

Research Article

Cost-Effective Energy Usage in a Microgrid Using a Learning Algorithm

Chaimaa Essayeh , Mohammed Raiss El-Fenni, and Hamza Dahmouni

Department of Communication Systems, INPT, Madinat Al Irfane, Rabat, Morocco

Correspondence should be addressed to Chaimaa Essayeh; chaimaa.essayeh@gmail.com

Received 1 January 2018; Accepted 12 March 2018; Published 22 April 2018

Academic Editor: Qiangfu Zhao

Copyright © 2018 Chaimaa Essayeh et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The microgrid is a new concept of integrating the distributed energy resources (DER) within the grid. The management of the heterogeneous sources of energy presents a challenge, especially as most of the DER are unpredictable. Besides, implementing microgrids should be economically beneficial to the customer; this will raise the challenge of decreasing the costs while ensuring the energy balance. In this paper, we used a stochastic approach based on a model-free Markov decision process (MDP) to derive the optimal strategy for the home energy management system. The approach aims to decrease the energy bill while taking into account the intermittency of the renewable energy resources (DER) and other constraints. While other proposals charge the battery from the utility energy, making the state of charge (SOC) of the battery a deterministic variable, our work adopts a scenario where the battery is charged from the excess of the generated energy, which makes the SOC a nondeterministic variable affected by the uncertain character of the renewable energy. Therefore, our model considers the randomness at two levels: renewable energy level and battery SOC level. We take into account the complexity of the solution, and we propose a simple strategy that can be implemented easily in microgrids.

1. Introduction

The electric grid is one of the big consumers of the fuel. Such type of energy source is exhaustible and is getting to disappear in the too near future. Electricity suppliers across the world are now searching for new alternatives to compensate it and are making solid steps to incorporate new technologies in many aspects of the grid.

With the raise of the new green technologies such as the PV panels, the wind turbines, and the electrical batteries, new ways of consuming energy emerge. In particular, the incorporation of such technologies with an information infrastructure created the concept of the smart grid [1]. Though it has not yet a standardized definition or a defined architecture, the smart grid can be seen as the innovation that will transform the electric grid from centralized and producer-controlled to a distributed and consumer-driven grid. The opportunities offered are countless, and the integration of the ICT infrastructure leads to a rich content and multiple scenarios of use. In fact, with the help of the ICT

infrastructure, more information is provided and this will enable an efficient control and monitoring of the grid system.

The home energy management system (HEMS) is one of the research fields associated with the smart grid concept. How to manage the energy efficiently and optimally using the new resources of energy and taking advantage of the ICT infrastructure is the core topic of the HEMS. Several distributed architectures of the HEMS are suggested in the literature; some of them propose a multiagent system [2, 3] to monitor the different technologies (power electronics, telecommunications, generation, and storage energy systems) that compose the energy management system [4]. Others proposed architectures that manage the energy flow at the substations; the local generated energy is sold to the utility instead of being consumed locally. The utility integrates then controlling strategies at substations to prioritize the DER and make them the first energy supply for the customers [5, 6]. Other works have suggested the use of the smart microgrid (μ grid) architecture (e.g., [7–10]). The US Department of Energy (DOE) defines the μ grid as “a group of interconnected

loads and distributed energy resources (DER) with clearly defined electrical boundaries that acts as a single controllable entity with respect to the grid [and can] connect and disconnect from the grid to enable it to operate in both grid-connected or island mode.” In the μ grid, the management of the heterogeneous sources of energy is done locally, thus, presenting a challenge, especially as most of the DER are unpredictable. Besides, implementing microgrids should be economically beneficial to the customer. And this will raise the challenge of decreasing the costs while ensuring the energy balance.

As aforementioned, the μ grid system works normally in an “on-grid” mode (or the connected mode). In this mode, the μ grid is connected to the utility grid. It can either injects the local generated energy into or take off energy from the utility grid. When a malfunction occurs in the grid, the μ grid disconnects from the grid and turns to the autonomous mode (off-grid mode, islanded mode). In this mode, the μ grid starts working autonomously and uses the local energy only. In previous works [11, 12], we proposed a load shedding/shifting strategy for the off-grid and the on-grid mode, respectively. The proposed strategy for the off-grid mode aims to reschedule efficiently the home loads to meet the limited amount of energy available during the off-grid switching, whereas in [12], we suggested a strategy for the on-grid mode that aimed to reduce the energy bill. The optimization strategy was modelled as an MDP and the model was solved using the value iteration algorithm. We assumed that the transition function and the reward function are known and used sample values in the simulation. However, in more realistic scenarios, we cannot predict those functions. Notably, a μ grid system relies on a complex stochastic structure, which is not amenable to closed form analysis. On the one hand, renewable energy depends on the most random natural phenomena. Research has shown that even the most robust forecasting models derive a forecast error of 10% when the time scale of forecasting exceeds 6 hours. On the other hand, the underlying structure and probability distributions for the consumer data D are based on the preferences and behaviour of individual consumers which are generally unknown. Therefore, it is not possible to find a closed form expression for the transition function. This paper is an improvement of [12]; it models the system as a model-free MDP and thus does not need to specify the transition function, and it solves it by the reinforcement learning. It proposes a strategy that returns an optimal charging/discharging cycle of battery while considering the intermittency of the renewable energy resources (RER) and the reduction of the utility energy consumption during peak hours and ensuring the user comfort. The rest of the paper is organized as follows. Next section presents a literature review of the existing solutions. Section 3 presents the grid architecture on which we work. The decision making approach as an MDP model is formulated in Section 4, and its resolution is presented in Section 5. The simulation results and the evaluation of the proposed solution are given in Section 6. The final section concludes this paper and discusses future works.

2. Related Work

A variety of energy management applications was widely addressed for the residential side. Many papers have treated the problem of efficiency and optimality in the μ grid systems [13, 14], and explored the role of the Demand Side Management (DSM) to reduce the electrical energy consumption, minimize the energy cost, and reduce environmental degradation. Several papers consider the optimal exploitation of renewable energy [5, 15, 16]. Some papers investigate the optimal exploitation of renewable energy considering cost-efficient energy scheduling for residential μ grids equipped with a centralized renewable energy source [5, 15]. Others focus on maintaining a continuous energy balance by ensuring uninterrupted energy delivery to end users [16]. Another approach to manage the μ grids is the load shifting approach. The approach consists of separating the residential loads into different classes (e.g., shiftable/nonshiftable loads [17], priority based classification [18]) and rescheduling the loads to achieve significant savings while ensuring the user comfort. The works introduce two techniques of shifting: day-ahead load scheduling and real-time load scheduling. The scenarios proposed for the demand shifting are multiple and include electrical storage shifting (e.g., batteries, electrical vehicles) [17, 19], thermal storage shifting [17, 20], or electrical equipment shifting [18]. Soares et al. [21] combined the different provided approaches. They present an evolutionary algorithm that optimizes the usage of heterogeneous energy resources (local generation and storage systems) and proposes an optimal strategy for the loads (shiftable loads and thermostatically controllable loads). Thus, a multiobjective model was developed to minimize the energy cost and to minimize the end-user’s dissatisfaction associated with the suggested management strategies.

