



# Optimal multi-criteria management of energy storage systems in a micro-grid

Nathan Célié, Margot Gaetani-Liseo, Amine Lahyani, Ali Sari

## ► To cite this version:

Nathan Célié, Margot Gaetani-Liseo, Amine Lahyani, Ali Sari. Optimal multi-criteria management of energy storage systems in a micro-grid. 2023 IEEE Vehicle Power and Propulsion (IEEE VPPC 2023), IEEE VTS, Oct 2023, Milan, Italy. <10.1109/VPPC60535.2023.10403298>. <hal-04338784>

**HAL Id: hal-04338784**

**<https://hal.science/hal-04338784v1>**

Submitted on 12 Dec 2023

**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.



Distributed under a Creative Commons CC BY 4.0 - Attribution - International License

# Optimal multi-criteria management of energy storage systems in a micro-grid

Nathan CÉLIÉ

*Université Claude Bernard Lyon 1  
Ecole Centrale de Lyon, INSA Lyon, CNRS  
Ampère, UMR5005  
F-69100 Villeurbanne, France  
nathan.celie@insa-lyon.fr*

Ali SARI

*Université Claude Bernard Lyon 1  
Ecole Centrale de Lyon, INSA Lyon, CNRS  
Ampère, UMR5005  
F-69100 Villeurbanne, France  
ali.sari@univ-lyon1.fr*

Margot GAETANI-LISEO

*Université Claude Bernard Lyon 1  
Ecole Centrale de Lyon, INSA Lyon, CNRS  
Ampère, UMR5005  
F-69100 Villeurbanne, France  
margot.gaetani-liseo@univ-lyon1.fr*

Amine LAHYANI

*Laboratoire des Systèmes Electriques, LR11ES15  
INSAT of Tunis  
Tunis, Tunisia  
amine.lahyani@enit.utm.tn*

**Abstract**—With the aim of relocating consumption and production, energy management at the scale of a micro-grid seems to be a lever to improve system efficiency, lifespan and profitability. The purpose of our work is to implement an energy management policy using reinforcement learning techniques in order to maximize the profits and considering different time-varying phenomena: variable power production from a PV system, variable power consumption from a DC load and variable energy price. This work will also take into account the battery degradation as a cost to be minimized. In the near future, this application would enable a better integration of electric vehicles in the grid.

**Index Terms**—Charging Systems and Infrastructures, Micro-grids, Energy management system, Vehicule-To-Grid, Battery degradation, Reinforcement learning

## I. INTRODUCTION

The integration of Renewable Energies Sources (RES) in the power grid and electromobility seems to be a solution to reduce GreenHouse Gases emissions (GHG) and contribute to tackle climate changes and global warming. These considerations have led to major developments of renewable-based technologies such as PhotoVoltaic (PV) and wind power generation. However, these technologies are intermittent and do not have inertia, which leads to power grid instability and management issues. As a solution, Energy Storage Systems (ESS), such as batteries, are added in Micro-Grids (MG) connected to the grid, in order to improve flexibility and allow Energy Management System (EMS) with complex strategies making the all system more reliable and resilient [1]. Moreover, the environmental footprint of buildings is estimated at 30% of GHG emission and energy consumption [2]. More than being a solution for the building themselves, optimal EMS would enable a improve the ESS integration such as stationary batteries, and/or second life batteries and/or Electric Vehicles (EV) in the MG.

Our work aims at finding an optimal energy management policy on a given MG in order to improve its sustainability and its profitability. The MG considered in this paper had been scaled in articles [3] and [4].

Markov Decision Processes (MDP) and Agent-based algorithms such as Q-learning have shown a good efficiency in high dimension sequential optimisation problems and will be used in our study. This techniques are already used for solving such problems and they are mainly focused in the cost of sold and purchased energy. The purpose of our work is to add the degradation of the ESS both in the model and as a decision variable.

Section II presents a brief literature review on what have been done so far in terms of Reinforcement Learning (RL) techniques applied to EMS in MG. Section III aims at giving a model to the system considered in this study. Section IV gives the formulation of the problem that has to be solved by the RL algorithms. Section V presents some results, discussions and perspectives of this work. Section VI gives the conclusions.

