

## Article

# Usage of GAMS-Based Digital Twins and Clustering to Improve Energetic Systems Control

Timothé Gronier <sup>1,2,3</sup> , William Maréchal <sup>1,\*</sup>, Christophe Geissler <sup>1</sup> and Stéphane Gibout <sup>2,\*</sup> <sup>1</sup> Advestis, 75008 Paris, France<sup>2</sup> Université de Pau et des Pays de l'Adour, E2S UPPA, LaTEP, 64053 Pau, France<sup>3</sup> ADERA, 33608 Pessac, France

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

**Abstract:** With the 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. However, 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 however is 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. In some cases, they even show that it is worth punctually missing the H<sub>2</sub> production requirement in exchange for higher economic performance. However, they require fine-tuning as to not deteriorate the system performance.

**Keywords:** energy management system; digital twins; general additive models; green H<sub>2</sub>



**Citation:** Gronier, T.; Maréchal, W.; Geissler C.; Gibout, S. Usage of GAMS-Based Digital Twins and Clustering to Improve Energetic Systems Control. *Energies* **2023**, *16*, 123. <https://doi.org/10.3390/en16010123>

Academic Editors: Alfeu J. Sguarezi Filho, Jen-Hao Teng, Lakshmanan Padmavathi and Kin-Cheong Sou

Received: 29 October 2022

Revised: 29 November 2022

Accepted: 17 December 2022

Published: 22 December 2022



**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. Introduction

With the 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]. Unfortunately, the intermittent nature of most renewable energy sources makes it difficult for the residential consumer to self-consume it. Furthermore, small-scale energy storage is still an expensive investment in 2022 and the risk/reward ratio is not yet attractive enough for individual homeowners. The solution studied here dedicates excess energy to a collective use, which in this case is the daily filling of a hydrogen bus. This usage fits into the European plan for hydrogen development [3] and some hydrogen buses are already in service in Europe [4]. Following our previous work on this case study [5], 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 moving down to individual systems, such as individuals or groups of dwellings, or solar community. In a previous article [5] we studied the opportunity of integrating energy storage technologies for such scales. In this article we will focus on the generation of H<sub>2</sub> 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 ensure that  $H_2$  is delivered. Machine learning has shown its ability to take advantage of both volatile and exogenous information [6,7], the kind of which is increasingly common for energy markets and for the behavior of small systems.

### 1.2. Literature Review

Numerous methods have been developed to build EMS systems and to monitor them [8–10]. 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 [8]. Their development has given way to the rise of Deep Reinforcement Learning [11]. 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 (DTs) are a technology that has recently been brought to the front pages [12] of business deciders and is presented as a path to less waste, shorter times to market, and constant customer insights. In the literature [13–20], 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 [14]. Accurate models are not 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) [21–23] and/or transfer learning [24–28], but the technology is still maturing. For this study, Generalized 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 smoothing functions [29].

### 1.3. Contributions to Novelty

Several approaches are being developed for energy management systems [8–11,30,31]. The latest one, based on deep reinforced learning, while producing results and being developed in open-sourced frameworks [32–35], is considered a black-box model [36] 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 [37]. 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 [38,39].

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 used to define a policy. The policy is then applied throughout the runs of the simulation, but as the operating conditions—either exogenous (prices, weather conditions) or endogenous (equipment efficiency, users' behavior)—evolve over time, the digital twins need to be re-evaluated. This re-evaluation in turn leads 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, which have

potentially made the chosen actions non-optimal. In this work, the presented approach appears as a proxy for reinforcement learning.

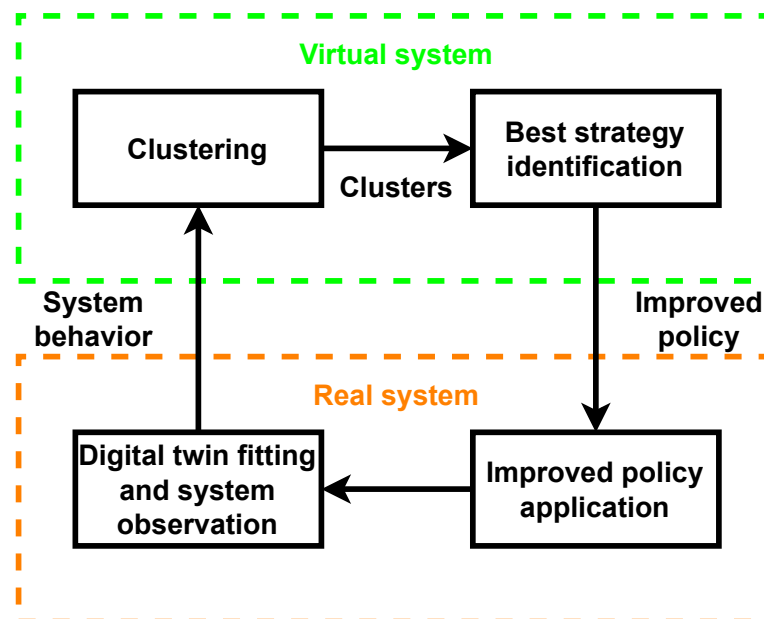
This approach is innovative as it uses techniques known by domain experts to approximate reinforcement learning methods, and can thus serve as a bridge between the two domains. This approach can serve as a benchmark against which Deep Learning methods can be compared in a way that can be interpreted 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 (Section 2). We will then present the chosen case study (Section 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 (Section 5) before drawing a conclusion.