Other researches reduce the problem of optimality in the battery and focus on the optimal charging/discharging battery cycles [22–24]. In fact, if used optimally, that is, charged or discharged at the right moment, the battery can ensure an optimal performance of the microgrid.

The authors of [22, 23] suggested the scenario that gives the battery the right to be charged from the utility grid only; the power flow from the battery to the grid is not permitted. The authors of [22] studied the control of the energy storage under price fluctuations and then derived the structure of the cost-minimizing storage policy using a model-based MDP framework. In [23], the resolution of the system was done using a model-free reinforcement learning. The objective of the optimal control in that paper was to find the optimal battery charging/discharging/idle control law which minimizes the total expense of the power from the grid while considering the battery limitations. Li and Jayaweera [24] proposed a scenario that consists of storing, fully or partially, the positive excess energy of a customer for its own use in the future or selling the excess to the utility, taking into account the possible time variant pricing information. The authors developed an HMDP (hidden MDP) model for customer real-time decision making.

In this paper, we discuss another scenario where the flow of energy from the μ grid to the utility is not allowed, and the

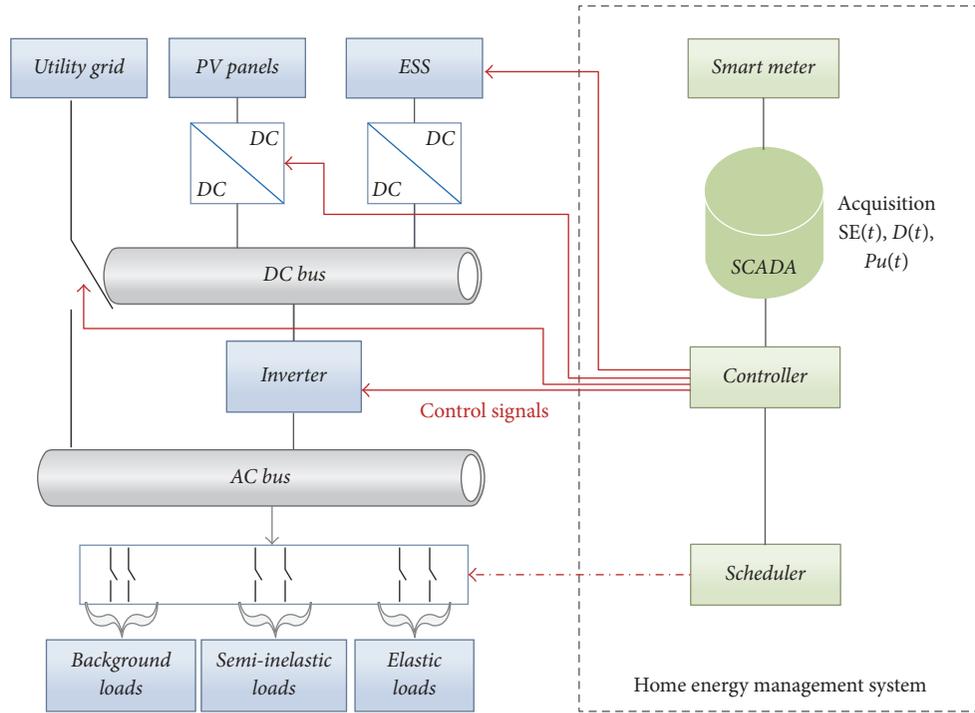


FIGURE 1: Smart microgrid architecture.

battery is charged from the excess of the generated energy. Thus, the battery SOC becomes a nondeterministic variable, affected by the uncertain character of the renewable energy. Therefore, our model takes into account the randomness at two levels: renewable energy level and the battery SOC level. Our approach aims to enhance cost-savings, while taking into account the following constraints:

- (1) the optimal exploitation of the battery which is charged by the renewable energy,
- (2) the management of utility energy especially during peak hours,
- (3) the maintainability of energy balance,
- (4) the ensuring of user comfort,
- (5) the integration of a dynamic pricing method.

We take into account the complexity of the solution, and we reduced the resolution into a simple strategy that can be implemented easily in the microgrid.

3. Smart Microgrid Architecture

A μ grid system can serve different electricity customers: a house, a residence, a small industry, and so on. We consider an on-/off-grid architecture as defined in Figure 1. The prosumer is connected to two types of energy resources: renewable energy resource (RER) and utility energy. The distributed RER is the primary energy supply and the utility grid is the backup energy supply [3]. The energy generated locally is directly consumed without being injected through the grid. The local energy supplying isolates the prosumer

appliances from the line and generators; this reduces the loss due to energy transport and provides benefits such as the reduction of signal distortion [5] and the reduction of costs of maintenance and management operations. Furthermore, the direct consumption of the local generated energy will be beneficial, since it will ensure the μ grid flexibility by allowing its operation under two different modes: the connected mode and the islanded mode.

We decided not to allow the energy flow from the μ grid to the utility, since this is the most adopted case in a great number of countries where the power injection to the utility grid is not legalized yet. In fact, the power grid as it is implemented currently in most developed countries has a hierarchical structure: we have the centralized power plants, the HV stations (transmission networks), the MV stations (distribution networks), and the low-voltage (220/230 V) that run into our houses. The transformer stations that connect the different levels of power voltage often have some kind of decoupling mechanism that prevent the bidirectional flow of energy, especially from a lower level to a higher level as it would through a linear transformer. In particular, this case is suitable for Morocco, and the study is applied on the institute implemented μ grid. However, the energy bidirectional flow between the μ grid and the utility grid can be treated but only in the theoretical level and will be the subject of a further research.

In order to take a full advantage of the RER, we integrate an ESS (Energy Storage System) into the model. The ESS is used to store the excess generated energy and work as a source of electricity supply in case the local energy is not sufficient to meet the load demand. This component will help increase the autonomy rate of the system and reduce the dependence

of the prosumer on the utility grid. In the off-grid mode, the ESS is considered as the backup supply that performs the energy balance. When there is an excess of local energy, that is, if the local generated energy feeds the load and fully charges the battery, the RER is disconnected not to harm house equipment.

