


Article

Usage of GAMS-based digital twins and clustering to improve energetic systems control

Timothé Gronier ^{1,2,3} , William Maréchal ^{1*}, Christophe Geissler ¹ and Stéphane Gibout ² *

¹ Advestis, France

² Université de Pau et des Pays de l'Adour, E2S UPPA, LaTEP, Pau, France

³ ADERA, France

* Correspondence: wmarechal@advestis.com, W. M.; stephane.gibout@univ-pau.fr, S. G.)

Abstract: With increasing constraints on energy and resource markets and the non-decreasing trend in energy demand, the need for relevant clean energy generation and storage solutions is growing and is gradually reaching the individual home. But small-scale energy storage is still an expensive investment in 2022 and the risk/reward ratio is not yet attractive enough for individual homeowners. One solution is for homeowners not to store excess clean energy individually but to produce hydrogen for mutual use. In this paper a collective production of hydrogen for a daily filling of a bus is considered. Following our previous work on the subject, the investigation consists of finding an optimal buy/sell rule to the grid, and the use of the energy with an additional objective: mobility. The dominant technique in the energy community is reinforcement learning, which is however difficult to use when the learning data is limited as in our study. We chose a less data-intensive and yet technically well-documented approach. Our results show that rulebooks, different but more interesting than the usual robust rule, exist and can be cost-effective. But they require fine-tuning as to not deteriorate system performance. In some cases, it is worth missing the H₂ production requirement in exchange for higher economic performance.

Keywords: Energy Management System, Digital Twins, General Additive Models, Green H₂.

1. Introduction

With increasing constraints on energy and resource markets and the non-decreasing trend in energy demand, the need for relevant clean energy generation and storage solutions is growing and is gradually reaching the individual home [1,2]. But small-scale energy storage is still an expensive investment in 2022 and the risk/reward ratio is not yet attractive enough for individual homeowners. One solution is for homeowners not to store excess clean energy individually but to produce hydrogen for mutual use.

In this paper a collective production of hydrogen for a daily filling of a bus is considered. Following our previous work on the subject [3], the investigation consists of finding an optimal buy/sell rule to the grid, and the use of the energy with an additional objective: mobility. The dominant technique in the energy community is reinforcement learning, which is however difficult to use when the learning data is limited as in our study. We therefore chose a less data-intensive and yet technically well-documented set of tools.

1.1. Context

The importance of Energy Management Systems (EMS) has increased over the years and is drilling down to individual systems, like individuals or groups of dwellings, or solar community. In a previous article [3] we have studied the opportunity of integrating energy storage technologies for such scales. In this article we will focus on the generation of H₂ in sufficient volume to refuel a bus on a daily basis. We also investigate if an EMS could be developed in order to deliver more interesting results than the usual robust strategy, which does not seek to maximize profits but to make sure H₂ is delivered. Machine learning has shown its ability to take advantage of both volatile and exogenous information [4,5], the kind of which is increasingly common for energy markets and for the behaviour of small systems.



Citation: Gronier, T.; Maréchal, W.; Geissler C.; Gibout, S. . *Preprints* **2022**, *1*, 0. <https://doi.org/>

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1.2. Literature review

Numerous methods have been developed to build EMS systems and to monitor them [6–8]. They can be regrouped in classical programming methods that have proven their worth and are well understood, but suffer from convergence difficulties. Evolutionary and stochastic methods are robust but can be costly in computation time. The Fuzzy logic methods have been developed to manage the multi-control functions and to represent uncertainty in the measured data. Neural Networks are used for control and energy management due to their reliability, computational capability, and adaptability to handle complex nonlinear systems [6]. Their development has given way to the rise of Deep Reinforcement Learning [9]. The technique of Reinforcement Learning is employed to solve sequential decision-making problems that learn by iterative trial and error through a carefully defined reward scheme.

Digital Twins (DT) are a technology that has recently been brought to the front pages [?] of business deciders and is presented as a path to less waste, shorter times to market, constant customer insights. In the literature [10–17], DTs are realized using either two approaches model-based and data-based. Model-based DT requires a fine physical model of the system, while the data-based or machine learning (ML) one requires sufficiently good quality data on the real system, as well as an adequate fine-tuning of the hyperparameters, both of which can be hard to achieve [11]. Accurate models aren't always available for the system of which we want to build a DT for its operating conditions. The problem of the availability of data could possibly be solved using Generative Adversarial Networks (GANs) [18–20] and/or transfer learning [21–25], but the technology is still maturing. For this study, General Additive Models (GAMs) are used as they are a category of ML that requires few to none hyperparameters tuning to approximate nonlinear relationships with a combination of linear formulation of a series of smoothening functions [26].

1.3. Contributions to novelty

Several approaches are being developed for energy management systems [6–9,27,28]. The latest one, based on deep reinforced learning, while producing results and being developed in open-sourced frameworks [29–32] is considered a black-box model [33] and the results are hard to explain, analyze and validate for domain experts. The approach developed in this paper can be considered as a proxy for Off-Policy reinforcement learning [34]. Off-Policy learning algorithms evaluate and improve a policy, called target policy, distinct from the policy used for action selection, called behavior policy. An initial policy is given and used by the system, usually a greedy policy. Then over the iterations, the result of the actions taken following the behavior policy are evaluated as to update the target policy that eventually becomes the new behavior policy [35,36].

Our approach is to create a digital twin of the system that is then used to evaluate different trajectories of the system in the range of possible states that are then classified. The best trajectory or strategy for each class is then chosen, and defines then a policy. The policy is then applied throughout the runs of the simulation, but as operating conditions either exogenous (prices, weather conditions), either endogenous (equipment efficiency, users' behavior), evolve over time, the digital twins need to be re-evaluated. This re-evaluation leads in turn to the need to re-evaluate the virtual system and its classification, thus updating the policy. The trigger condition for such a policy update is not the discovery by the system of better actions to take, but rather the evolution of the operating conditions, that have potentially made the chosen actions non-optimal. In this work, the presented approach appears as a proxy for reinforcement learning.

This approach is innovating in that it uses techniques known by domain experts to approximate reinforcement learning methods, and thus can serve as a bridge between the two domains. This approach can serve as a benchmark against which Deep Learning methods can be compared in a way that can be interpreted by experts in the field.

After the presentation of the general context and the corresponding state of the art, we will present the general methodology retained in our approach (§2). We will then present

the chosen case study (§3), on which the general methodology will be applied; section §4 will thus be dedicated to the details of this implementation. We will then analyze the results obtained (§5) before drawing a conclusion.

2. Methodology

As exposed above, the goal of this article is to propose a method to improve the control of a microgrid and to test it on a case study. The global idea is to create a virtual version of an industrial system in order to create and evaluate an improved policy.