## 2. Methodology

As explained 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.

The general methodology is shown in Figure 1. It is divided into four steps that allow for 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.



**Figure 1.** Proposed methodology.

These steps are repeated at intervals for two reasons. First, the systems themselves tend to evolve through time (typically because of their performances decline) and thus the 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 (such as 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 by the 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 that allow to present a larger set of situations to a clustering algorithm.

Then comes the clustering phase: each situation is associated with 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 previously created cluster, 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 Figure 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). This configuration, in the French context, can take the form of a “renewable energy community”, an entity authorized to consume, store, or sell the renewable energy it produces and to share it freely among its members [40]. The energy produced by the system has three 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 H<sub>2</sub> 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/m<sup>2</sup>/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 hyperparameters and to compare the trained strategy to the basic one.

Concerning data, we rely on three years of data (2019–2021). It is important to note that the same dataset is used both during the sizing and the policy training: the studied microgrid should therefore offer little room for maneuver. Consumption data are extracted from the IHOGA software [41], electricity prices from EPEX spot, and meteorological data from Ref. [42], using a tool recalculating (and publishing freely) weather conditions worldwide using real data.

This case is very similar to the one investigated in a previous study. Then, to get further details about the final cost of hydrogen and the self-sufficiency of the microgrid or physical models (apart from the wind turbine, the electrolyser, and the compressor); the interested reader can refer to Ref. [5]. The same stands for the private homemade Python software used for simulation.

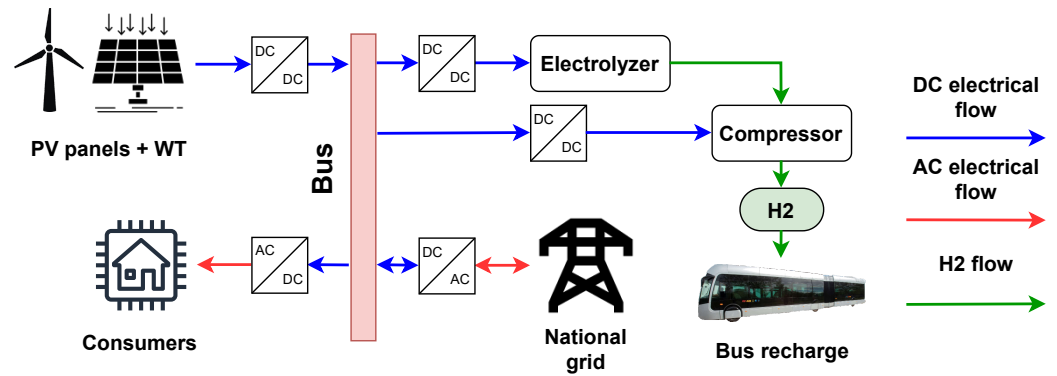


Figure 2. Case illustration.

