. Energy Management in Solar Microgrid via Reinforcement Learning, 2016
[7]: L. Breiman et al. Classification and Regression Trees, 1984