## II. STATE OF THE ART

Most of the subjects which deal with optimal EMS in MG are mainly focused on the amount of purchases and sells of energy. For example, the work presented in [5] contributed to show the interest of Q-learning methods in energy management problems with economic criteria, compared to deterministic methods and the relevance of a good discrete size. The work done in [6] on the time basis of a day consisted in a comparison between a trading and a non-trading case, and proved the relevance of the latter. Moreover, in this type of optimisation problem the aging aspect is barely considered and could represent a non-negligible potential blow on a long term basis.

Other non-economic criteria are often taken into account without being considered as part of the objective function. For instance, the study carried out for the stoRES project in 2019 was aiming at maximizing the self-consumption ratio on a bunch of sites in Europe considering the evolution of batteries efficiency [7].

Hence, the purpose of our work is to take into account the ESS degradation both as a part of the cost and as a component of a state.

### III. SYSTEM MODELING

The MG considered in this paper consists of a PV production system (i.e. solar panels), an ESS (i.e. a battery pack), an electric DC load (i.e. a lighting network with a set of LEDs) and a connection to a large scale network which enables imports and exports of electrical energy. In Fig. 1, we show the different potential power flows in the MG.

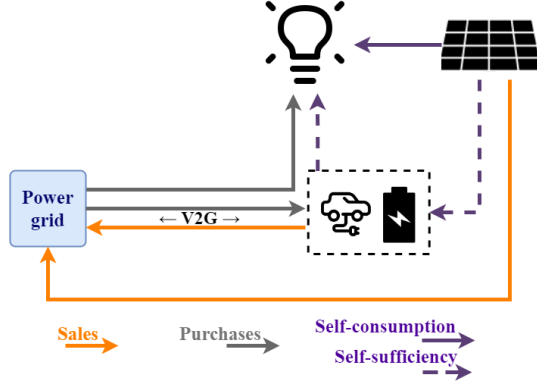


Fig. 1. Micro-grid structure and power flows

The possibility of a Vehicle-To-Grid (V2G) configuration could also be considered with a direct flow between the power grid and the ESS, in both directions. In a first approach this possibility is not implemented.

In our case, the algorithms and the model are implemented in Python language.

#### A. Micro-grid specifications

The PV and ESS specifications have already been set in [3]. This work suggests a new optimal sizing algorithm using a power/energy consideration. An optimal compromise between PV and ESS sizes is proposed with  $0.65kWp$  STC<sup>1</sup> of PV panels and an ESS with a nominal energy  $E_0^B$ , of  $1.87kWh$ . The ESS energy efficiency,  $\eta^B$ , is 0.83, considered as constant. The DC-DC converter connected between the ESS and the DC bus has an efficiency  $\eta^{DC-DC}$  of 0.96 as measured in [8].

The amounts of production and consumption power taken from data have already been affected by the efficiency of the converters.

In the configuration considered in this work, the PV installation had been sized for the PV installation enable full self-consumption [3].

#### B. Input data

The input data of PV power production come from the database of the building integrated photovoltaic (BiPV) described in [3], as well as the power consumption data available in [4] and [3]. The power load profile is the consumption of

the DC lighting network [3]. The database consists of two years (730 days) of readings, collected at a 1 min time-step, between 2016-01-01 and 2017-12-31. PV power production and DC load power consumption are respectively denoted  $p_t^{PV}$  and  $p_t^{load}$ . These powers are both given in kiloWatts [kW].

A database description is detailed in table I.

TABLE I  
INPUT DATA : PV PRODUCTION AND LOAD CONSUMPTION

	$p_t^{PV}$	$p_t^{load}$
count	1036800	1051200
mean	0.0838 kW	0.006 kW
std	0.144 kW	0.017 kW
min	0.0 kW	0.099 kW
max	0.578 kW	0.0 kW

Data related to PV production have been verified thanks to another database extracted from the software PVGIS [9] with the appropriate location and installation characteristics.