### 3.2. Additional Modeling Information

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

#### 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 Ref. [43]:

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

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

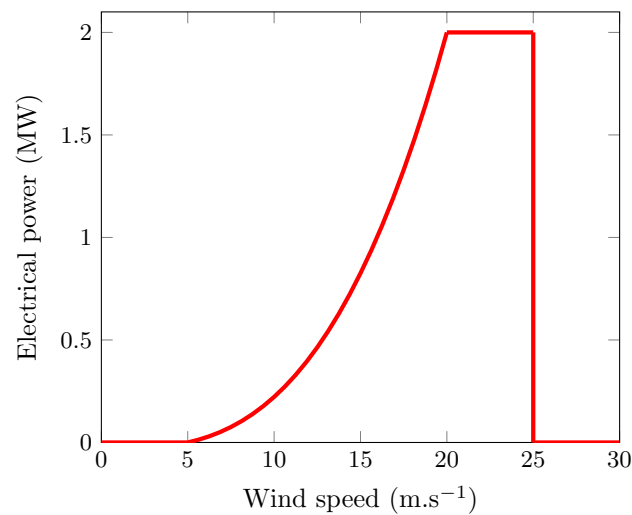


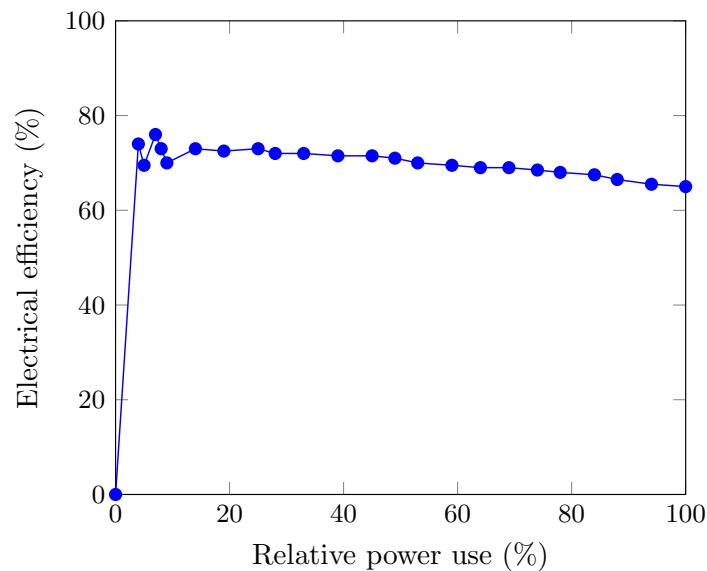
Figure 3. Wind turbine power curve.

#### 3.2.2. Electrolyser

The electrolyser is characterized by its efficiency  $\eta_{el}$ :

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

where  $\dot{m}_{el}$  is the mass flow rate of produced hydrogen and  $P_{el}$  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 in Figure 4: the efficiency rises quickly from 0 to 5% of relative power usage to reach 76% of efficiency. Then, the efficiency decreases slowly to 65% when the electrolyser works at full power. We took data from Ref. [44] and interpolated it with the dedicated function of the Scipy Python library.



**Figure 4.** Experimental electrolyser efficiency curve.

### 3.2.3. Compressor

In general, the  $H_2$  pressure level at the electrolyser 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 (the 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  $P_{st}$ :

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

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 Ref. [5].

### 3.3. Sizing of the System

The sizing of the system, realized with the Skopt Python library, has two steps: a simplex optimization that is followed by a Bayesian optimization.

Bayesian optimization [45] uses a prior belief about the objective function  $f_{obj}$  and updates the prior with samples drawn by an acquisition function with direct sampling in towards an area where an improvement over the current best observation is likely. The skopt package is used 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, such as the lower confidence bound and the probability of improvement. Bayesian optimization also uses an surrogate model, often a Gaussian process, to approximate the objective function.

The Bayesian optimization quickly converges 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 optimizer, 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 [46]: 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 to encircle the optimal location. The different transformations, affecting only one vertex at a time, can be found in the literature [47].

The final optimization criterion,  $\epsilon_{Tot}$ , must be minimized and is the combination of three criteria:  $\epsilon_{houses}$  seeks to minimize the number of houses needed,  $\epsilon_{service}$  inflicts a heavy penalty if there is not enough  $H_2$  to refuel the bus and  $\epsilon_{Eco}$  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 shown in Equation (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:

- Ninety houses, each house 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 us to create a virtual system, which is used to identify the improved monitoring strategies. Households consumption, electricity prices, and meteorological data 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 GAMs in Python. GAMs are a type of 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 have 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 Algorithm 1, has the following priorities: first, self-consumption; second, H<sub>2</sub> delivery and, third, selling energy to the grid. The basic policy has a robust approach, as it tends to fill the H<sub>2</sub> tank as soon as possible 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, Algorithm 2, where all of the energy is sold to the grid. The idea is that, whenever it is relevant, the trained policy should be able to profit from the price variability.

---

##### Algorithm 1 Storing strategy

---

```

 $\Delta E \leftarrow C - \mathcal{P}r$ 
if  $\Delta E > 0$ , energy is available
  then
     $E^{H2} \leftarrow E^{H2} + \min(\Delta E, \mathcal{P}_{max}^{H2} \cdot \Delta t)$ 
     $\Delta E \leftarrow \max(\Delta E - \mathcal{P}_{max}^{H2} \cdot \Delta t, 0)$ 
  else
    Nothing, energy difference is bought from the grid
  end if
 $E_{exchanges} \leftarrow \Delta E$ 

```

---



---

##### Algorithm 2 Selling strategy

---

```

 $\Delta E \leftarrow C - \mathcal{P}r$ 
 $E_{exchanges} \leftarrow \Delta E$ , energy difference is either bought or sold to the grid, no H2 is produced

```

---

#### 4.3. Clustering

In this part we define three items: the elaboration of the situations that 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 the 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 long sequence of several years.

This way allows us to increase the variety of situations around the one already encountered while avoiding irrelevant situations, such as 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 it into little sequences. These little sequences are the ones gathered in clusters. Each sequence is described by three parameters:

- Energy available: the quantity of energy produced minus the consumption, i.e., the energy to be distributed between H<sub>2</sub> production and electricity selling;
- Electricity price: the average electricity price over the sequence;
- H<sub>2</sub> mass stored: the quantity of H<sub>2</sub> 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 directly use cluster centers to train the policy. Instead, for each cluster, we use the closest sequence to each center: the distance is obviously calculated with the least square, as it is the norm 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 Equation (6), and we keep the best one. The evaluation is based on the financial gain and the production of hydrogen. The parameter  $C_{H2}$  sets the relative importance of H<sub>2</sub> production compared to financial gain: the higher it is, the more H<sub>2</sub> production is favored. 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.