We integrate a HEMS infrastructure within the electric grid. The HEMS is equipped with a smart meter, which is a component of the AMI (Advanced Metering Infrastructure), to enable a two-way communication. On the one hand, the smart meter ensures the communication from the HEMS to the control center by transmitting the detailed measurement of electricity consumption in a household. On the other hand, it serves as a gateway that transmits the information about the utility prices from the control center to the HEMS. The smart meter spots the power losses and helps identify the devices that consume the most electricity. The data collected in the smart meter is transferred to the Supervisory Control and Data Acquisition (SCADA) Center for storage and processing. The data gathered in the SCADA can be used for many purposes. The data archive of the generated energy and the power consumption can help elaborate efficient models of forecasting. Meanwhile, the real-time data is processed to elaborate controlling algorithms including the home energy management strategies. The controller is a subfeature of the SCADA, it is used to communicate the instructions of the SCADA to the components of the electric system through control signals. The above architecture was implemented in INPT at a larger scale for the MDE smart μ grid project [25]. The institute is composed of different buildings (cf. Figure 2); each building is considered as a μ grid and has all the electrical components that we cited in Figure 1.

4. MDP Based Approach Formulation

A Markov decision process is defined by a 4-tuple $\{S, A, P, R\}$, where S is a finite set of states, A is a finite set of actions, P is the transition function, and R is the reward function. A transition element $P_a(s, s')$ of the transition function defines the probability that action a in state s at time t will lead to state s' at time $t + 1$. It requires well knowledge of the system states which are composed (as we will see later) by renewable generation and electric consumption. Hence, renewable energy generation forecasting faces several mathematical and statistical challenges [26, 27] in the technical level (material, technology, meteorological estimations, etc.). Besides, The consumer electric behaviour cannot be predicted; it depends on one's preferences and can vary depending on time period (for instance, the consumption differs remarkably between summer days and winter days). Therefore, we will use the model-free reinforcement learning to learn the state-transition function and solve the system. The three remaining elements (S, A, R) will be projected on our case study and detailed in the current section.

4.1. System States: S . In each time-slot t , the system state is defined by the triplet: $S(t) = \{D(t), SE(t), P_u(t)\}$, with the following:

- (i) $D(t)$ is the net demand of the prosumer, $D(t) = L(t) - G(t)$, where $L(t)$ is the demand load of the user at time t and $G(t)$ is the amount of local produced energy at time t ;
- (ii) $SE(t)$ is the amount of stored energy in the battery at time t ;
- (iii) $P_u(t)$ is the price of the utility energy at time t .

The battery state of charge ($SE(t)$) is not deterministic. The stochasticity in the battery state of charge is derived from the stochasticity of the renewable energy generation, since the battery is charged from the excess of the generated energy.

The stored energy SE varies in the range $[C_{\min} C_{\max}]$ with C_{\max} being the maximum capacity of the battery and C_{\min} being the minimum capacity that has to remain in the battery, in order to extend its life time. The net demand, on the other hand, varies in the range of $[-G_{\max} L_{\max}]$. The lower bound of the net demand corresponds to the case where the generated energy is the highest and there is no load demand (this case is not realistic since there will always be the background appliances running in the house but we take it as a reference case to limit the demand load range), and the upper bound corresponds to the case where the load is in the peak but there is no generated energy.

Dynamic pricing models have been adopted in the latest smart grid strategies. Methods of dynamic pricing include among others time of use pricing (TOUP), real-time pricing (RTP), and critical peak pricing (CPP).

TOUP determines two or three levels of energy price, each level for a certain period of time of the day. The price levels are predetermined and can be changed only once or twice a year (summer period TOUP, winter period TOUP). In RTP, instead of predetermining the price levels, the exact price value for each period is calculated and announced to the user only at the beginning of the operation period. CPP is an event-driven pricing method, a CPP event is activated when the utility determines that there is a need to call on customers for temporary reductions in electricity use. In this study, we consider a TOUP price with three levels: off-peak price, mid-peak price, and on-peak price changing once a year (one model for the summer period, and another for the winter period). Our state space is then continuous, which makes the problem resolution impossible. We use a grid to discretize the state space and approximate the continuous MDP state space via a discrete space as shown in Table 1. The discrete state space is presented by (1), where i refers to the index of net demand interval to which $D(t)$ belongs, j refers to the index of state of charge interval to which $SE(t)$ belongs, and k is the price level.

$$S_{\text{discrete}} = \left\{ \left(s_{i,j,k} \right) \right\}, \quad (1)$$

$$1 \leq i \leq n, \quad 1 \leq j \leq m, \quad k \in \{1, 2, 3\}.$$

4.2. Actions: A . The decision making process will be limited to the states where $D(t) > 0$, because once the generated energy exceeds the demand load (i.e., $D(t) \leq 0$), the best action is to use only this generated energy, which is obviously

TABLE 1: Grid of substate space for a given price k .

	$[SE_1, SE_2[$	\dots	$[SE_j, SE_{j+1}[$	\dots	$[SE_{m-1}, SE_m[$
$[D_1, D_2[$	$s_{1,1,k}$	\dots	$s_{1,j,k}$	\dots	$s_{1,m,k}$
\vdots	\ddots	\ddots	\ddots	\ddots	\ddots
$[D_i, D_{i+1}[$	$s_{i,1,k}$	\dots	$s_{i,j,k}$	\dots	$s_{i,m,k}$
\vdots	\ddots	\ddots	\ddots	\ddots	\ddots
$[D_{n-1}, D_n[$	$s_{n,1,k}$	\dots	$s_{n,j,k}$	\dots	$s_{n,m,k}$