In order to find some patterns, each day has been divided into 48 time intervals where each time interval has a duration  $\Delta t$  taken at 30 min and the recordings of  $p_t^{PV}$  and  $p_t^{load}$  in the input data have been averaged for every time-step interval.

The power profile  $p_t^{PV}$ ,  $p_t^{load}$ , as well as all time-dependant variables are discrete-time variables with a step time  $\Delta t$ . Hence the notation  $t$  corresponds to one of the 48 time intervals of a given day as described in equation (1).

$$\forall t, t = n\Delta t \mid n \in \mathbb{N} \quad (1)$$

#### C. Power flows along the DC bus

The electrical architecture of the study MG is represented in Fig. 2.

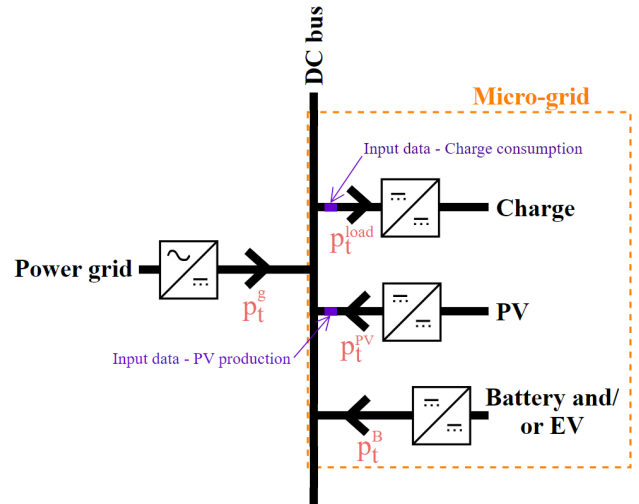


Fig. 2. Schematic and power flow for the considered MG

The power produced by the PV system  $p_t^{PV}$  can only be positive or null. The power received by the DC load  $p_t^{load}$

<sup>1</sup>Standard Test Conditions, i.e.  $1000W/m^2$ , AM1.5,  $25^\circ C$

can only be positive or null. The ESS can both receive or deliver power and we set that  $p_t^B$  is positive when battery is discharging and negative when it is charging. The power that is purchased or sold from and to the power grid, called  $p_t^g$ , is taken positive when energy is bought and negative when energy is sold. These conventions are illustrated in the schematic diagram of Fig. 2.

A balance power,  $p_t^{bal}$ , is derived from Fig. 2 as a difference between PV power and load demand, and defined with equation (2).

$$p_t^{bal} = p_t^{PV} - p_t^{load} \quad (2)$$

If we sum up the powers along the DC bus we get (3).

$$p_t^g + p_t^B + p_t^{PV} - p_t^{load} = 0 \quad (3)$$

And then equation (3) is simplified with equation (2) to get (4).

$$p_t^g + p_t^B + p_t^{bal} = 0 \quad (4)$$

In our case, only full power flows are considered which means that the power  $p_t^B$  going to or from the ESS cannot be split or lowered. The same rule is applied to the powered exchanged with the power grid  $p_t^g$ . Hence  $p_t^B$  and  $p_t^g$  cannot be non-zero on a same time interval.

#### D. Battery model

In our work, the ESS charge level is given in terms of energy as a State of Energy (SoE) which is given by equation 5 during charge and by equation 6 during discharge operating modes.

$$SoE_{t+\Delta t} = SoE_t - \eta^{DC-DC} \times \eta^B \times \frac{P_t^B \times \Delta t}{E_0^B} \quad (5)$$

$$SoE_{t+\Delta t} = SoE_t - \frac{1}{\eta^{DC-DC}} \times \frac{P_t^B \times \Delta t}{E_0^B} \quad (6)$$

Where  $E_0^B$  denotes the initial amount of energy in the ESS.

The ESS State of Health (SoH), with energy consideration is given by equation 7, so called  $SoH_t^E$ .

$$SoH_t^E = \frac{E_0^B - E_t^{B,loss}}{E_0^B} \quad (7)$$

#### IV. PROBLEM FORMULATION

The EMS aims at finding an optimal policy for battery charge and discharge in order to ensure system profitability taking as inputs the production and consumption powers as well as ESS ageing. The goal of this optimization can be expressed in the paradigm of the Q-learning as a *Reward* to be maximized.