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

$$\epsilon_{H2} = \frac{\Delta m_{H2}^{tank}}{m_{H2}^{bus}} \cdot \frac{l}{24} \quad (6b)$$

$$\epsilon_{tot}^{strategy} = \epsilon_{Eco} + \epsilon_{H2} \cdot C_{H2} \quad (6c)$$

#### 4.5. Performance Assessment

Performance assessment consists of a two-year-long simulation relying on 2020 and 2021 data, applying a trained policy and assessing its score according to two metrics.

As stated previously, the main goal of the system is to produce enough hydrogen for the daily bus recharge. Thus, a heavy penalty is inflicted when hydrogen is lacking. The secondary objective is to maximize the financial earnings by selling electricity to the grid at the best moments. As the raw variation of money flow makes 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<sub>2</sub> 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<sub>2</sub>;
4. Deduce the minimum loss, i.e., the minimum money not sold to the grid to the production of H<sub>2</sub>;
5. Calculate the theoretical best money flow by subtracting the money loss due to H<sub>2</sub> production from the money earned when no H<sub>2</sub> 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 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<sub>2</sub> production: the relative importance of H<sub>2</sub> production when identifying the best strategy.

Table 1 presents the different values assessed for each parameter, for a total of 96 runs. Concerning sequence length, there are two things that should be known: the software uses a 1-h long time step and the price variations tend to follow a daily pattern. Thus, 2 h appears as the lowest sequence to consider (as it has only 2 steps) while 24 h 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 creating a lot of situations that the system will not encounter, which will thus dilute the useful information. It ranges from 0, with an absolute lower bound of 0.3, which is a value sufficient to observe the dilution effect.

Last, the relative weight of H<sub>2</sub> production ranges from 0 (where H<sub>2</sub> production does not count in the strategy assessment) to 0.75, where strategy always prefer to produce H<sub>2</sub>.

**Table 1.** Values explored in the sensitivity analysis.

Sequence Length (h)	Standard Deviation (-)	Relative Weight of H <sub>2</sub> production (-)
2, 3, 6, 12, 18, 24	0, 0.1, 0.2, 0.3	0, 0.25, 0.5, 0.75

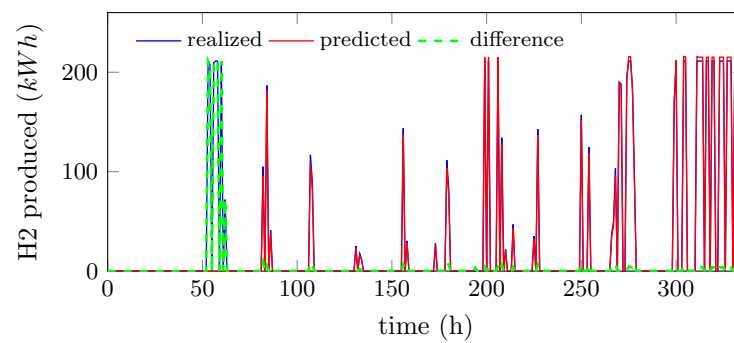
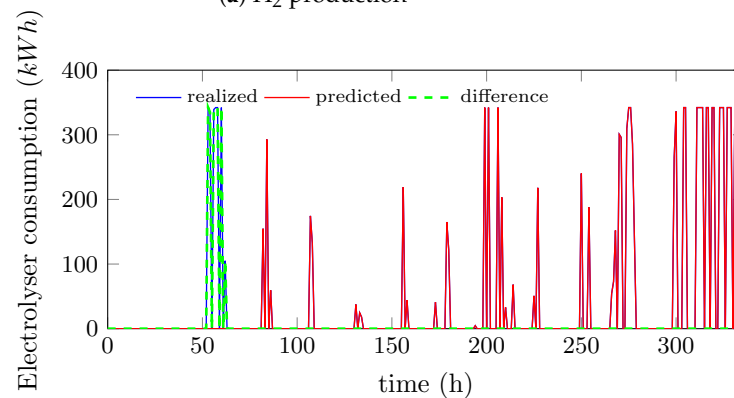
## 5. Results and Discussion

### 5.1. Digital Twins Fitting

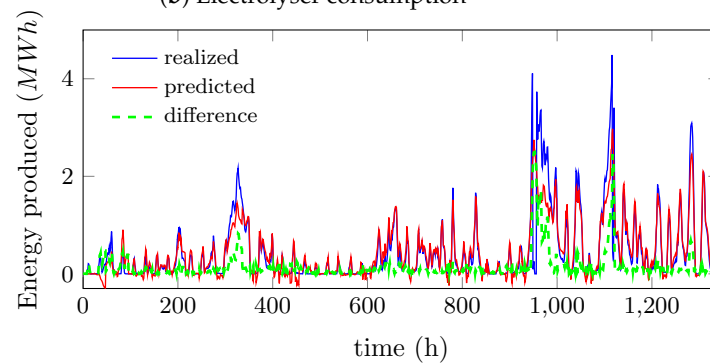
The first step of our approach is to create efficient digital twins. Figure 5 presents the digital twins predictions compared to the realized values. In this setting, the digital twin model is re-evaluated each day.