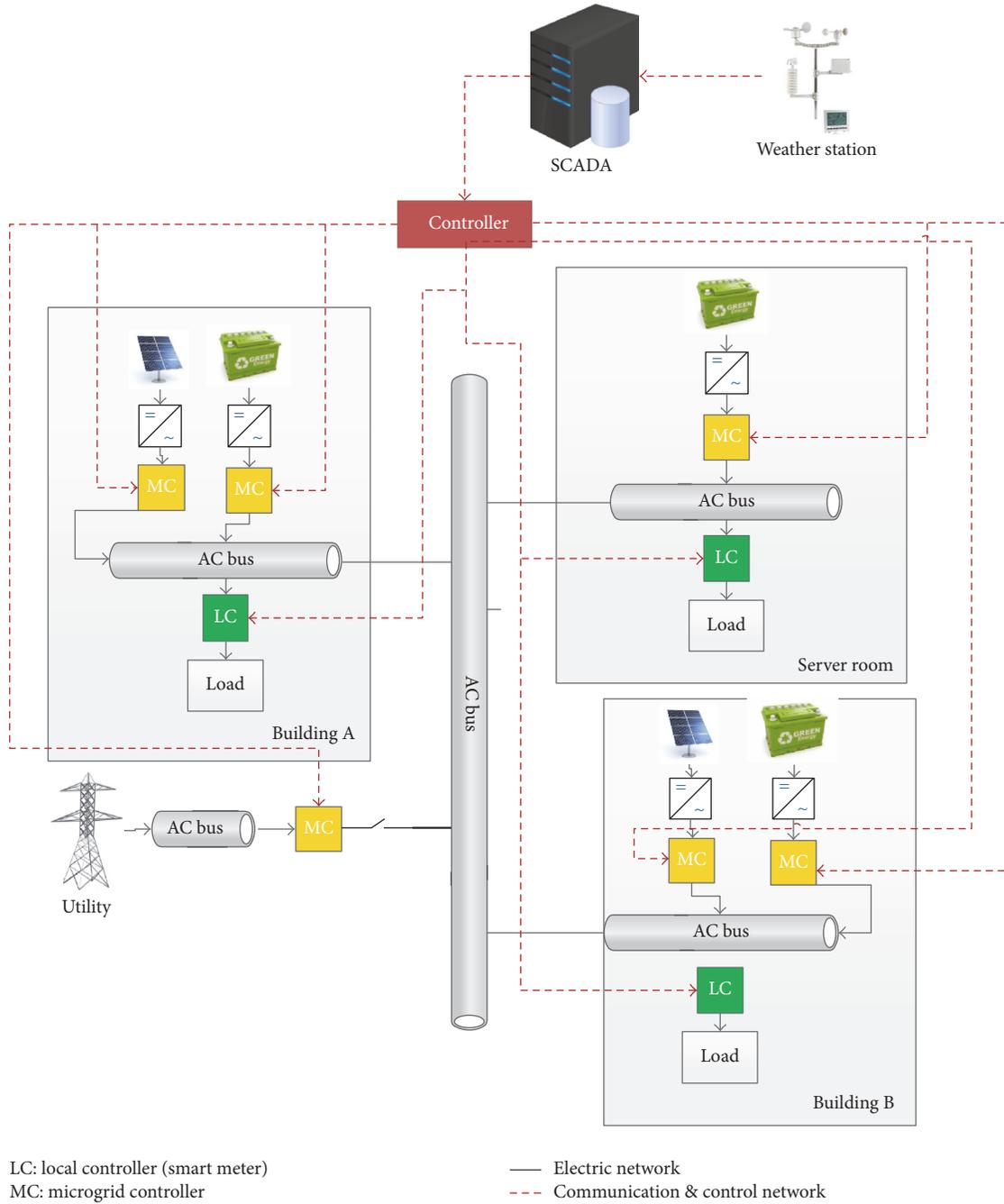


FIGURE 2: INPT microgrid architecture.

the most natural choice, since we have enough cheap energy to fulfill the load and to charge the battery.

a_0 : use the generated energy to fulfill the load and to charge the battery.

Depending on the actual state $S(t) = \{D(t) > 0, SE(t), P_u(t)\}$, the system chooses between the following actions:

a_1 : feed the net demand D with only the utility energy.

a_2 : use the battery to feed the net demand and complete with the utility energy if the stored energy is not sufficient.

The battery state (state of charge or depth of discharge) at time $t + 1$ depends on the action taken at the time t and the preceding battery state. In fact, the action a_1 does not affect the state of the battery since it does not use it neither in charge nor in discharge. In this case, an amount of the stored energy is decremented in the battery by the self-discharge phenomenon. Yet, since the battery does not remain stagnant for a long period of time, we can neglect this effect. This is a reasonable assumption when the time scale over which the loss caused by the self-discharge takes place is much larger than that of interest to us. The battery will be discharged if a_2 is chosen. The amount of energy discharged from the battery depends on the battery parameters and the available energy.

$$SE(t+1) = \begin{cases} \min(SE(t) + \min(r, -D(t)), C_{\max}) & \text{if } D(t) \leq 0, a = a_0 \\ SE(t) - \min(D(t), r, SE(t) - C_{\min}) & \text{if } D(t) > 0, a = a_2 \\ SE(t) & \text{if } D(t) > 0, a = a_1, \end{cases} \quad (2)$$

with r being the maximum charging/discharging rate.

4.3. Reward: R . The reward is a bounded, real-valued function that expresses the system preferences. One common form of the reward is the one consisting of two components, a positive component representing the utility U and a second component corresponding to the cost: $R = U - \text{Cost}$

In this paper, the cost function, Cost , corresponds to a fictive monetary cost of the consumed energy having chosen the action a :

$$\text{Cost}(S(t), a(t)) = \begin{cases} -P_b(t) \min[r, \min(-D(t), C_{\max} - SE(t))] & \text{if } a_0 \\ P_u(t) D(t) & \text{if } a_1 \\ P_b(t) B(t) + P_u(t) (D(t) - B(t)) & \text{if } a_2, \end{cases} \quad (3)$$

where $P_b(t)$ and $P_u(t)$ represent the prices, at time t , of the stored energy and the utility energy, respectively, and $B(t)$ is the amount of stored energy consumed at t .

The goal of MDP planning is to compute the optimal policy π^* , that is, the policy that maximizes the expected sum of rewards:

$$\pi^* = \arg \max_{\pi} \mathbf{E} \left[\sum_t R_t(S(t), a(t), S(t+1)) \mid \pi \right]. \quad (4)$$

An optimal stationary policy is the policy that is not worse than any other policy at any state and any time.

5. Model Resolution: Optimization of Energy Use

Reinforcement learning is a branch of the machine learning and also a branch of the artificial intelligence. It allows machines to define the best behaviour to enhance their performance. The machine learns—via a trial-error series—its behaviour based on the feedback (in a type of reward or punishment) of the environment.

Among the reinforcement learning algorithms existing in the literature, the Q-learning, SARSA, actor-critic learning, and the average reward difference temporal learning are the most used in the MDP context.