A global *Reward* over the whole time of the experience, noted  $T$ , has to be maximized. It is defined as  $r_{total}$  in equation (8).

$$r_{total} = \sum_{t=0}^T r_t \quad (8)$$

Where each *Reward*,  $r_t$ , is the opposite of a cost, noted  $c_t$ , as defined in equation (9). The cost  $c_t$  is split into three components, namely the cost of the power sold to the grid,  $c_t^S$ , the cost of the power purchased from the grid,  $c_t^P$ , and the cost of the battery degradation  $c_t^{deg}$ , as described in equation (10). Those costs have not been normalized in order to show their actual impact on the policy.

$$r_t = -c_t \quad (9)$$

$$c_t = -c_t^S + c_t^P + c_t^{deg} \quad (10)$$

The ESS degradation cost will be explained in subsection IV-A. The costs  $c_t^S$  and  $c_t^P$  are both defined as follows :

$$c_t^S = p_t^{g,S} \times rate_t^S \times \Delta t \quad (11a)$$

$$c_t^P = p_t^{g,P} \times rate_t^P \times \Delta t \quad (11b)$$

Where the power exchanged with the power grid  $p_t^g$  is split into two variables  $p_t^{g,S}$  and  $p_t^{g,P}$  for sale and purchase modes, respectively.

Rates of sold and purchased energy, respectively  $rate_t^S$  and  $rate_t^P$  are based on the french market which uses an *on-peak / off-peak* system for purchased energy and a constant rate for sold energy.

The energy sales rate is considered constant during the period of the study at  $rate_t^S = 0.133\text{€}/\text{kWh}$ , which is the rate for the first semester of 2023 in France, for a PV installation lower than  $3\text{kWp}$ , according to [10].

In order to highlight the influence of time periods in energy purchase from the main grid, a fluctuation of energy cost during the day has to be considered. The french system imposes a format of alternating between *on-peak* and *off-peak* hours. In order for these rates fluctuations to be relevant to the algorithms, they must occur during system's operating periods, i.e. during times when load consumption and PV panels production are not zero (approximately between 7 a.m. and 6 p.m. each day). The chosen format for *on* and *off-peak* hours is given in TABLE II.

TABLE II  
ON AND OFF -PEAK RATES FOR DAILY TIME INTERVALS

Time interval	Rate (period 1, period 2, period 3) [€/kWh]
00:00 - 06:00	off-peak rate (0.111, 0.127, 0.125)
06:00 - 12:00	on-peak rate (0.160, 0.156, 0.159)
12:00 - 13:00	off-peak rate (0.111, 0.127, 0.125)
13:00 - 23:00	on-peak rate (0.160, 0.156, 0.159)
23:00 - 00:00	off-peak rate (0.111, 0.127, 0.125)

Considering an installation of an apparent power lower than 6kVA and for the energy rates in France for 2022-2023, we get the hourly rates from [11].

#### A. Cost of battery degradation

The novelty of this article is to consider the battery degradation cost  $c_t^{deg}$  in the reward function, using the energy loss,  $\Delta E_t^{B,loss}$ , in the amount of "storable" energy. The integration of the ESS degradation is based on the degradation model of [12]. Hence the loss of energy can be interpreted as an aging cost that is updated at each time-step according to equation (12) and inspired by the work done in [13].

$$c_{total}^{deg} = \sum_0^T c_B \times \frac{\Delta E_t^{B,loss}}{(1 - SoH_{lim}^E) \times E_0^B} = \sum_0^T c_t^{deg} \quad (12)$$

Thus, the degradation cost at each time interval is given by (13).

$$c_t^{deg} = c^B \times \frac{\Delta E_t^{B,loss}}{(1 - SoH_{lim}^E) \times E_0^B} \quad (13)$$

In equations (12) and (13),  $c^B$  stands for the investment cost of a new ESS (a set of batteries pack) and the factor  $(1 - SoH_{lim}^E)$  represents the admissible relative loss in term of energy. This way, if the  $E_t^{B,loss}$  energy losses are added up after each time interval, when we reach a total loss of  $(1 - SoH_{lim}^E)$  compared to the nominal energy available  $E_0^B$  (i.e. it remains  $SoH_{lim}^E$  of  $E_0^B$ ), then the ESS has to be changed at the price of the initial cost  $c_B$ . The parameter  $SoH_{lim}^E$  is set to 0.7 in this study.