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

Concerning the H<sub>2</sub> production cost and the electrolyser consumption, 1 week of training (168 h) 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 H<sub>2</sub> 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.

(a) H<sub>2</sub> production

(b) Electrolyser consumption



(c) Renewable energy production

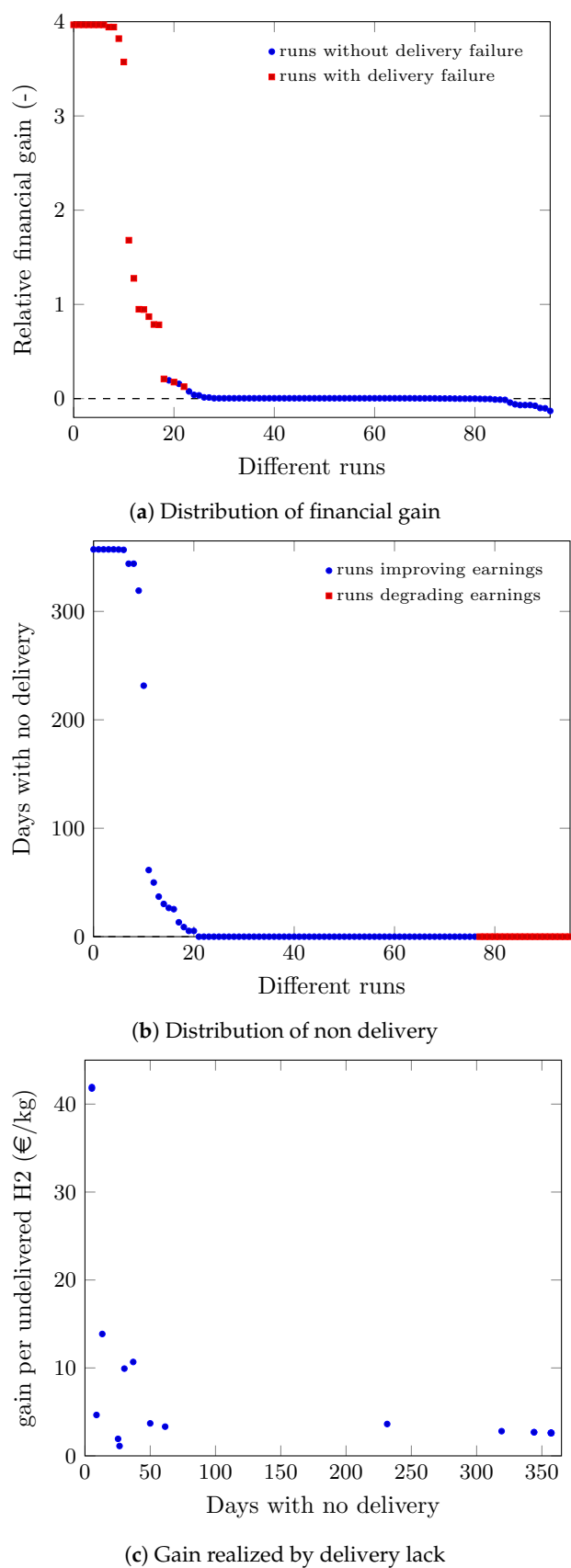
**Figure 5.** Comparison of values realized and predicted by the digital twins.

### 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 analyzed: the financial gain and the potential default of H<sub>2</sub> delivery.

#### 5.2.1. Overall Performance

At first glance, in 21 cases out of 96, H<sub>2</sub> 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 is observable in Figure 6a,b, there is no run accumulating undelivered H<sub>2</sub> 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, or they sell too much and do not produce enough H<sub>2</sub>.



**Figure 6.** Overall results.

More worryingly, in 64 runs (67%), the financial difference is lower than 5% variation, meaning that the training has no sensible effect. From the observations made, 3 runs out of

96 significantly improved 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: as a reminder, the method relies on a statistical approach based 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.

At second glance, knowing that the renewable H<sub>2</sub> price revolves around 3–8 \$/kg [48,49], it can be interesting to accept punctual delivery default: in some cases, the financial gain reaches up to 42 €/kg as shown in Figure 6c, i.e., at least 5 times H<sub>2</sub> price. Moreover, it can be seen that these high values are found for relatively low values of undelivered H<sub>2</sub> (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 H<sub>2</sub> by other means or use electricity bought from the grid to produce a bit of H<sub>2</sub> (while keeping in mind the requirements for “green” H<sub>2</sub> labels) to benefit from these policy advantages without their disadvantages.