5.1. Q-Learning. Q-Learning is an off-policy algorithm for temporal difference learning. It is exploration-insensitive, which means that it will converge to the optimal policy regardless of the exploration policy being followed, under the assumption that each state-action pair is visited an infinite number of times, and the learning parameter α is decreased appropriately. The basic idea of the Q-learning is to compute a running average of the Q-value. At each iteration, the system faces new experience. The Q-learning averages the Q-value of the new experience with the current running average:

$$Q(s, a) \leftarrow (1 - \alpha) \underbrace{Q(s, a)}_{\text{average}} + \alpha \underbrace{\left[r + \gamma \max_{a'} Q(s', a') \right]}_{\text{new experience}}, \quad (5)$$

where γ is the discounted factor; it is a number in the range $[0, 1]$ used to discount reward's value. It is a reasonable way to encourage sooner rewards than later ones. The γ factor also helps the algorithm to converge since it enforces the sum of *discounted* rewards to be bounded. In fact, $\sum_t \gamma^t \times r_t \leq R_{\max}/(1 - \gamma)$. The closer γ to 0 is, the smaller the horizon is; we say that it is a shorter term evaluation. In the other case (γ closer to 1), the evaluation is said to be “far-sighted.”

α is called the learning rate; it has to decrease as the learning process progresses, in order to give the last estimates the highest weights. Some commonly used examples include $\alpha(t) = a/(b + t)$, where $a < b$ are real numbers and t is the iteration step. The rule $1/t$ where $a = 1$ and $b = 0$ may not lead to good behaviour [28], although it is still widely used. We suggest the use of the log rule

$$\alpha(t) = \frac{\log(t)}{t}, \quad (6)$$

since it converges to 0 slower than the $a/(b + t)$ rule for many values of a and b . This will help our algorithm to get the right estimate as the learning progresses.

5.2. Exploitation versus Exploration. During the learning process, the system discovers through acting the “goodness” or the “badness” of a state-action pair. On the one hand, the system tries to choose safe actions, which report a positive reward (exploiting the solution); on the other hand, it would explore all the states in order not to skip or neglect a better solution in terms of reward! Therefore, the system must take risks and visit the unknown stat-action pairs.

```

 $\forall(s, a) Q(s, a) = 0.$ 
 $iter \leftarrow 1$ 
Set  $\zeta$  and  $\gamma$  ▷ The convergence condition and the learning
parameter
Initialize  $s(iter = 1)$ 
repeat
  if  $D(iter) < 0$  then
    Charge the battery Eq. (2)
  else
    define  $\alpha(iter), P_{iter}(s, a)$ 
    generate random number  $\epsilon$ 
    if  $\epsilon \leq P_{iter}(s)$  then
       $\pi(s) \leftarrow \arg \max_a Q(s, a)$ 
    else  $\pi(s) \leftarrow$  random action else than  $\arg \max_a Q(s, a)$ 
    end if
    Execute  $\pi(s)$ 
    Update the state of the battery Eq. (2)
  end if
  Observe  $Reward(s, \pi(s), s(iter + 1))$  and Update Q;
  Equation (5)
   $iter \leftarrow iter + 1$ 
until  $\forall(s, a) : |Q_{updated}(s, a) - Q(s, a)| < \zeta$ 

```

ALGORITHM 1: Q-Learning.

The exploitation-exploration dilemma is typical for a reinforcement learning problem. The solution is to choose randomly an action while focusing on the actions that currently have better rewards. Moreover, the exploration rate should decrease as long as we progress in the learning process. The Boltzmann probability distribution verifies the two conditions: (1) focus on the good states, and (2) decrease as the learning progress.

$$P(a | s) = \frac{\exp(Q(s, a) / T)}{\sum_{a'} \exp(Q(s, a') / T)}, \quad (7)$$

where T is the Boltzmann temperature. T is initialized to a large value making all the state-actions equiprobable. By reducing T from a large initial value to a small final value, the system will focus on good states. At a temperature of 0 the exploration-exploitation rule becomes deterministic and the system chooses only the state-action that returns the best value of the q -function.

5.3. Final Executed Algorithm. Algorithm 1 describes the main steps of the model resolution. The algorithm starts by observing the initial state of the system. if the net demand is negative, the battery will be charged, and the system will get a reward for charging the battery. In the other case, the system generates a random number and compares it to the exploration rate. At the beginning of the algorithm execution, the exploration rate is equal for all actions (with the function Q being initialized at 0 for all the pairs state-action), and the system is more likely to choose a random action. As the learning progresses, the exploration rate decreases and its probability to be less than the random number decreases too. The system starts focusing on the good actions returning the

TABLE 2: Solar panel parameters.

Location and station identification	
Requested location	Helena
Weather data source	(TMY2) Helena, MT 2.8 mi
Latitude	46.6°N
Longitude	112°N
PV system specifications (<i>residential</i>)	
DC system size	1 KW
Module type	Standard
Array type	Fixed (open rack)
Array tilt	20°
Array azimuth	180°
System losses	14%
Inverter efficiency	96%
Dc to AC size ratio	1.1

best rewards. After executing the action, the system receives a reward and updates the state of charge of the battery according to (2). The system observes the next state and updates the Q -function (5).

6. Results

6.1. Simulation Parameters. To test the reliability of our model, we choose to simulate the results with the following.

Statistical Data. We consider statistical data for the generated renewable energy [29] and the demand load [30]. The solar panel technical characteristics are shown in Table 2. Figure 3 displays the hourly net demand of the user $D(t) = L(t) - G(t)$

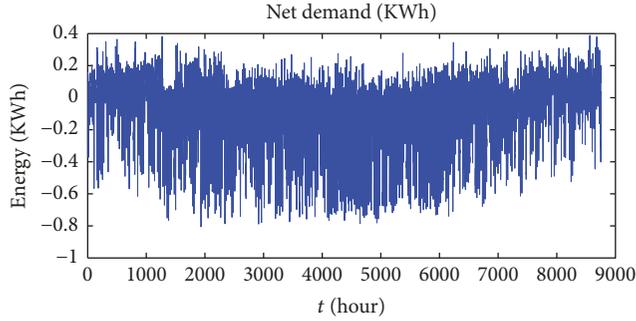


FIGURE 3: Simulation data.

during one year. *Solar Battery*. Regarding type of batteries suitable for the photovoltaic systems which allows a large number of charging/discharging cycles, we choose for the simulation a battery with 5 KWh (5280 Wh) as a maximum capacity. Yet, to extend the life time of the battery, we will use only 60% of this capacity (cf. Figure 4), which means that only 3168 Wh will be used. The minimum capacity is then equal to 2112 Wh. The maximum charging/discharging rate varies between $C_{\max}/5$ and $C_{\max}/20$, depending on the state of charge (SOC)/depth of discharge (DOD) of the battery. For simplification, we took a maximum charging rate fixed to $C_{\max}/10$ which returns 0.528 Wh. *Time of Use Price*. The TOUP model determines the price of the utility grid; Figure 5 shows the utility price variation depending on the time of the day and the day of the year.