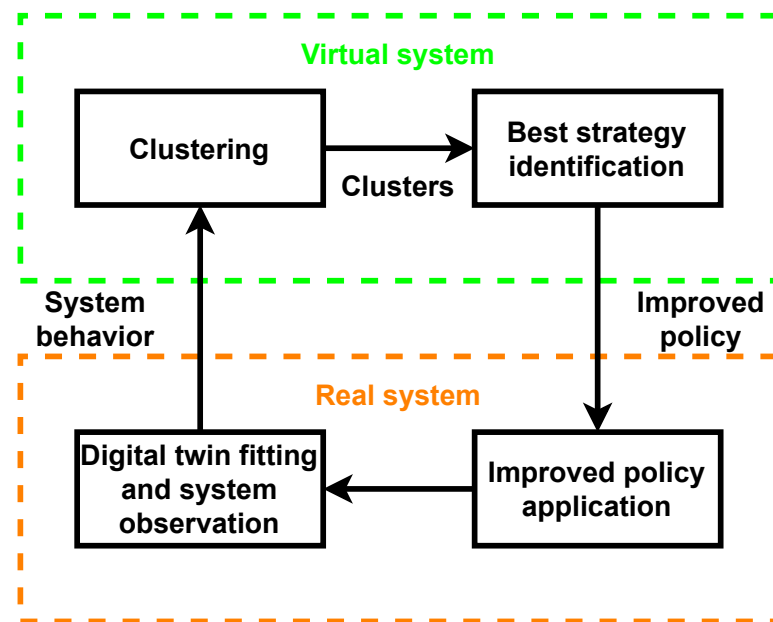


Figure 1. Proposed methodology

The general methodology is shown in Fig. 1. It is divided into four steps which allow the identification of optimal control strategies (clustering and best strategy identification) based on the observation of the real system. The system evolves under the effect of the external instructions but also of the piloting orders recommended by the algorithm (improved policy application). Since the optimization steps require the evaluation of a large number of future trajectories, digital twins are used to represent the behavior of the observed system.

These steps are repeated at intervals for two reasons. First, systems itself tend to evolve through time (typically because their performances decline) and thus digital twins need to be refitted. Second, the "optimal trajectory" evolves according to external conditions that are out of control e.g., the actual electricity price rise and the foreseen reduction of heating consumption are significant environmental changes.

2.1. Digital-twin fitting and system observation

The first step consists in the creation of the virtual system. The idea is to identify the autonomous subsystems composing the microgrid. Each subsystem is then represented by a digital twin. Variables can be exterior inputs to the microgrid (e.g meteorological values) or decisions taken by the manager (send energy to one device or another). The role of the digital twin is to return an approximation of the response of the sub-system (like the temperature and the debit of a fluid flow outside of a heat pump).

During the same phase, it is important to record data on the environment and the state of the system. They will serve as the basis to elaborate the clustered situations, as detailed in the next section. These situations include: meteorological state, economic environment and consumer needs inside the microgrid.

2.2. Clustering

Once digital twins have been trained accurately, and enough observations on the system environment have been gathered, it is possible to build a virtual system. This virtual version is confronted to a wide variety of situations it can encounter, based on the observations made: relying only on real data is an option, but there exist numerous methods of data augmentation allowing to present a larger set of situations to a clustering algorithm.

Then comes the clustering phase: to each situation is associated a preferentially small number of characteristics. Next, according to these characteristics, the situations are gathered into a reduced number of typical groups called clusters. These clusters will serve to identify the strategy to apply in a given situation.

2.3. Policy training

The policy training is quite simple: to each cluster previously created, the best strategy to adopt is identified. Thus, during the daily monitoring, the only role of the manager is to determine which cluster the current situation belongs to, and to apply the strategy associated to this cluster.

3. Case study

3.1. Description of the microgrid

As illustrated in Fig. 2, the studied system is a purely electrical microgrid composed of three elements: a 2 MW wind turbine, prosumers with roof PV and a system producing hydrogen (an electrolyser and a storage tank). The energy produced by the system has 3 usages. The absolute priority is to cover the consumption of the households. Then, the system has to deliver 30 kg of hydrogen each day in order to refuel a bus. The surplus of energy, if any, is sold to the national electrical grid on the intraday market. Therefore, it is possible to optimize the H₂ production in order to benefit from the variability of electricity prices. This installation is located in North-Western France (oceanic climate, quite windy, around 1 MWh/m²/year of solar energy).

It is important to note that in our case, the "real system" is in fact simulated with physical models and experimental data. This approach allows to test different hyper-parameters and to compare the trained strategy to the basic one.

Concerning data, we rely on 3 years of data (2019-2021). It is important to note that the same data-set is used both during the sizing and the policy training: then, the studied microgrid should offer little room for maneuver. Consumption data are extracted from the IHOGA software ([37]), electricity prices from EPEX spot and meteorological data from [38], a tool recalculating (and publishing freely) weather conditions worldwide using real data.

This case is very similar to the one investigated in a previous work. Then, to get further details about final cost of hydrogen, self-sufficiency of the microgrid or physical models (apart from the wind turbine, the electrolyser and the compressor), the reader is invited to consult [3]. The same stands for the private homemade Python software used for simulation.

Figure 2. Case illustration

3.2. Additional modeling information

The wind turbine, the electrolyser and the compressor being the only elements different from our previous study, they are fully presented here.

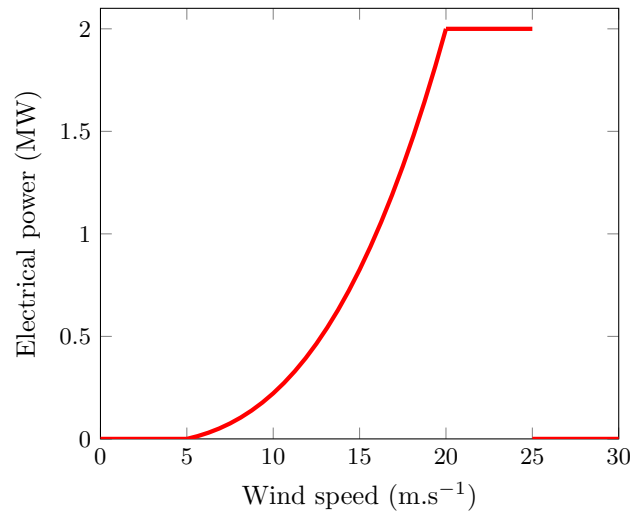


Figure 3. Wind turbine power curve

3.2.1. Wind turbine

The wind turbine is characterized by its nominal power $\mathcal{P}_{WT,nom}$ at the rated wind speed $U_{WT,nom}$, as well as the start-up ($U_{WT,min}$) and safety cut-off ($U_{WT,max}$) speeds. The relation giving the power as a function of the wind speed is taken from [39]:

$$\mathcal{P}_{WT} = \begin{cases} 0 & \text{if } U \leq U_{WT,min} \\ \mathcal{P}_{WT,nom} \frac{U^3 - U_{WT,min}^3}{U_{WT,nom}^3 - U_{WT,min}^3} & \text{if } U_{WT,min} < U \leq U_{WT,nom} \\ \mathcal{P}_{WT,nom} & \text{if } U_{WT,nom} < U \leq U_{WT,max} \\ 0 & \text{if } U \geq U_{WT,max} \end{cases} \quad (1a)$$

These equations lead to a power curve presented in Fig. 3.