### 5.2.2. Sensitivity to Parameters

First, the impact on sequence length can seem to be unclear regarding the performance of the trained policies, as it can be seen in Table 2: for the three metrics, the evolution is not monotonous. Meanwhile, it is possible to distribute the simulation results in three sequence groups (separated by dashed lines): 2 h; 3 and 6 h; 12, 18 and 24 h. 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 H<sub>2</sub> quantity goes from 3 days/year to 79. If raw financial gain is globally increasing with sequence length, the gain per undelivered H<sub>2</sub> mass is far higher for 2 h-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 H<sub>2</sub> is interesting only if it can earn more than by selling H<sub>2</sub>: as green H<sub>2</sub> prices are around 3–8 \$/kg, earning 3€ instead of delivering H<sub>2</sub> is nonsense.

**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 <sub>2</sub> (€/kg)	17.0	3.3	2.6	2.7	2.7	2.9

Second, as observed in Table 3, standard deviation can have a positive effect on performance but has to be used with caution.

**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 <sub>2</sub> (€/kg)	2.9	3.8	3.0	2.9

With 10 and 20% of standard deviation, undelivered H<sub>2</sub> reduces 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 H<sub>2</sub> 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 H<sub>2</sub> is still stable (2.9 €/kg). Undelivered H<sub>2</sub>, 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 many different situations and it is no longer possible to assign a single strategy efficient for all situations.

Consequently, using data augmentation methods, at least in the studied case, can improve policy performances as long as it does not hamper with the 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 <sub>2</sub> (€/kg)	2.9	3.7	N/A	N/A

The stock coefficient has the clearer impact on policies performance (Table 4). When the stock coefficient is equal to 0, it means that H<sub>2</sub> production is not taken into account. Unsurprisingly, it results in a very high quantity of undelivered H<sub>2</sub>, of 151 days/year, as the policy tends to sell all electricity to the grid. On the other side of the spectrum, for coefficients of 0.5 and 0.75, all H<sub>2</sub> is delivered but no financial gain is made: the trained policy produces as much H<sub>2</sub> as possible and imitates the basic policy.

Consequently, 0.25 seems to be the best value as it avoids these two extremes.

## 6. Conclusions

In this paper we have developed a methodology to evaluate and choose the 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 the 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. Even though in several cases, the trained policy has degraded the system's financial performance, this study shows that it is possible to outperform an energy conservative strategy. More studies are needed to establish a robust approach.

**Author Contributions:** Funding acquisition: S.G. and C.G. investigation: T.G., W.M. project administration: C.G. software: T.G. supervision: S.G. visualization: T.G. writing: T.G., W.M. and S.G. 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.

**Data Availability Statement:** Not applicable.

**Acknowledgments:** We thank 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, R. Dufo-Lopez, from the *University of Zaragoza*, for letting us use load data extracted from his iHOGA software and the team of 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

$A$	surface area, $m^2$
$C$	consumption, kWh
$C_p$	power coefficient, —
$c$	specific heat capacity, $J K^{-1} kg^{-1}$
$E$	total energy, J or kWh
$h$	specific enthalpy, $J kg^{-1}$
$l$	length of a sequence in clustering, h
$LHV$	lower heating value, $J m^{-3}$ or $J kg^{-1}$
$m$	mass, kg
$M$	money flow, €
$\hat{M}$	molar mass, $kg mol^{-1}$
OPEX	operational expenditures, €
$P$	pressure, Pa
$\mathcal{P}$	power, W
$\mathcal{P}_r$	production, kWh
$P$	price, €
$Q$	heat, J or kWh
$q$	flow-rate, $m^3 s^{-1}$ or $L s^{-1}$
$R$	individual ideal gas constant, $J K^{-1} kg^{-1}$
SOC	state of charge, %
$T$	temperature, K or °C
$t$	time, s or min
$V$	volume, $m^3$

### Greek symbols

$\alpha$	ground roughness factor, —
$\eta$	efficiency, — or %
$\rho$	density, $kg m^{-3}$
$\tau$	time constant, s
$\epsilon$	optimisation criterion, —

### Subscripts and superscripts

a	air
amb	ambient
bot	bottom
bui	building
ch	charge
cut	cutoff
dis	discharge
el	electrical
Q	heat