6.2. Results Discussion. The continuous interval $]0 L_{\max}]$ of $D(t)$ is divided into two intervals.

The state of charge of the battery is divided into intervals of the range $C_{\max}/3$.

Since we have 3 levels of the utility price, we will have 18 states for the case $D(t) > 0$ and 1 state for the case $D(t) \leq 0$. Table 3 summarizes the values of the different parameters of the simulation. Table 4 shows the optimal policy of the suggested system. The resolution is reduced to two actions which are in fact a simplification of the following actions: energy generation, use of the utility energy as source, use of the batteries as source, and the combination of these last two possibilities. The goal is to reduce the complexity while keeping all the necessary elements for the model to be realistic. For the summer case, some of the states are never visited, and this is due to the large amount of generated energy provided in this period. In fact, during the summer period the net demand $D(t)$ decreases and the battery is almost nearly charged. These states are marked by a “-” symbol in the table of optimal policy.

We compare the suggested solution to two other scenarios:

- (i) Classic scenario: in this case, the battery is always given a priority over the utility energy. When the generated energy is not sufficient, the system chooses to use the battery and compensate with the utility energy if the sum of generated and stored energy is not sufficient to fulfill the user load.

TABLE 3: Simulation parameters.

P_u	P_u^1	8.7
	P_u^2	13.2
	P_u^3	15
<i>Solar Battery</i>	C_{\max}	5.28 KWh
	DoD	40%
	r maximum charging rate	0.528
Algorithm parameters	P_b	10
	γ	0.85
	Convergence condition ζ	0.001

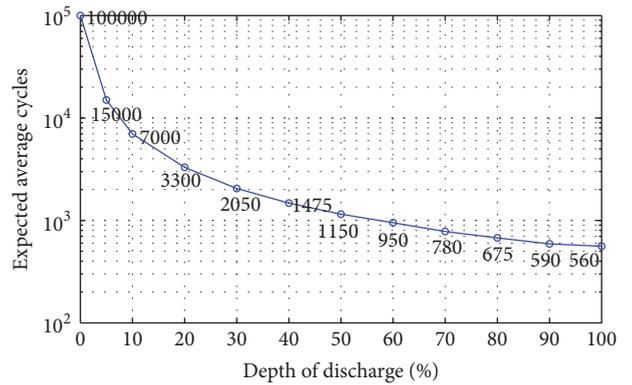


FIGURE 4: Cycle lifetime as function of depth of discharge.

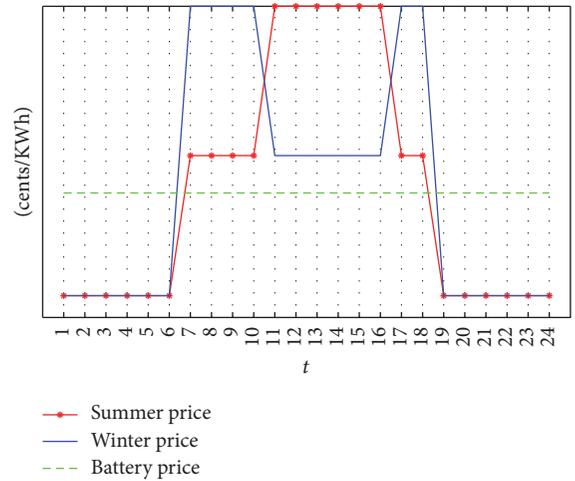


FIGURE 5: Time of use price.

- (ii) MDP-finite horizon scenario: we have proposed this case study in [12]. In this work, we considered that the data (the consumer total load and the predicted local generated energy) is updated daily. Therefore, the resolution program was done every 24 hours. We chose to work with the value iteration method, since it is the most adopted for resolving MDPs in finite horizon [31].

Figure 6 visualizes the monthly cost of the three scenarios. The total cost for the classic scenario is 3448.6 cents, while the

TABLE 4: Optimal policy.

D	SE	P_u	State	Optimal action (summer)	Optimal action (winter)
D_1	SE ₁	P_u^1	$s_{1,1,1}$	-	a_1
		P_u^2	$s_{1,1,2}$	-	a_1
		P_u^3	$s_{1,1,3}$	-	a_2
	SE ₂	P_u^1	$s_{1,2,1}$	-	a_1
		P_u^2	$s_{1,2,2}$	-	a_2
		P_u^3	$s_{1,2,3}$	-	a_2
	SE ₃	P_u^1	$s_{1,3,1}$	a_1	a_1
		P_u^2	$s_{1,3,2}$	a_2	a_2
		P_u^3	$s_{1,3,3}$	a_2	a_2
D_2	SE ₁	P_u^1	$s_{2,1,1}$	-	a_1
		P_u^2	$s_{2,1,2}$	-	a_1
		P_u^3	$s_{2,1,3}$	-	a_2
	SE ₂	P_u^1	$s_{2,2,1}$	-	a_1
		P_u^2	$s_{2,2,2}$	-	a_2
		P_u^3	$s_{2,2,3}$	-	a_2
	SE ₃	P_u^1	$s_{2,3,1}$	a_1	a_1
		P_u^2	$s_{2,3,2}$	a_2	a_2
		P_u^3	$s_{2,3,3}$	-	a_2

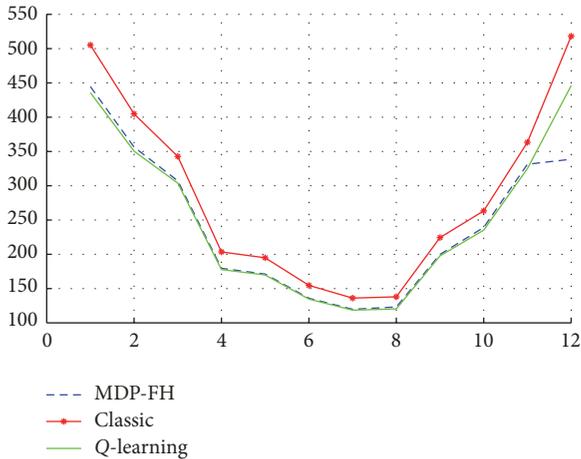
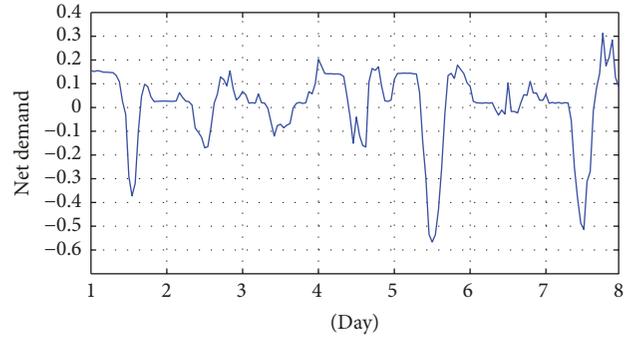
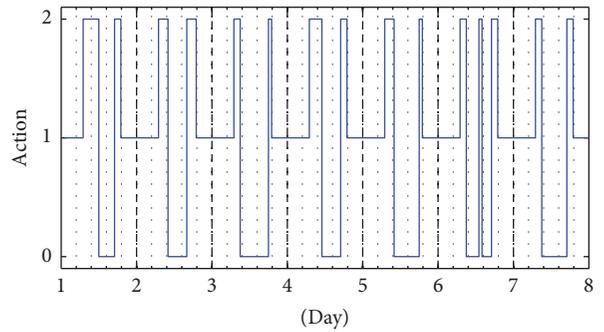


FIGURE 6: Cost of the different scenarios.