3.2.2. Electrolyser

The electrolyzer is characterized by its efficiency η_{el} :

$$\eta_{el} = \frac{\dot{m}_{el} HHV}{\mathcal{P}_{el}} \quad (2)$$

where \dot{m}_{el} is the mass flow rate of produced hydrogen and \mathcal{P}_{el} the power consumption. In our previous work, the electrolyser was modeled with a static efficiency. In this work, we chose to use a dynamic value of efficiency relying on experimental data. The different values are presented on Fig. 4: efficiency rises quickly from 0 to 5% of relative power usage to reach 76% of efficiency. Then, efficiency decreases slowly to 65% when the electrolyser works at full power. We took data from [40] and interpolated it with the dedicated function of the Scipy Python library.

3.2.3. Compressor

In general, the H_2 pressure level at the electrolyzer outlet is not compatible with the pressure in the storage unit (which varies over time). The compression requires mechanical work which adds to the energy balance of the system. This additional consumption is calculated by applying an isentropic ratio (ratio between the specific work in the ideal-reversible case and in the real case) to the reversible adiabatic (i.e. isentropic) transformation from the pressure P_{el} (outlet of the electrolyser) to the pressure in the stock P_{st} :

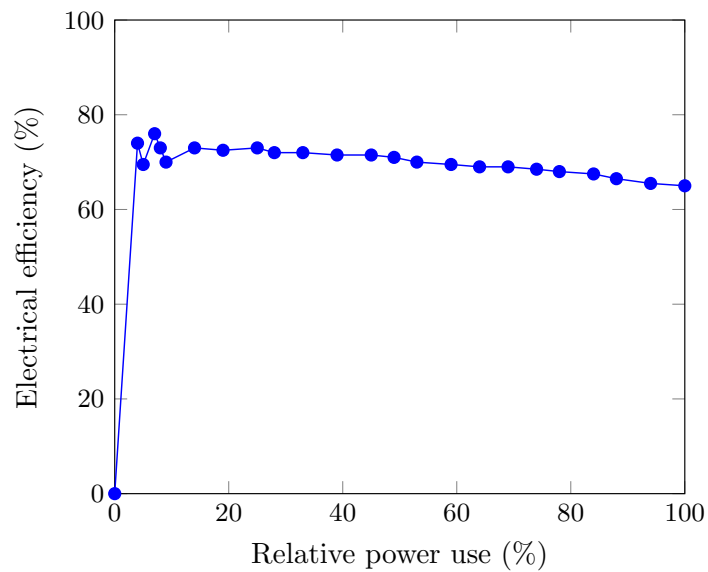


Figure 4. Experimental electrolyser efficiency curve

$$P_{comp} = \frac{\gamma}{\gamma - 1} \cdot R \cdot T \cdot \left[\left(\frac{P_{st}}{P_{el}} \right)^{\frac{\gamma-1}{\gamma}} - 1 \right] \cdot \frac{\dot{m}_{el}}{\hat{M}} \quad (3)$$

The evolution of the pressure P_{st} in the stock evolves dynamically as a function of hydrogen production and consumption. The model is described in [3].

3.3. Sizing of the system

The sizing of the system, realized with the Skopt Python library, has 2 steps: a simplex optimisation follows a bayesian one.

Bayesian optimization [41] use a prior belief about the objective function f_{obj} and updates the prior with samples drawn by an acquisition function that direct sampling in towards an area where an improvement over the current best observation is likely. The skopt package use by default for the acquisition function an expected improvement.

$$-EI(x) = -\mathbb{E}[f_{obj}(x) - f_{obj}(x_{best})] \quad (4)$$

Other acquisition functions are possible like the lower confidence bound and the probability of improvement. Bayesian optimisation also uses an surrogate model, often a Gaussian process, to approximate the objective function.

The Bayesian optimization quickly converge near the optimum but depending on its hyper-parameters, we found that the final steps can be hard. To increase the robustness of our approach, we limit the number of iterations of the Bayesian optimiser, wait until it is near the optimum, then run a simplex algorithm to reach convergence on the optimum.

The Nelder-Mead algorithm is a classical adaptive-size Simplex algorithm [42]: In a N -dimensions search domain, the $N + 1$ vertex of a simplex (*i.e* the generalization of a triangle to an arbitrary dimension) is deformed and moved to the algorithm try encircle the optimal location. The different transformations, affecting only one vertex at a time, can be found in the literature [43].

The final optimisation criterion, ϵ_{Tot} , must be minimized and is the combination of three criteria: ϵ_{houses} seeks to minimise the number of houses needed, $\epsilon_{service}$ inflicts a

heavy penalty if there is not enough H_2 to refuel the bus and ϵ_{Eco} measures the economic performance of the system. The H_2 storage tank is empty at the beginning of the simulation: otherwise, the sizing could lead to a solution with a huge storage tank without any H_2 production. Each criterion is multiplied by a coefficient C to balance the importance of each criterion, as exposed in Eq. 5d. Their value was determined in the previous study: C_{houses} value is 0.001, $C_{service}$ is 0.8 and C_{money} is 1.

$$\epsilon_{houses} = N_{houses} \cdot C_{houses} \quad (5a)$$

$$\epsilon_{service} = \frac{m_{H_2}^{lacking}}{m_{H_2}^{bus}} \cdot C_{service} \quad (5b)$$

$$\epsilon_{Eco} = \frac{M_{realized}}{M_{ref}} \cdot C_{money} \quad (5c)$$

$$\epsilon_{tot}^{sizing} = \epsilon_{houses} + \epsilon_{service} - \epsilon_{Eco} \quad (5d)$$

The sizing retained is the following:

- 90 houses, each one corresponding to 17 kW of solar power peak;
- an electrolyser of 325 kW;
- a tank accepting up to 1200 kg of hydrogen.

4. Method implementation

4.1. Digital twins fitting

Digital twins allow to create a virtual system which is the one used to identify improved monitoring strategies. Households consumption, electricity prices and meteorological were simulated with 2019 data. Thus, it was necessary to get three different digital twins:

- Energy produced: given the meteorological state (irradiation, wind speed, temperature, ambient pressure), it returns the renewable energy production.
- H_2 production: the H_2 production for the total electricity consumed (electrolyser and compressor)
- Electrolyser consumption: for a total electricity used to produce H_2 , this digital twin returns the consumption of the electrolyser branch. It is separated from the previous digital twin as it is needed to know which quantity to send in the electrolyser branch and in the compressor branch separately.