The determination of  $\Delta E_t^{B,loss}$  [%] is based on results given in [12] which shows the evolution of this loss as a function of full equivalent cycles (FEC) in a Lithium Iron Phosphate (LFP) storage system. Only a characteristic at low current solicitation is considered. The original FEC data has been translated in terms of exchanged energy, considering the initial energy of our storage system,  $E_0^B$ . This conversion enables to give a  $\Delta E_{t(carac)}^{B,loss}$  at each time interval, as shown on Fig.3 with the black curve.

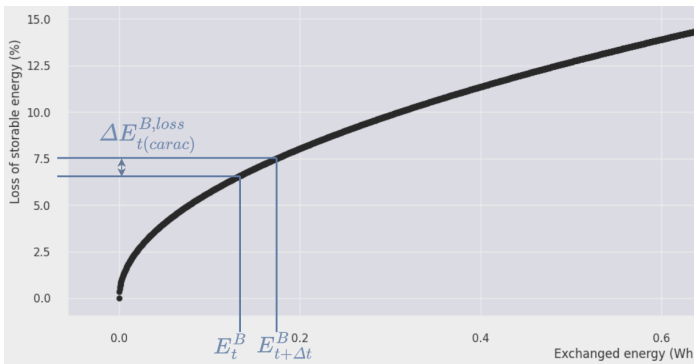


Fig. 3. Loss of energy characteristic and determination of  $\Delta E_{t(carac)}^{B,loss}$

Our aging model takes as inputs the power exchanged by the ESS and its  $SoE$ . Only the cycling aging is taken into account in our study. At each times step, the loss  $\Delta E_t^{B,loss}$  is calculated and a cost of degradation is computed according to 13. This loss depends on the amount of energy that has been exchanged since the beginning of the simulation  $E_t^B$ , the amount of energy that is going to be exchanged between  $t$  and  $t + \Delta t$ . The determination of  $\Delta E_{t(carac)}^{B,loss}$  is represented by the blue lines on Fig3 which is performed at each time interval. The storable amount of energy in the ESS is decreasing along the simulation.

The characteristic above gives a percentage of loss. Equation 14 converts it in terms of actual amount of energy.

$$\Delta E_t^{B,loss} = \Delta E_{t(carac)}^{B,loss} \times E_0^B \quad (14)$$

Then the loss calculated in 14 enables to update  $SoH_t^E$  at each time step as already described in 7 and 13.

#### B. Specifications of the Q-learning algorithms

The formalism of RL algorithms relies on definition of the different parameters as close as possible to the physics of the MG.

In order for the future policy not to end up in forbidden states, some global boundaries (i.e. external of the algorithms) have to be defined, such that the ESS remains between its energy limits  $E_{min}^B$  and  $E_{max}^B$ . The flowchart of fig 4 gives the method of determination of the allowed action in each situation.

After the execution of the algorithm described above, the returned actions are given as inputs to the RL algorithm.

The maximum amounts of power that can be admitted and delivered by the ESS in a time-step  $\Delta t$ , namely at  $P_{max}^{PV}$  and  $P_{max}^{load}$  are also set, according to the manufacturer specifications.