ht	heating
inv	inverter
nom	nominal
out	outdoor
pan	panel
ref	reference

#### Abbreviations

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

- Mitali, J.; Dhinakaran, S.; Mohamad, A. Energy Storage Systems: A Review. *Energy Storage Sav.* **2022**, *1*, 166–216. <https://doi.org/10.1016/j.enss.2022.07.002>.
- 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. *Adv. Appl. Energy* **2021**, *4*, 100060. <https://doi.org/10.1016/j.adapen.2021.100060>.
- A Hydrogen Strategy for a Climate-Neutral Europe; Technical Report; European Commission: Luxembourg, 2020.
- Vodovozov, V.; Raud, Z.; Petlenkov, E. Review of Energy Challenges and Horizons of Hydrogen City Buses. *Energies* **2022**, *15*, 6945. <https://doi.org/10.3390/en15196945>.
- Gronier, T.; Maréchal, W.; Gibout, S.; Geissler, C. Relevance of Optimized Low-Scale Green H<sub>2</sub> Systems in a French Context: Two Case Studies. *Energies* **2022**, *15*, 3731. <https://doi.org/10.3390/en15103731>.
- Petrozziello, A.; Troiano, L.; Serra, A.; Jordanov, I.; Storti, G.; Tagliaferri, R.; La Rocca, M. Deep learning for volatility forecasting in asset management. *Soft Comput.* **2022**, *26*, 8553–8574.
- Christensen, K.; Siggaard, M.; Veliyev, B. A Machine Learning Approach to Volatility Forecasting. 2021. Available online: [https://papers.ssrn.com/sol3/Delivery.cfm/SSRN\\_ID3766999\\_code414727.pdf?abstractid=3766999&mirid=1](https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3766999_code414727.pdf?abstractid=3766999&mirid=1) (accessed on 29 October 2022).
- Sadaqat, A.; Zheng, Z.; Aillerie, M.; Péra, M.C.; Hissel, D. A Review of DC Microgrid Energy Management Systems Dedicated to Residential Applications. *Energies* **2021**, *14*, 4308.
- Sang, J.; Sun, H.; Kou, L. Deep Reinforcement Learning Microgrid Optimization Strategy Considering Priority Flexible Demand Side. *Sensors* **2022**, *22*, 2256. <https://doi.org/10.3390/s22062256>.
- Nyong-Bassey, B.E. A Concise Review of Energy Management Strategies for Hybrid Energy Storage Systems. *Eur. J. Eng. Technol. Res.* **2022**, *7*, 77–81. <https://doi.org/10.24018/ejeng.2022.7.3.2815>.
- 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. <https://doi.org/10.1109/ACCESS.2020.3038735>.
- Borden, K.; Herit, A.; Company, M.. Digital Twins: What Could They Do for Your Business? 2022. Available online: <https://www.mckinsey.com/capabilities/operations/our-insights/digital-twins-what-could-they-do-for-your-business> (accessed on 1 August 2022).
- Rasheed, A.; San, O.; Kvamsdal, T. Digital Twin: Values, Challenges and Enablers from a Modeling Perspective. *IEEE Access* **2020**, *8*, 21980–22012. <https://doi.org/10.1109/ACCESS.2020.2970143>.
- 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. *Arch. Comput. Methods Eng.* **2020**, *27*, 105–134.
- Tao, F.; Xiao, B.; Qi, Q.; Cheng, J.; Ji, P. Digital twin modeling. *J. Manuf. Syst.* **2022**, *64*, 372–389.
- Errandonea, I.; Beltrán, S.; Arrizabalaga, S. Digital Twin for maintenance: A literature review. *Comput. Ind.* **2020**, *123*, 103316.
- Dembski, F.; Wössner, U.; Yamu, C. Digital twin. Virtual Reality and Space Syntax: Civic Engagement and Decision Support for Smart, Sustainable Cities. In Proceedings of the 12th International Space Syntax Conference, Beijing, China, 8–13 July 2019; pp. 8–13.
- Singh, M.; Fuenmayor, E.; Hinchy, E.P.; Qiao, Y.; Murray, N.; Devine, D. Digital twin: Origin to future. *Appl. Syst. Innov.* **2021**, *4*, 36.
- Qi, Q.; Tao, F.; Hu, T.; Anwer, N.; Liu, A.; Wei, Y.; Wang, L.; Nee, A. Enabling technologies and tools for digital twin. *J. Manuf. Syst.* **2021**, *58*, 3–21.
- Zheng, Y.; Yang, S.; Cheng, H. An application framework of digital twin and its case study. *J. Ambient. Intell. Humaniz. Comput.* **2019**, *10*, 1141–1153.
- Lu, H.; Du, M.; Qian, K.; He, X.; Wang, K. GAN-based data augmentation strategy for sensor anomaly detection in industrial robots. *IEEE Sens. J.* **2021**, *22*, 17464–17474.
- Gan, J.; Smith, C. Drivers for renewable energy: A comparison among OECD countries. *Biomass Bioenergy* **2011**, *35*, 4497–4503.