We used the Pygam Python library to fit the digital twins. This library is an implementation of Generalized Additive Models (GAMs) in Python. GAMs are kind of an extension of linear models: they try to build a model with more classical functions than polynomial ones. Pygam proposes splines, exponential terms and a bi-variate product. It is able to seek a function by itself but the user can indicate the form of the function to earn time or to improve accuracy. Meanwhile, an important caveat with GAMs is that they tend to overfit, *i.e* there is no guarantee of precision on predictions made outside of the space they already explored.

Another choice made here is to limit the data taken into account to one year, for two reasons: it reduces the computation time and it allows to follow the aging of the system.

4.2. Control strategies

Control strategies are the algorithms applied to manage the energy available for distribution. The strategy used by the basic policy, detailed in Alg. 1, has the following priorities: first, self-consumption; second, H_2 delivery and, third, selling energy to the grid. The basic policy has a robust approach, as it tends to fill as soon as possible the H_2 tank to secure the bus refueling. Unfortunately, this policy does not exploit the variations of prices. In order to keep it simple and understandable, the trained strategy will only have the choice

between two algorithms: the one previously presented and another one, Alg. 2, where all energy is sold to the grid. The idea is that, whenever it is relevant, the trained policy should be able to profit from the prices variability.

Algorithm 1 Storing strategy

```

 $\Delta E \leftarrow \mathcal{C} - \mathcal{P}r$ 
if  $\Delta E > 0$  then
   $\Delta E \leftarrow \max(\Delta E - \mathcal{P}_{max}^{H_2} \cdot \Delta t, 0)$ 
end if
 $p_{exchanges} \leftarrow \Delta E$ 
  
```

Algorithm 2 Selling strategy

```

 $\Delta E \leftarrow \mathcal{C} - \mathcal{P}r$ 
if  $\Delta E > 0$  then
  Do Nothing
end if
 $p_{exchanges} \leftarrow \Delta E$ 
  
```

4.3. Clustering

In this part are defined three items: the elaboration of the situations the system could encounter, the criteria used for clustering and the clustering method.

In our case, the situations encountered are a mix of weather, households consumption, electricity prices and storage level. We had the following constraints: limit the observation time needed to apply the methodology and explore numerous but probable situations. Consequently, we retained the following method:

1. Observe the system during one full year, as it can be considered as a complete cycle for meteorological and consumption variations.
Concerning the prices, they have been quite unstable in Europe over the last year and it would be far beyond the scope of this article to make assumptions on their evolution. Thus, the one-year-long observation is considered good enough, as it contains at least the daily and seasonal variations.
2. Replicate this single year of observation several times.
3. Apply a gaussian noise to the several years long sequence.

This way allows to increase the variety of situations around the one already encountered while avoiding irrelevant situations, like coinciding consumption and PV production peaks (usually located respectively during winter evenings and summer afternoons). In our case, we took data corresponding to the year 2019.

Once this long sequence is defined, we need to cut in little sequences. These little sequences are the ones gathered in clusters. Each sequence is described by 3 parameters:

- Energy available: the quantity of energy produced minus the consumption, *i.e.* the energy to be distributed between H_2 production and electricity selling.
- Electricity price: the average electricity price over the sequence.
- H_2 mass stored: the quantity of H_2 available at the beginning of the session.

Last, the chosen clustering method is a classical least square k-means from the SciPy python library. Once the different clusters are obtained, a difficulty persists: each cluster represents a sequence, *i.e.* several consecutive values of available energy and prices while their centers are only scalar values. Thus, we cannot use directly clusters centers to train the policy. Instead, for each cluster, we use the sequence being the closest one to each center: the distance is obviously calculated with the least square, as it is the norm used for clustering.

4.4. Policy training

For each cluster, we assess the performance of the two strategies defined in section 4.2 using the criterion defined in Eq. 6 and we keep the best one. The evaluation is based on the financial gain and the production of hydrogen. The parameter C_{H_2} sets the relative importance of H_2 production compared to financial gain: the higher, the more H_2 production is favoured. This approach is made possible because of the limited number of strategies. In a more complex case, optimization tools could have been used to identify the relevant strategy.

$$\epsilon_{Eco} = \frac{M_{realized}}{M_{ref}} \quad (6a)$$

$$\epsilon_{H_2} = \frac{\Delta m_{H_2}^{tank}}{m_{H_2}^{bus}} \cdot \frac{1}{24} \quad (6b)$$

$$\epsilon_{tot}^{strategy} = \epsilon_{Eco} + \epsilon_{H_2} \cdot C_{H_2} \quad (6c)$$

4.5. Performance assessment

Performance assessment consists in a 2-year-long simulation relying on 2020 and 2021 data applying a trained policy and assessing its score according to two metrics. As stated before, the main goal of the system is to produce enough hydrogen to the daily bus recharge. Thus, an heavy penalty is inflicted when hydrogen is lacking. The secondary objective is to maximise financial earnings, by selling electricity to the grid at the best moments. As the raw variation of money flows is of little sense, the achievable financial gain is estimated relative to a defined baseline using the following procedure:

1. Calculate the money earned if no H_2 was produced and all electricity was sold to the grid
2. Sort the simulation rounds by increasing electricity prices
3. Calculate the total energy needed to produce H_2
4. Deduce the minimum loss, *i.e.* the minimum money not sold to the grid to the production of H_2
5. Calculate the theoretical best money flow by subtracting the money loss due to H_2 production from the money earned when no H_2 is produced
6. The over-performance is thus equal to the difference between the theoretical best money flow and the reference run money flow

This method ensures that the needed mass of hydrogen is produced overall but does not take into account nor the daily delivery constraint nor the maximum storage limit: . Consequently, as this gain may not be achievable, it's not the maximum but only an upper bound of the best gain. In our case, this estimation is equal to 7k€ on 2 years.

4.6. Simulation plan

An important question in this work is the influence of parameters chosen for the process. Three main parameters have been identified:

- Sequence length: the length of the sequences on which clustering is made
- Standard deviation: the standard deviation used in the gaussian noise applied to the clustering sequence
- Relative weight of H_2 production: the relative importance of H_2 production when identifying the best strategy

Tab. 1 presents the different values assessed for each parameter, for a total of 96 runs. Concerning sequence length, there are 2 things to know: the software uses 1-hour long time step and prices variations tend to follow a daily pattern. Thus, 2 hours appears as the lowest sequence to consider (as it has only 2 steps) while 24 hours represents the maximum length for financial gain optimization.