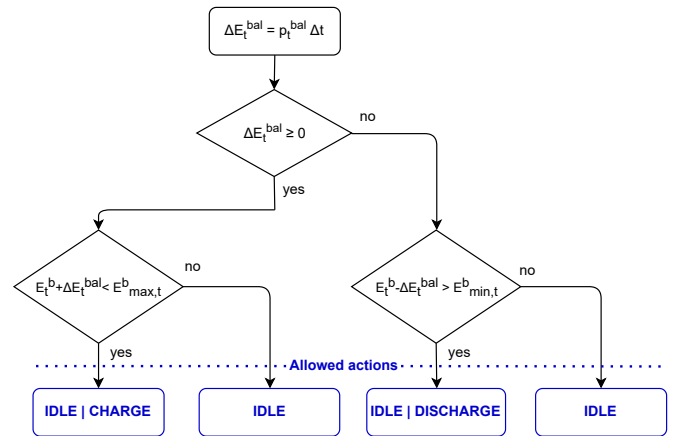


Fig. 4. Flowchart - allowed actions

Each state  $s_t$  is defined by a timing component, some non-controllable exogenous components and controllable components [14] and is given according to equation (15).

$$s_t \in S \mid s = \left( t, d_j, SoE_t, p_t^{PV}, p_t^{load}, rate_t^P, SoH_t^E \right) \quad (15)$$

Where  $t$  is the  $t^{th}$  time-step of a day (out of 48 for a 30-minute time-interval),  $d_j$  is the  $j^{th}$  day of a year,  $SoE_t$  is the level of energy of the ESS at this time-step,  $p_t^{PV}$  is the level of power production at this time-step,  $p_t^{load}$  is the level of power demand during this time-step and  $SoH_t^E$  is the actual battery state of health in term of energy.

The latter component  $SoH_t^E$  has not been used so far and might have its importance in the chosen policy.

Each component has an *a priori* importance in the actions chosen on a specific state and we want to know the influence of each one in the global optimal policy.

In order to keep a reasonable number of states, the levels  $SoE_t$ ,  $p_t^{PV}$  and  $p_t^{load}$  only take a finite number of values. The databases have been simplified in order to keep a reasonable number of states. Thus, the powers have been rounded to the closest tens of watts and the whole database have been shifted to start from zero. Since  $rate_t^S$  is considered as constant, it will not be regarded as a part of a state.

The formulation described in (15) will conduct to a large amount of states but not every single state will eventually be explored by the agent, as long as the trade-off between exploration and exploitation [15] is appropriated to the problem. The  $\epsilon$  parameter has been set to decrease exponentially so that more and more importance is given to exploitation.

Three actions depending on the operating state of the battery are considered:  $a \in A = [-1; 0; 1]$ , defined as follows:

- action = -1 "Discharge"
- action = 0 "Idle"
- action = 1 "Charge"

## V. RESULTS, DISCUSSIONS AND PERSPECTIVES

The first operation consists in the establishment of the optimal policy, for a specific day. This day have been chosen for its representativeness of the whole period. The algorithm performs 5000 episodes over this day and a *Return* is computed for each of them. The evolution of the *Return* after each episode is given in Fig. 5. Each thin blue line represents the occurrence of a maximum return and shows the infinite occurrence of the best policy after some episodes. In order to avoid the algorithm to choose a non-optimal policy, the return is checked backward from the last iteration such that the last occurrence of the maximum value sets the optimal policy.

Starting from this initial optimal policy, each following day is executed as a single episode, so that the policy can adapt to the whole time of the experiment.

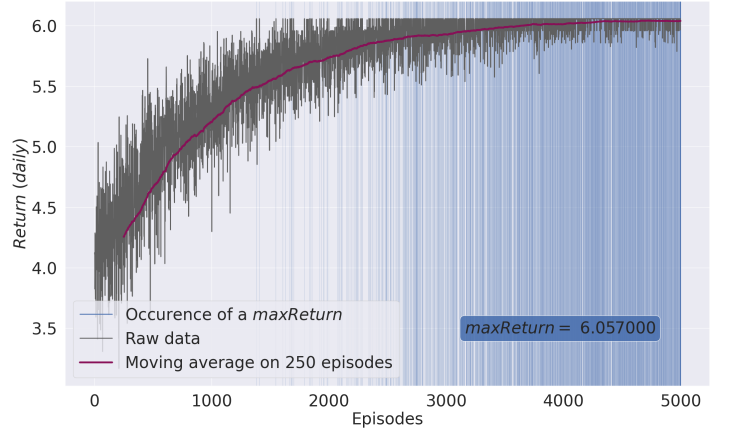


Fig. 5. Evolution of the return over the episodes on a specific day - moving average (250) and occurrences of the maximum *Return*