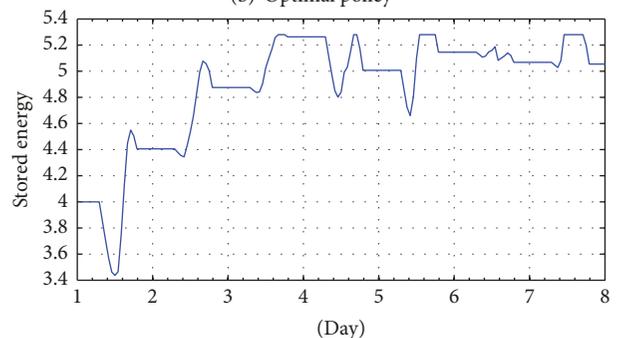
one of the MDP-FH scenario is 2946.3 cents. The Q Learning resolution stands with a total cost of 3014.3 cents, which means that the MDP-FH outperforms the results of the Q learning. Nevertheless, the MDP-FH scenario has a greater cost of computational complexity. In fact, the MDP-FH treats each day individually and is launched at the beginning of the day, contrary to the q learning that solves the system independently of the time of day and is launched only for once. In addition, the transition matrix for the MDP-FH was chosen arbitrarily. In more realistic scenario, this matrix may contain a margin of error due to the use of a statistical estimation method, which makes it very critical to use in a real implemented system. For greater readability, we show the results for restricted periods. Figures 7 and 8 show the system behaviour for two different weeks, a winter week and a summer week. Once the generated energy exceeds the demand load, the best action is to use only this generated energy, which is obviously the most natural choice since we have enough cheap energy to be used and to charge the



(a) Winter data



(b) Optimal policy



(c) Battery state of charge

FIGURE 7: Simulation for a winter week.

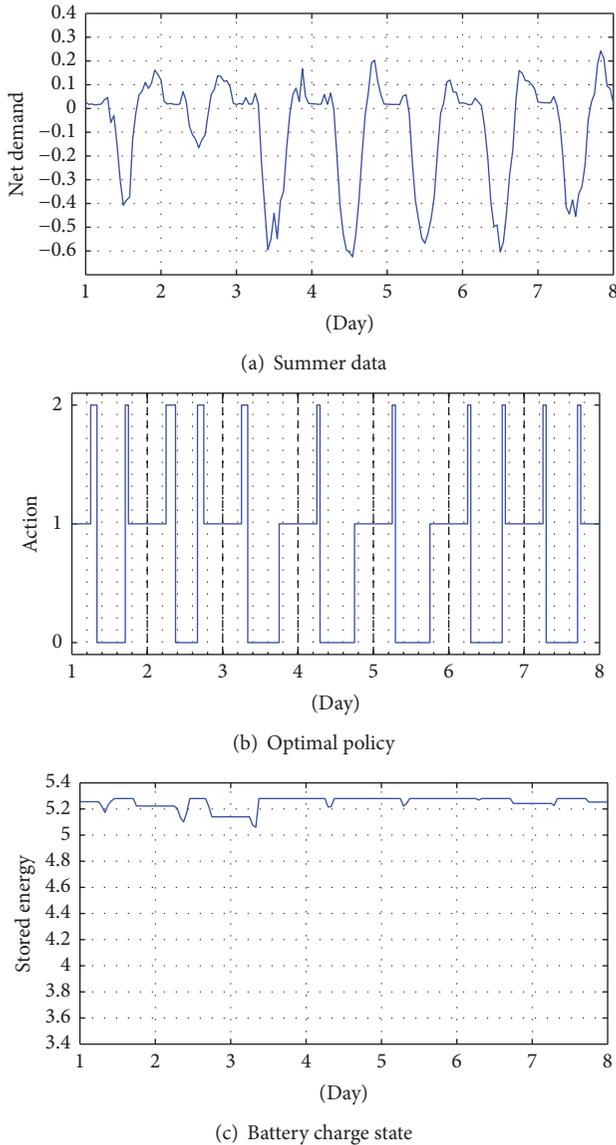


FIGURE 8: Simulation for a summer week.

battery. The period of energy excess differs from 5 hours in the winter day to 13 hours in the summer day. In this period, the battery starts charging (cf. Figures 7(c) and 8(c)). The battery is rarely utilized in the summer day, due to the fact that the peak period of the summer days coincides with the period of the generated energy excess, and the period of higher net demand coincides with the lower utility price.

7. Conclusion and Future Work

This paper proposes an intelligent home energy management model that manages the use of the intermittent distributed energy under several constraints in order to improve the balance of usage of heterogeneous energy resources. The proposed HEM model takes into account the elaboration of an effective scheduling of charging/discharging cycles of the battery and returns the optimal strategy that reduces the

annual energy bill. The effectiveness of the suggested model for finding an optimal solution is demonstrated through extensive simulations. These latter indicate that the model-free based resolution approaches the results of the model-based resolution in terms of cost optimization. The results could be more satisfying if a μ grid sizing study was carried on. Therefore, and in order to increase the consumer benefit, an optimal sizing of the μ grid architecture presented in the present paper will be performed.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was financially supported by the National Center of Scientific and Technical Research (CNRST), Morocco. It is also a part of the MDE project “INPT Smart Microgrid” funded by the Moroccan Ministry of Environment.