The standard deviation set here is the one applied to environment variables (weather, prices and consumption) during clustering. When it increases, situations used for clustering will be more diverse but also more distant from the observed data. On the one hand, too low a value limits the policy capacity to adapt to new environments. On the other hand, too high a value leads to create a lot of situations the system will not encounter and thus dilute the useful information. It ranges from 0, the absolute lower bound to 0.3, a value sufficient to observe the dilution effect.

Last, the relative weight of H₂ production ranges from 0 (where H₂ production does not count in the strategy assessment) to 0.75, where strategy always prefer to produce H₂.

Table 1. Values explored in the sensitivity analysis

Sequence length (h)	Standard deviation (-)	Relative weight of H ₂ production (-)
2, 3, 6, 12, 18, 24	0, 0.1, 0.2, 0.3	0, 0.25, 0.5, 0.75

5. Results and discussion

5.1. Digital twins fitting

The first step of our approach is to create efficient digital twins. In Fig. 5 is presented the digital twins predictions compared to the realized values. In this setting, the digital twin model is re-evaluated each day.

We observed that for energy production, 2 months were enough to reach a satisfying accuracy, as illustrated by Fig. 5c. Meanwhile, as it can be seen around 1000h (approximately 7 weeks), GAMs only give satisfactory results in already encountered weather contexts.

Concerning the H₂ production cost and the electrolyser consumption, 1 week of training (168h) seems to be enough: unlike irradiation or wind, the variables are bounded by the equipment limits (max power for the electrolyser and storage capacity for the H₂ tank). Then, the usage of the electrolyser at its full power during the first day allows the digital twins to rapidly converge on a high level of precision.

5.2. Results of policy training

As explained before, the goal of this work is to evaluate the efficiency of the proposed method and its robustness to parameters settings. Two aspects are analysed: the financial gain and the potential default of H₂ delivery.

5.2.1. Overall performance

At first glance, in 21 cases out of 96, H₂ is not fully delivered. On the financial side, in 8 runs, the trained policy is more than 5% worse compared to the estimated best over-performance. This means that in these 29 runs (30%), the trained policy degrades the microgrid performance. As observable on Fig. 6a and 6b, there is no run cumulating undelivered H₂ and poor financial performance. This phenomenon comes from the nature of strategies proposed to the policy: either they store too much and thus degrade their financial gain, either they sell too much and do not produce enough H₂.

More worryingly, in 64 runs (67%), the financial difference is lower than 5% variation, meaning that the training has no sensible effect. These observations made, 3 runs out of 96 improve significantly the performance of the microgrid, by improving its gain by 19, 16 and 8% of the best performance.

While the proportion of successful runs is very low, the proposed method is still able to improve the system efficiency: for reminder, the method relies on a statistical approach based itself on an approximation of the microgrid behavior. This result is even more interesting if we consider the fact that the system had little room to maneuver. The system operates exactly in the environment its sizing was optimized on, which is quite unlikely for a real system, as industrial systems are often oversized to ensure a minimal performance and enhance reliability.

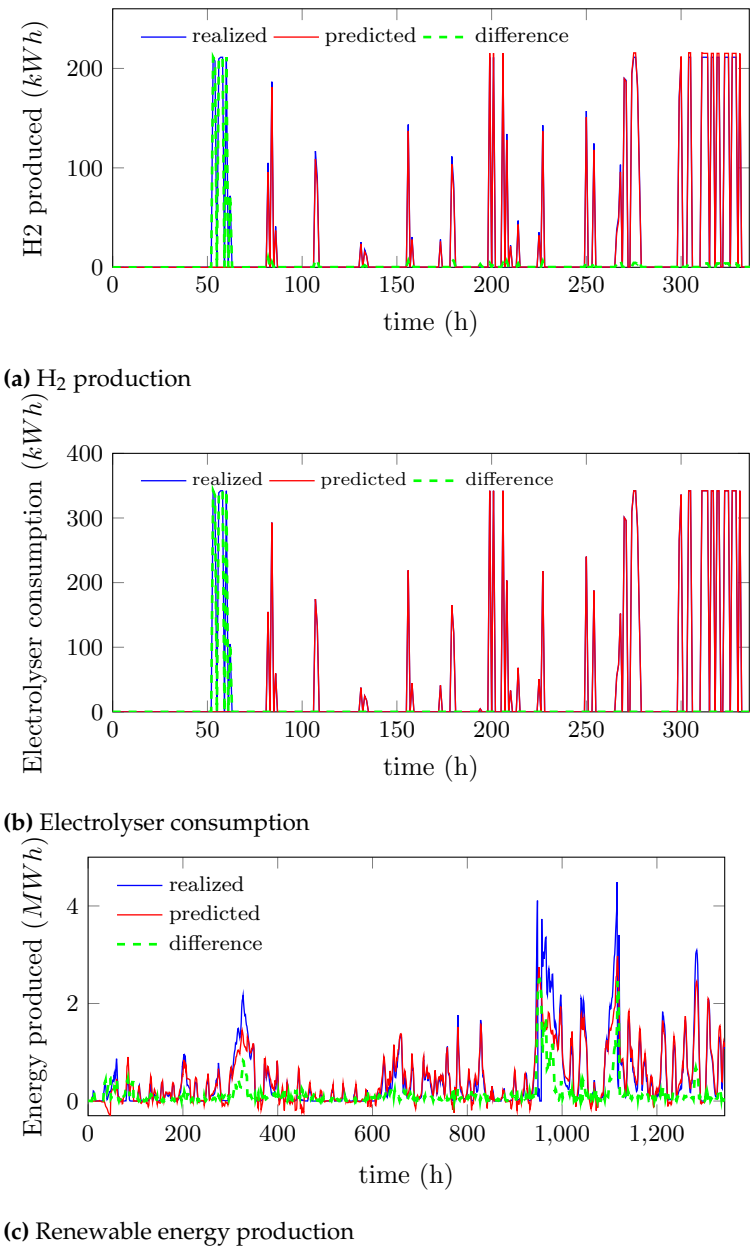
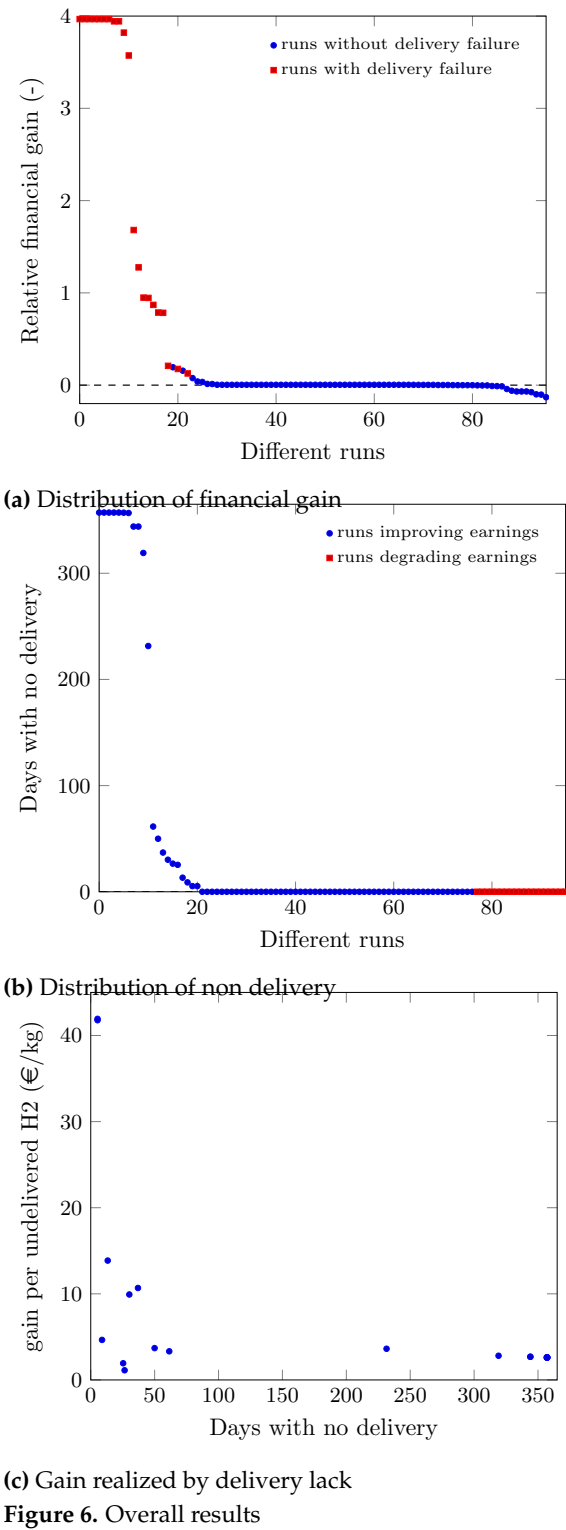