In order to compare the Q-Learning results in the same conditions, three system configurations with different EMS strategies have been implemented :

- A passive installation, without any PV platform nor ESS is studied for energy cost basic analysis.
- A second topology integrating PV production is considered and enables purchases and sells of energy surplus or deficits.
- A third configuration with the full MG described in Fig. 1, with basic rules EMS (called in literature naive, rules-based or decision tree strategies). It prioritizes ESS charge and discharge at each time step, based on the  $SoE_t$  value limits and on the  $p_t^{bal}$  sign according to equation (4). If the  $SoE_t$  limits are reach it uses the main power grid.

These configurations led to first results concerning the amounts of sold and purchased electrical energy during the experiment time period. The basic EMS strategies can be compared with the strategy derived from the Q-learning algorithms but we have to consider in this last configuration the impact of the exchanges of energy on its aging in order to achieve the actual comparison.

However, these first results showed the interest of a storage system in minimizing the amount of purchased energy as shown in the second and third rows of the table III.

TABLE III  
STANDARD CONFIGURATIONS RESULTS

	Sales of energy	Purchases of energy	Profits
<b>Passive installation</b>	0 €	26.58 €	-26.58 €
<b>Self-consumption (excess sells)</b>	176.05 €	1.81 €	174.24 €
<b>Full micro-grid with standard rules</b>	173.88 €	0.01 €	173.87 €

Concerning the Q-Learning results, the table below gives

the distribution between each action in terms of occurrences over the whole time of the study.

TABLE IV  
PROPORTION OF OCCURRENCES OF EACH ACTION WITH THE OPTIMAL POLICY

Action	Proportion of occurrences
IDLE	0.720
CHARGE	0.104
DISCHARGE	0.176

However, this distribution highly depends on the a priori balance between the different costs. Since non-normalized data have been used, some bias can influence the optimal policy, such as the initial price of the ESS or the typical cost of energy. The choice of a specific country, with its own rules and rates of electrical energy would change the optimal policy as well as the quantitative results.

Furthermore, a specific day is chosen for the establishment of the initial optimal policy. The choice of this day might influence the establishment of the policy. These different conditions and can have major impacts on the quantitative results and it would be worth doing an impact analysis of each of them.

As perspectives of this work and in the case of an application including EV as a temporary storage system, some parameters would enable the micro-grid to charge the EV before the use in priority. Moreover, the ESS is fully described with its level of energy but some other parameters like the State of Charge (*SoC*) could also be taken into account in further developments.

We could imagine splitting  $p_t^{bal}$  between the ESS and the grid in a specific amount determined by an optimal balance parameter. In this case, the self-consumption wouldn't be to be maximized.

We also plan to implement a complete aging model coupling calendar and cycling effects based on the Eyring laws and the works [16], [17].

## VI. CONCLUSIONS

The work presented in this paper aims at improving integration of renewable energy systems and optimising usage of ESS in MG. An optimal EMS based on RL is proposed. Considering the degradation of the storage system both as a cost and as a decision variable is one of the key points of our work.

## VII. ACKNOWLEDGEMENTS

We would like to express our sincere thanks and gratitude to the members of the ISGE team at the LAAS-CNRS lab for letting us use their readings and data. We would also like to thank "Collegium de Lyon for their technical and financial support.