23. Lu, J.; Yi, S. Autoencoding Conditional GAN for Portfolio Allocation Diversification. *arXiv* **2022**, arXiv:2207.05701.
24. Xu, Y.; Sun, Y.; Liu, X.; Zheng, Y. A digital-twin-assisted fault diagnosis using deep transfer learning. *IEEE Access* **2019**, *7*, 19990–19999.
25. Maschler, B.; Braun, D.; Jazdi, N.; Weyrich, M. Transfer learning as an enabler of the intelligent digital twin. *Procedia CIRP* **2021**, *100*, 127–132.
26. 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. *Reliab. Eng. Syst. Saf.* **2021**, *215*, 107938.
27. Deebak, B.; Al-Turjman, F. Digital-twin assisted: Fault diagnosis using deep transfer learning for machining tool condition. *Int. J. Intell. Syst.* **2021**, 1–28. <https://doi.org/10.1002/int.22493>.
28. Liu, M.; Fang, S.; Dong, H.; Xu, C. Review of digital twin about concepts, technologies, and industrial applications. *J. Manuf. Syst.* **2021**, *58*, 346–361.
29. Naser, M. Digital twin for next gen concretes: On-demand tuning of vulnerable mixtures through Explainable and Anomalous Machine Learning. *Cem. Concr. Compos.* **2022**, *132*, 104640.
30. 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*, 10331. <https://doi.org/10.3390/su131810331>.
31. Battula, A.R.; Vuddanti, S.; Salkuti, S.R. Review of Energy Management System Approaches in Microgrids. *Energies* **2021**, *14*, 5459. <https://doi.org/10.3390/en14175459>.
32. 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. *Mach. Learn.* **2021**, *110*, 2335–2387. <https://doi.org/10.1007/s10994-021-06020-8>.
33. 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. Available online: <https://hdl.handle.net/2268/264828> (accessed on 1 September 2022).
34. Bolland, A.; Boukas, I.; Berger, M.; Ernst, D. Jointly Learning Environments and Control Policies with Projected Stochastic Gradient Ascent. *J. Artif. Intell. Res.* **2022**, *73*, 117–171. <https://doi.org/10.1007/s10994-021-06020-8>.
35. Henry, R.; Ernst, D. Gym-ANM: Reinforcement learning environments for active network management tasks in electricity distribution systems. *Energy AI* **2021**, *5*, 100092. <https://doi.org/10.1016/j.egyai.2021.100092>.
36. Yang, Y.; Li, H.; Shen, B.; Pei, W.; Peng, D. Microgrid Energy Management Strategy Base on UCB-A3C Learning. *Front. Energy Res.* **2022**, *10*, 858895. <https://doi.org/10.3389/fenrg.2022.858895>.
37. Munos, R.; Stepletn, T.; Harutyunyan, A.; Bellemare, M.G. Safe and efficient off-policy reinforcement learning. *Advances in Neural Information Processing Systems 29 (NIPS 2016)*; 2016. Available online: <https://proceedings.neurips.cc/paper/2016/file/c3992e9a68c5ae12bd18488bc579b30d-Paper.pdf> (accessed on 1 January 2019).
38. Sutton, R.S.; Barto, A.G. *Reinforcement Learning: An Introduction*; The MIT Press: Cambridge, MA, USA, 2018.
39. Rodrigues, C.; Gérard, P.; Rouveirol, C. On and Off-Policy Relational Reinforcement Learning. 2008. Available online: <https://lipn.univ-paris13.fr/~gerard/docs/publications/rodrigues-ger-rou-rlp08-submit.pdf> (accessed on 1 July 2017).
40. LegiFrance. Articles L291-1 et L291-2 du code de l'énergie. 2021. Available online: [https://www.legifrance.gouv.fr/codes/article\\_lc/LEGIARTI000043976710](https://www.legifrance.gouv.fr/codes/article_lc/LEGIARTI000043976710) (accessed on 1 November 2022).
41. 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. *Appl. Energy* **2011**, *88*, 4033–4041. <https://doi.org/10.1016/j.apenergy.2011.04.019>.
42. 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). *J. Clim.* **2017**, *30*, 5419–5454. <https://doi.org/10.1175/JCLI-D-16-0758.1>.
43. Burton, T.; Jenkins, N.; Sharpe, D.; Bossanyi, E. *Wind Energy Handbook*, 2nd ed.; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2011.
44. Yue, M.; Lambert, H.; Pahon, E.; Roche, R.; Jemei, S.; Hissel, D. Hydrogen energy systems: A critical review of technologies, applications, trends and challenges. *Renew. Sustain. Energy Rev.* **2021**, *146*, 111180. <https://doi.org/10.1016/j.rser.2021.111180>.
45. Pelikan, M. Bayesian optimization algorithm. In *Hierarchical Bayesian Optimization Algorithm*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 31–48.
46. Maréchal, W. Utilisation de Méthodes Inverses Pour la caractéRisation de matéRiaux à Changement de Phase (MCP). Ph.D. Thesis, 2014. Thèse de Doctorat de l'Université de Pau et des Pays de l'Adour, Pau, France.
47. 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, FL, USA, 1991.
48. *The Future of Hydrogen*; Technical Report; International Energy Agency: Paris, France, 2019.
49. IRENA. *Hydrogen: A Renewable Energy Perspective*; Technical Report; International Renewable Energy Agency: Abu Dhabi, United Arab Emirates, 2019.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.