Figure 5. Comparison of values realized and predicted by the digital twins



(c) Gain realized by delivery lack
Figure 6. Overall results

At second glance, knowing that the renewable H₂ price revolves around 3-8\$/kg [44,45], it can be interesting to accept punctual delivery default: in some cases, the financial gain reaches up to 42 €/kg as shown on Fig. 6c, *i.e* at least 5 times H₂ price. Moreover, it can be seen that these high values are found for relatively low values of undelivered H₂ (around 5 days/year of delivery failure). Consequently, it is very possible that the microgrid manager finds that these policies are more interesting than the basic one: they could buy H₂ by other means or use electricity bought from the grid to produce a bit of H₂ (while keeping in mind requirements for "green" H₂ labels) to benefit from these policies advantages without their disadvantages.

5.2.2. Sensitivity to parameters

Table 2. Sensitivity to sequence length

Sequence length (h)	2	3	6	12	18	24
Financial gain (k€)	1.7	2.6	1.8	5.4	4.1	6.8
Non-delivery (days/year)	3	26	23	67	52	79
Gain per undelivered H ₂ (€/kg)	17.0	3.3	2.6	2.7	2.7	2.9

First, the impact on sequence length can seem to be unclear regarding the performance of the trained policies, as it can be seen in Tab. 2: for the 3 metrics, the evolution is not monotonous. Meanwhile, it is possible to distribute the simulation results in 3 sequence groups (separated by dashed lines): 2h; 3 and 6h; 12, 18 and 24h. With this grouping, it is possible to identify a tendency: the shorter the better. Beginning with bus refueling, shorter sequences lead to safer strategies, as undelivered H₂ quantity goes from 3 days/year to 79. If raw financial gain is globally increasing with sequence length, the gain per undelivered H₂ mass is far higher for 2h-long sequences than for the others (17 €/kg compared to approximately 3). As explained in the previous section, from a microgrid manager point of view, not delivering all the H₂ is interesting only if it can earn more than by selling H₂: as green H₂ prices are around 3-8\$/kg, earning 3€ instead of delivering H₂ is a nonsense.

Table 3. Sensitivity to standard deviation

Standard deviation	0	0.1	0.2	0.3
Financial gain (k€)	6.7	1.4	1.6	5.2
Non-delivery (days/year)	77	13	18	59
Gain per undelivered H ₂ (€/kg)	2.9	3.8	3.0	2.9

Second, as observed in Tab. 3, standard deviation can have a positive effect on performance but has to be used with caution. With 10 and 20% of standard deviation, undelivered H₂ reduce significantly: from 77 days/year with 0 standard deviation to 13 and 18 respectively. Meanwhile, if the financial gain seems lesser (from 6.7 k€ to 1.4 and 1.6), the relative gain per undelivered H₂ is stable (2.9 €/kg compared to 3.8 and 3.0). Thus, they globally improve the efficiency of trained policies. When standard deviation reaches 30%, however, performance decreases. If financial performance is better than for 10 and 20% standard deviation (5.2 k€), the marginal gain per undelivered H₂ is still stable (2.9 €/kg). Undelivered H₂, meanwhile, increases to 59 days/year, far more than the 10-20 days observed for lower standard deviations. In fact, in this context, we observe that the situations used for clustering are too far from the ones encountered. Thus, clusters cover too different situations and it is not possible anymore to assign a single strategy efficient for all situations. Consequently, using data augmentation methods, at least in the studied case, can improve policies performance as long as it does not hamper strategy identification for clusters.

Table 4. Sensitivity to stock coefficient

Stock coefficient	0	0.25	0.5	0.75
Financial gain (k€)	13.1	1.8	0.01	0
Non-delivery (days/year)	151	16	0	0
Gain per undelivered H ₂ (€/kg)	2.9	3.7	N/A	N/A

The stock coefficient has the clearer impact on policies performance. When stock coefficient is equal to 0, it means that H₂ production is not taken into account. Unsurprisingly, it results in a very high quantity of undelivered H₂, 151 days/year, as the policy tends to sell all electricity to the grid. On the other side of the specter, for coefficients of 0.5 and 0.75, all H₂ is delivered but no financial gain is made: the trained policy produces as much H₂ as possible and imitates the basic policy. Consequently, 0.25 seems to be the best value as it avoids these 2 extremes.

Conclusion

In this paper we have developed a methodology to evaluate and choose a best policy regarding economic performance, product delivery and the evolution of the system degradation over time. While the retained approach is not reinforcement learning, it is a proxy for it or an explainable intermediary step to it, as it defines its policies with the current state of the system and its updated estimated performance.