References

- [1] Prepared for the U.S. Department of Energy by Litos Strategic Communication under contract No. DE-AC26-04NT41817, The Smart Grid: An introduction, 2011.
- [2] L. Hernandez, C. Baladron, and J. M. Aguiar, “A multi-agent system architecture for smart grid management and forecasting of energy demand in virtual power plants,” *IEEE Communications Magazine*, vol. 51, no. 1, pp. 106–113, 2013.
- [3] L. Wang, Z. Wang, and R. Yang, “Intelligent multiagent control system for energy and comfort management in smart and sustainable buildings,” *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 605–617, 2012.
- [4] Z. Jun, L. Junfeng, W. Jie, and H. W. Ngan, “A multi-agent solution to energy management in hybrid renewable energy generation system,” *Journal of Renewable Energy*, vol. 36, no. 5, pp. 1352–1363, 2011.
- [5] A. Karabiber, C. Keles, A. Kaygusuz, and B. B. Alagoz, “An approach for the integration of renewable distributed generation in hybrid DC/AC microgrids,” *Journal of Renewable Energy*, vol. 52, pp. 251–259, 2013.
- [6] S. Misra, A. Mondal, S. Banik, M. Khatua, S. Bera, and M. S. Obaidat, “Residential energy management in smart grid: A markov decision process-based approach,” in *Proceedings of the 2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing, GreenCom-iThings-CPSCOM 2013*, pp. 1152–1157, August 2013.
- [7] T. Logenthiran, D. Srinivasan, A. M. Khambadkone, and H. N. Aung, “Multiagent system for real-time operation of a microgrid in real-time digital simulator,” *IEEE Transactions on Smart Grid*, vol. 3, no. 2, pp. 925–933, 2012.

- [8] M. Pipattanasomporn, H. Feroze, and S. Rahman, "Multi-agent systems in a distributed smart grid: Design and implementation," in *Proceedings of the IEEE/PES Power Systems Conference and Exposition (PSCE '09)*, pp. 1–8, March 2009.
- [9] A. Mehrizi-Sani and R. Iravani, "Potential-function based control of a microgrid in islanded and grid-connected modes," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1883–1891, 2010.
- [10] T. L. Vandoorn, B. Meersman, J. D. M. De Kooning, and L. Vandevelde, "Transition from islanded to grid-connected mode of microgrids with voltage-based droop control," *IEEE Transactions on Power Systems*, vol. 28, no. 3, pp. 2545–2553, 2013.
- [11] C. Essayeh, M. Raiss El-Fenni, and H. Dahmouni, "Load management strategy for an OFF-grid switched mode in smart MicroGrid systems," in *Proceedings of the 2016 International Conference on Wireless Networks and Mobile Communications, WINCOM 2016*, pp. 52–57, October 2016.
- [12] C. Essayeh, M. Raiss El-Fenni, and H. Dahmouni, "Towards an intelligent home energy management system for smart Microgrid applications," in *Proceedings of the 12th IEEE International Wireless Communications and Mobile Computing Conference, IWCMC 2016*, pp. 1051–1056, September 2016.
- [13] R. H. Lasseter and P. Paigi, "Microgrid: A conceptual solution," in *Proceedings of the 2004 IEEE 35th Annual Power Electronics Specialists Conference, PESC04*, vol. 6, pp. 4285–4290, June 2004.
- [14] D. E. Olivares, A. Mehrizi-Sani, A. H. Etemadi et al., "Trends in microgrid control," *IEEE Transactions on Smart Grid*, vol. 5, no. 4, pp. 1905–1919, 2014.
- [15] Y. Wu, V. K. N. Lau, D. H. K. Tsang, L. P. Qian, and L. Meng, "Optimal energy scheduling for residential smart grid with centralized renewable energy source," *IEEE Systems Journal*, vol. 8, no. 2, pp. 562–576, 2014.
- [16] B. B. Alagoz, A. Kaygusuz, and A. Karabiber, "A user-mode distributed energy management architecture for smart grid applications," *Energy*, vol. 44, no. 1, pp. 167–177, 2012.
- [17] M. A. A. Pedrasa, T. D. Spooner, and I. F. MacGill, "Coordinated scheduling of residential distributed energy resources to optimize smart home energy services," *IEEE Transactions on Smart Grid*, vol. 1, no. 2, pp. 134–143, 2010.
- [18] M. Pipattanasomporn, M. Kuzlu, and S. Rahman, "An algorithm for intelligent home energy management and demand response analysis," *IEEE Transactions on Smart Grid*, vol. 3, no. 4, pp. 2166–2173, 2012.
- [19] A. Di Giorgio and F. Liberati, "Near real time load shifting control for residential electricity prosumers under designed and market indexed pricing models," *Applied Energy*, vol. 128, pp. 119–132, 2014.
- [20] S. A. Pourmousavi, S. N. Patrick, and M. H. Nehrir, "Real-time demand response through aggregate electric water heaters for load shifting and balancing wind generation," *IEEE Transactions on Smart Grid*, vol. 5, no. 2, pp. 769–778, 2014.
- [21] A. Soares, A. Gomes, C. H. Antunes, and C. Oliveira, "A customized evolutionary algorithm for multiobjective management of residential energy resources," *IEEE Transactions on Industrial Informatics*, vol. 99, 2016.
- [22] N. Hegde, L. Massoulié, and T. Salonidis, "Optimal control of residential energy storage under price fluctuations," *ENERGY*, 2011.
- [23] Q. Wei, D. Liu, and G. Shi, "A novel dual iterative Q-learning method for optimal battery management in smart residential environments," *IEEE Transactions on Industrial Electronics*, vol. 62, no. 4, pp. 2509–2518, 2015.
- [24] D. Li and S. K. Jayaweera, "Machine-Learning Aided Optimal Customer Decisions for an Interactive Smart Grid," *IEEE Systems Journal*, vol. 9, no. 4, pp. 1529–1540, 2015.
- [25] <http://www.inpt.ma/Pages/Actualite.aspx?ActuID=199>.
- [26] P. Pinson, "Wind energy: forecasting challenges for its operational management," *Statistical Science*, vol. 28, no. 4, pp. 564–585, 2013.
- [27] C. Candelise, M. Winkler, and R. J. K. Gross, "The dynamics of solar PV costs and prices as a challenge for technology forecasting," *Renewable & Sustainable Energy Reviews*, vol. 26, pp. 96–107, 2013.
- [28] A. Gosavi, "On step sizes, stochastic shortest paths, and survival probabilities in Reinforcement Learning," in *Proceedings of the 2008 Winter Simulation Conference (WSC)*, pp. 525–531, Miami, FL, USA, December 2008.
- [29] "NREL pv watts calculator," <http://pvwatts.nrel.gov/pvwatts.php>.
- [30] M. Lichman, *UCI machine learning repository*, 2013, <http://archive.ics.uci.edu/ml>.
- [31] M. L. Puterman, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, Wiley Series in Probability and Mathematical Statistics: Applied Probability and Statistics, Wiley-Interscience, 2005.



Hindawi

Submit your manuscripts at
www.hindawi.com