## REFERENCES

- [1] S. Choudhury, "Review of energy storage system technologies integration to microgrid: Types, control strategies, issues, and future prospects," *Journal of Energy Storage*, vol. 48, p. 103966, Apr. 2022.
- [2] A. Mukherji, P. Thorne, W. W. L. Cheung, S. L. Connors, M. Garschagen, O. Geden, B. Hayward, N. P. Simpson, E. Totin, K. Blok, S. Eriksen, E. Fischer, G. Garner, C. Guivarch, M. Haasnoot, T. Hermans, D. Ley, J. Lewis, Z. Nicholls, L. Niamir, S. Szopa, B. Trewin, M. Howden, C. Méndez, J. Pereira, R. Pichs, S. K. Rose, Y. Saheb, R. Sánchez, C. Xiao, and N. Yassaa, "SYNTHESIS REPORT OF THE IPCC SIXTH ASSESSMENT REPORT (AR6)," 2022.
- [3] M. Gaetani-Liseo, C. Alonso, and B. Jammes, "Identification of ESS Degradations Related to their Uses in Micro-Grids: application to a building lighting network with VRLA batteries," *European Journal of Electrical Engineering*, vol. 23, no. 6, p. 455, 2021.
- [4] M. Gaetani-Liseo, *Consideration of energy storage systems and their degradation in microgrids sizing and energy management : influence of model accuracy (Prise en compte des systèmes de stockage de l'énergie et de leurs dégradations dans la gestion et le dimensionnement des micro-réseaux : influence de la précision des modèles, fr)*. phdthesis, Université Paul Sabatier - Toulouse III, Dec. 2021.
- [5] S. Abedi, S. W. Yoon, and S. Kwon, "Battery energy storage control using a reinforcement learning approach with cyclic time-dependent Markov process," *International Journal of Electrical Power & Energy Systems*, vol. 134, p. 107368, Jan. 2022.
- [6] G. Muriithi and S. Chowdhury, "Optimal Energy Management of a Grid-Tied Solar PV-Battery Microgrid: A Reinforcement Learning Approach," *Energies*, vol. 14, p. 2700, May 2021.
- [7] StoRES Project, "LIVING LAB developed within the StoRES project Computation details and hypothesis," 2019.
- [8] M. Gaetani-Liseo, C. Alonso, L. Segulier, and B. Jammes, "Impact on Energy Saving of Active Phase Count Control to a DC/DC Converter in a DC Micro Grid," (Paris), pp. 511–516, IEEE, Oct. 2018.
- [9] "JRC Photovoltaic Geographical Information System (PVGIS) - European Commission."
- [10] In sun we trust | Otovo, "Tarif Rachat Photovoltaïque au 1er Trimestre 2023 | EDF OA," Jan. 2022.
- [11] data.gouv.fr, "History of regulated electricity sales rates for residential consumers (Historique des tarifs réglementés de vente d'électricité pour les consommateurs résidentiels, fr)"
- [12] S. Sun, T. Guan, X. Cheng, P. Zuo, Y. Gao, C. Du, and G. Yin, "Accelerated aging and degradation mechanism of LiFePO<sub>4</sub>/graphite batteries cycled at high discharge rates," *RSC Advances*, vol. 8, no. 45, pp. 25695–25703, 2018.
- [13] T. Dragicevic, H. Pandzic, D. Skrlac, I. Kuzle, J. M. Guerrero, and D. S. Kirschen, "Capacity Optimization of Renewable Energy Sources and Battery Storage in an Autonomous Telecommunication Facility," *IEEE Transactions on Sustainable Energy*, vol. 5, pp. 1367–1378, Oct. 2014.
- [14] B. Mbuwir, F. Ruelens, F. Spiessens, and G. Deconinck, "Battery Energy Management in a Microgrid Using Batch Reinforcement Learning," *Energies*, vol. 10, p. 1846, Nov. 2017.
- [15] B. Lennartson, "Introduction to Discrete Event Systems - Letcure notes," 2009. with contributions from Martin Fabian and Knut Åkesson.
- [16] E. Redondo-Iglesias, P. Venet, and S. Pelissier, "Modelling Lithium-Ion Battery Ageing in Electric Vehicle Applications—Calendar and Cycling Ageing Combination Effects," *Batteries*, vol. 6, p. 14, Feb. 2020.
- [17] A. Houbbadi, E. Redondo-Iglesias, S. Pelissier, R. Trigui, and T. Bouton, "Smart charging of electric bus fleet minimizing battery degradation at extreme temperature conditions," in *2021 IEEE Vehicle Power and Propulsion Conference (VPPC)*, (Gijon, Spain), pp. 1–6, IEEE, Oct. 2021.