We have shown the existence of policies whose definition is not straightforward, but prove to be economically interesting for the micro-grid operator. It remains true that the identification of their optimal parameters does not obey to any obvious rule. In some cases, it is interesting to not always deliver the promised fuel, but instead to support the grid for financial gains. More than policies, this means that the methodology applied in this study is able to suggest shifts in the very business model of the micro-grid manager. Otherwise, this information can be used to determine the penalty to apply for non-delivery.

The usage of GAMs model as numerical-twins is promising as it is fairly easy to deploy (few to none hyperparameters) and has shown its ability to quickly gain a satisfactory accuracy.

Depending on the use case, we have shown that it is possible to obtain a performing strategy allowing, with a system sized around a nominal case, to obtain increased performances in real non-nominal conditions. A remaining difficulty is that, in the current state of our investigations, a prior definition of the parameters set of the method still seems out of reach. However, these results clearly confirm our conviction, namely that it is possible to improve the use of an existing system in an optimal way (according to a user-defined criterion) under conditions different from those used for its design.

Author Contributions:

Funding acquisition: Stéphane Gibout and Christophe Geissler

Investigation: Timothé Gronier, William Maréchal

Project Administration: Christophe Geissler

Software: Timothé Gronier

Supervision: Stéphane Gibout

Visualization: Timothé Gronier

Writing: Timothé Gronier, William Maréchal and Stéphane Gibout

All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by France Relance, especially through their participation to the funding of the post-doctoral stage of Timothé Gronier.

Acknowledgments: We thank Dr. S. Chabab, from the *Université de Pau et des Pays de l'Adour*, for his help in the choice of the equation of state used for hydrogen, Dr. R. Dufo-Lopez, from the *University of Zaragoza*, for letting us use load data extracted from his iHOGA software and the team of Pr. D. Hissel, from the *University Bourgogne France-Comté*, for the re-use of an experimental curve describing the electrolyser efficiency.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Nomenclature

Latin symbols		Greek symbols	
A	surface area, m ²	α	ground roughness factor, –
C	consumption, kWh	η	efficiency, – or %
C_P	power coefficient, –	ρ	density, kg m ⁻³
c	specific heat capacity, J K ⁻¹ kg ⁻¹	τ	time constant, s
E	total energy, J or kWh	ϵ	optimisation criterion, –
h	specific enthalpy, J kg ⁻¹	Subscripts and superscripts	
l	length of a sequence in clustering, h	a	air
LHV	lower heating value, J m ⁻³ or J kg ⁻¹	amb	ambient
m	mass, kg	bot	bottom
M	money flow, €	bui	building
\hat{M}	molar mass, kg mol ⁻¹	ch	charge
OPEX	operational expenditures, €	cut	cutoff
P	pressure, Pa	dis	discharge
\mathcal{P}	power, W	el	electrical
$\mathcal{P}r$	production, kWh	Q	heat
P	price, €	ht	heating
Q	heat, J or kWh	inv	inverter
q	flow-rate, m ³ s ⁻¹ or L s ⁻¹	nom	nominal
R	individual ideal gas constant, J K ⁻¹ kg ⁻¹	out	outdoor
SOC	state of charge, %	pan	panel
T	temperature, K or °C	ref	reference
t	time, s or min		
V	volume, m ³		

Abbreviations

The following abbreviations are used in this manuscript:

DT	digital twin
EMS	energy management system
EPEX	european power exchange
GAMs	generalized additive models
ML	machine learning
PV	photo-voltaic
WT	wind turbine

References

1. Mitali, J.; Dhinakaran, S.; Mohamad, A. Energy Storage Systems: A Review. *Energy Storage and Saving* **2022**, *1*, 166–216. doi:10.1016/j.enss.2022.07.002.

2. He, W.; King, M.; Luo, X.; Dooner, M.; Li, D.; Wang, J. Technologies and Economics of Electric Energy Storages in Power Systems: Review and Perspective. *Advances in Applied Energy* **2021**, *4*, 100060. doi:10.1016/j.adapen.2021.100060.

3. Gronier, T.; Maréchal, W.; Gibout, S.; Geissler, C. Relevance of Optimized Low-Scale Green H2 Systems in a French Context: Two Case Studies. *Energies* **2022**, *15*. doi:10.3390/en15103731.

4. Petrozziello, A.; Troiano, L.; Serra, A.; Jordanov, I.; Storti, G.; Tagliaferri, R.; La Rocca, M. Deep learning for volatility forecasting in asset management. *Soft Computing* **2022**, *26*, 8553–8574.

5. Christensen, K.; Siggaard, M.; Veliyev, B. A machine learning approach to volatility forecasting. *Available at SSRN* **2021**.

6. Sadaqat, A.; Zheng, Z.; Aillerie, M.; Péra, M.C.; Hissel, D. A Review of DC Microgrid Energy Management Systems Dedicated to Residential Applications. *MDPI energies* **2021**.

7. Sang, J.; Sun, H.; Kou, L. Deep Reinforcement Learning Microgrid Optimization Strategy Considering Priority Flexible Demand Side. *Sensors* **2022**, *22*. doi:10.3390/s22062256.

8. Nyong-Bassey, B.E. A Concise Review of Energy Management Strategies for Hybrid Energy Storage Systems. *European Journal of Engineering and Technology Research* **2022**, *7*. doi:10.24018/ejeng.2022.7.3.2815.

9. Arwa, E.O.; Folly, K.A. Reinforcement Learning Techniques for Optimal Power Control in Grid-Connected Microgrids: A Comprehensive Review. *IEEE Access* **2020**, *8*, 208992–209007. doi:10.1109/ACCESS.2020.3038735.

10. Rasheed, A.; San, O.; Kvamsdal, T. Digital Twin: Values, Challenges and Enablers From a Modeling Perspective. *IEEE Access* **2020**, *8*, 21980–22012. doi:10.1109/ACCESS.2020.2970143.

11. Chinesta, F.; Cueto, E.; Abisset-Chavanne, E.; Duval, J.L.; Khaldi, F.E. Virtual, digital and hybrid twins: a new paradigm in data-based engineering and engineered data. *Archives of computational methods in engineering* **2020**, *27*, 105–134.
12. Tao, F.; Xiao, B.; Qi, Q.; Cheng, J.; Ji, P. Digital twin modeling. *Journal of Manufacturing Systems* **2022**, *64*, 372–389.
13. Errandonea, I.; Beltrán, S.; Arrizabalaga, S. Digital Twin for maintenance: A literature review. *Computers in Industry* **2020**, *123*, 103316.
14. Dembski, F.; Wössner, U.; Yamu, C. Digital twin. Virtual Reality and Space Syntax: Civic Engagement and Decision Support for Smart, Sustainable Cities: Proceedings of the 12th International Space Syntax Conference, Beijing, China, 2019, pp. 8–13.
15. Singh, M.; Fuenmayor, E.; Hinchy, E.P.; Qiao, Y.; Murray, N.; Devine, D. Digital twin: Origin to future. *Applied System Innovation* **2021**, *4*, 36.
16. Qi, Q.; Tao, F.; Hu, T.; Anwer, N.; Liu, A.; Wei, Y.; Wang, L.; Nee, A. Enabling technologies and tools for digital twin. *Journal of Manufacturing Systems* **2021**, *58*, 3–21.
17. Zheng, Y.; Yang, S.; Cheng, H. An application framework of digital twin and its case study. *Journal of Ambient Intelligence and Humanized Computing* **2019**, *10*, 1141–1153.
18. Lu, H.; Du, M.; Qian, K.; He, X.; Wang, K. GAN-based data augmentation strategy for sensor anomaly detection in industrial robots. *IEEE Sensors Journal* **2021**.
19. Gan, J.; Smith, C. Drivers for renewable energy: A comparison among OECD countries. *Biomass and Bioenergy* **2011**, *35*, 4497–4503.
20. Lu, J.; Yi, S. Autoencoding Conditional GAN for Portfolio Allocation Diversification. *arXiv preprint arXiv:2207.05701* **2022**.
21. Xu, Y.; Sun, Y.; Liu, X.; Zheng, Y. A digital-twin-assisted fault diagnosis using deep transfer learning. *Ieee Access* **2019**, *7*, 19990–19999.
22. Maschler, B.; Braun, D.; Jazdi, N.; Weyrich, M. Transfer learning as an enabler of the intelligent digital twin. *Procedia CIRP* **2021**, *100*, 127–132.
23. Xia, M.; Shao, H.; Williams, D.; Lu, S.; Shu, L.; de Silva, C.W. Intelligent fault diagnosis of machinery using digital twin-assisted deep transfer learning. *Reliability Engineering & System Safety* **2021**, *215*, 107938.
24. Deebak, B.; Al-Turjman, F. Digital-twin assisted: Fault diagnosis using deep transfer learning for machining tool condition. *International Journal of Intelligent Systems* **2021**.
25. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *Journal of Manufacturing Systems* **2021**, *58*, 346–361.
26. Naser, M. Digital twin for next gen concretes: On-demand tuning of vulnerable mixtures through Explainable and Anomalous Machine Learning. *Cement and Concrete Composites* **2022**, *132*, 104640.
27. Jamal, S.; Tan, N.M.L.; Pasupuleti, J. A Review of Energy Management and Power Management Systems for Microgrid and Nanogrid Applications. *Sustainability* **2021**, *13*. doi:10.3390/su131810331.
28. Battula, A.R.; Vuddanti, S.; Salkuti, S.R. Review of Energy Management System Approaches in Microgrids. *Energies* **2021**, *14*. doi:10.3390/en14175459.
29. Boukas, I.; Ernst, D.; Théate, T.; Bolland, A.; Alexandre, H.; Martin, B.; Wynants, C.; Cornélusse, B. A deep reinforcement learning framework for continuous intraday market bidding. *Machine Learning* **2021**, p. 2335–2387. doi:10.1007/s10994-021-06020-8.
30. Aittahar, S.; Manuel de Villena Millan, M.; Derval, G.; Castronovo, M.; Boukas, I.; Gemine, Q.; Ernst, D. Optimal Control of Renewable Energy Communities with Controllable Assets, 2022.
31. Bolland, A.; Boukas, I.; Berger, M.; Ernst, D. ointly Learning Environments and Control Policies with Projected Stochastic Gradient Ascent. *Journal of Artificial Intelligence Research* **2022**, pp. 117–171. doi:arXiv:2006.01738.
32. Henry, R.; Ernst, D. Gym-ANM: Reinforcement learning environments for active network management tasks in electricity distribution systems. *Energy and AI* **2021**, *5*, 100092. doi:https://doi.org/10.1016/j.egyai.2021.100092.
33. Yang, Y.; Li, H.; Shen, B.; Pei, W.; Peng, D. Microgrid Energy Management Strategy Base on UCB-A3C Learning. *Frontiers in Energy Research* **2022**, *10*. doi:10.3389/fenrg.2022.858895.
34. Munos, R.; Stepletn, T.; Harutyunyan, A.; Bellemare, M.G. Safe and efficient off-policy reinforcement learning. *NIPS 2016*, , 2016.
35. Sutton, R.S.; Barto, A.G. Reinforcement Learning: An Introduction, 2015.
36. Rodrigues, C.; Gérard, P.; Rouveirol, C. On and Off-Policy Relational Reinforcement Learning, 2008.
37. Dufo-López, R.; Bernal-Agustín, J.L.; Yusta-Loyo, J.M.; Domínguez-Navarro, J.A.; Ramírez-Rosado, I.J.; Lujano, J.; Aso, I. Multi-Objective Optimization Minimizing Cost and Life Cycle Emissions of Stand-Alone PV–Wind–Diesel Systems with Batteries Storage. *Applied Energy* **2011**, *88*, 4033–4041. doi:10.1016/j.apenergy.2011.04.019.
38. Gelaro, R.; McCarty, W.; Suárez, M.J.; Todling, R.; Molod, A.; Takacs, L.; Randles, C.A.; Darmenov, A.; Bosilovich, M.G.; Reichle, R.; et al. The Modern-Era Retrospective Analysis for Research and Applications, Version 2 (MERRA-2). *Journal of Climate* **2017**, *30*, 5419 – 5454. doi:10.1175/JCLI-D-16-0758.1.
39. Burton, T.; Jenkins, N.; Sharpe, D.; Bossanyi, E. *Wind Energy Handbook*, second ed.; John Wiley & Sons, Ltd., 2011.
40. Yue, M.; Lambert, H.; Pahon, E.; Roche, R.; Jemei, S.; Hissel, D. Hydrogen energy systems: A critical review of technologies, applications, trends and challenges. *Renewable and Sustainable Energy Reviews* **2021**, *146*, 111180. doi:https://doi.org/10.1016/j.rser.2021.111180.
41. Pelikan, M. Bayesian optimization algorithm. In *Hierarchical Bayesian optimization algorithm*; Springer, 2005; pp. 31–48.
42. Maréchal, W. Utilisation de méthodes inverses pour la caractérisation de matériaux à changement de phase (MCP). PhD thesis, 2014. Thèse de Doctorat de l’université de Pau et des Pays de l’Adour encadrée par Pr. Jean-Pierre Dumas, Président du jury : Pr. Laurent Zalewski.

43. Walters, F.H., Ed. *Sequential Simplex Optimization: A Technique for Improving Quality and Productivity in Research, Development, and Manufacturing*; Chemometrics Series, CRC Press: Boca Raton, Fla, 1991.
44. The Future of Hydrogen. Technical report, International Energy Agency, 2019.
45. IRENA. Hydrogen: A renewable energy perspective. Technical report, 2019.