Electrical Power and Energy Systems 48 (2013) 148–160

Contents lists available at SciVerse ScienceDirect

Electrical Power and Energy Systems

journal homepage: www.elsevier.com/locate/ijepes

Action dependent heuristic dynamic programming for home energy resource scheduling

Danilo Fuselli [a], Francesco De Angelis [a], Matteo Boaro [a], Stefano Squartini [a,*], Qinglai Wei [c], Derong Liu [b], Francesco Piazza [a]

[a] Dipartimento di Ingegneria dell'Informazione, Universitá Politecnica delle Marche, Ancona, Italy
[b] Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, USA
[c] State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, China

ARTICLE INFO

Article history:
Received 10 July 2012
Received in revised form 17 November 2012
Accepted 25 November 2012
Available online 24 January 2013

Keywords:
Adaptive dynamic programming
Action dependent heuristic dynamic programming
Particle swarm optimization
Neural networks
Smart grid
Home energy management

ABSTRACT

Energy management in smart home environment is nowadays a crucial aspect on which technologies have been focusing on in order to save costs and minimize energy waste. This goal can be reached by means of an energy resource scheduling strategy provided by a suitable optimization technique. The proposed solution involves a class of Adaptive Critic Designs (ACDs) called Action Dependent Heuristic Dynamic Programming (ADHDP) that uses two neural networks, namely the Action and the Critic Network. This scheme is able to minimize a given Utility Function over a certain time horizon. In order to increase the performances of the ADHDP algorithm, suitable Particle Swarm Optimization (PSO) based procedures are used to pretrain the weights of the Action and the Critic networks. The results provided by PSO techniques and by a non-optimal baseline approach are also used as elements of comparison. Computer simulations have been carried out in different residential scenarios. An historical data set for solar irradiation has been used to simulate the behavior of a photovoltaic array to obtain renewable energy and the main grid is used to supply the load and charge the battery when necessary. The results confirm that the ADHDP is able to reduce the overall energy cost with respect to the baseline solution and the PSO techniques. Moreover, the validity of this method has also been shown in a more realistic context where only forecasted values of solar irradiation and electricity price can be used.

1. Introduction

In the last decade the price of oil and others fossil fuels has quickly increased and most of the world countries have been developed new policies to reduce the energy costs and the pollution impact. There is a certain variety of renewable resources that can provide energy production; however wind and solar power sources are undoubtedly the most used worldwide.

The former has recently increased its diffusion not only in U.S. [1] but also in several other countries. The main advantages of this technology are the high efficiency of the wind based systems, the absence of gas production, the short payback time and the very low costs of maintenance and dismantling of the wind system. On the other side, the main limitations are represented by the acoustic noise produced by the wind turbines and the fact that in several places the wind availability is sporadic or simply not enough.

The solar power is typically exploited by means of photovoltaic (PV) systems. With respect to other resources, the solar one has some relevant advantages, as the fact that all locations on earth receive predictable solar irradiation. Moreover, the PV systems can be easily scaled from large to small sizes, and they are often employed in home environments because they have no moving parts and they need only a little maintenance. Such systems typically present lower efficiency with respect to the ones exploiting wind power and thus a longer payback period. It follows that an important issue consists in optimizing the usage of produced PV electrical energy and its delivering to the grid in order to reduce the overall energy cost under different operating conditions.

Along this same direction, the aim of aligning the interests of electric utilities, consumers and environmentalists [2–9], the development of advanced technological solutions for the *smart grid* has encountered an increasing success within the research community in the last years. The advanced technologies find application in various grid environments, going from the micro scenarios, like houses and neighborhoods, to the large scale systems for production and distribution. Moreover, several commercial products have already appeared and more will come in the future, as confirmed by recent market trends [10].

This work deals with a domestic scenario where an electrical grid connection, a PV system and a suitable storage system (called

* Corresponding author. Tel.: +39 0712204381; fax: +39 0712204464.
E-mail address: s.squartini@univpm.it (S. Squartini).

0142-0615/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.ijepes.2012.11.023

simply battery from now on) coexist. The goal consists in satisfying the electrical load requirements over time by optimally managing the electrical power produced from the PV panel, the battery charge/discharge actions and the amount of electrical power acquired from the main grid. In this way the overall energy costs can be reduced.

In this paper we name the addressed problem as *Energy Resource Scheduling*, to differentiate it from the *Energy Task Scheduling* (or Energy Consumption Scheduling) one, whose main objective consists in allocating the temporal activity of household appliances according to some strategy. The large variety of solutions recently proposed for the Energy Task Scheduling issue, like [11–13] do not seem to include an optimal Energy Resource Scheduling policy therein. It represents thus an interesting topic to investigate in the near future.

Focusing on the specific problem addressed in this paper, several techniques have been proposed in the literature. A dynamic programming approach is used in [14,15] whereas a genetic algorithm is proposed in [16]. Moreover, Liu and Huang [17] have recently advanced an Adaptive Dynamic Programming (ADP) basic scheme using only a Critic Network and considering only three possible controls for the battery (charging mode, discharging mode, idle) choosing the best for every time slot. Venayagamoorthy and Welch also made use of the ADP paradigm to perform an Energy Resource Scheduling strategy in isolated electrical systems [18–20].

Up to the authors' knowledge, the most performing approach to deal with the Energy Resource Scheduling problem is the Particle Swarm Optimization (PSO), available in the literature in its online version: given a Utility Function to minimize, given some constraints, through certain "particles" that explore all the solutions space, the algorithm is able to find the best solution for the problem under test. Gudi et al. in [21] proposed an optimal management of renewable resources with PSO using an Utility Function able to charge the battery only from the PV system, not considering the charging from the main grid; furthermore the battery is constrained to discharge itself only in fixed "peak hours", when the electricity price or the load demand are high. Although this method provides a solution at a low computational cost, it is not able to work over an extended temporal horizon, but only step by step.

To reach the same goal an operational scheme with self-learning ability can be developed to optimize home energy systems according to system configuration and user demand. This method is based on a class of Adaptive Critic Designs (ACDs) called Action Dependent Heuristic Dynamic Programming (ADHDP) and it has the capability to learn from the environment [22]. The ADHDP uses two neural networks, an Action Network (which provides the control signals) and a Critic Network (which criticizes the Action Network performance). An optimal control policy is evolved by the Action Network over a time period using the feedback signals provided by the Critic Network. The goal of the control policy is to minimize the amount of energy imported from the grid, represented in a Utility Function, according to some given constraints, managing the battery action and knowing the state of the system in terms renewable energy resources, load profile and electricity price. The chosen Utility Function must be an index of system costs, so that when such a function is minimized the Action Network is able to provide an optimal control to keep the energy costs over the work horizon low. From this perspective and at the light of the theoretical properties of the ADHDP optimization framework [23,22], the implemented strategy for smart home Energy Resource Scheduling can be considered optimal.

The innovative solution advanced in this paper (and preliminary investigated in [24]) is to consider an ADHDP based algorithm where the neural networks weights are pre-trained with online or offline PSO algorithms. The first one, mentioned above, cannot work over an extended horizon but only step by step, while the second one, here proposed for the first time, has to minimize an extended Utility Function and it can work over a larger time horizon. This combined technique, compared with a baseline approach, is able to provide relevant results in terms of monetary saving. The proposed PSO methods are used in this paper to pre-train the ADHDP neural networks and also as element of comparison in performed computer simulations.

Computer simulations show that results provided by the ADHDP algorithm, pre-trained with PSO Offline, are the best in terms of saving, if compared to ADHDP pre-trained with PSO Online. Naturally the results outperform also the ones obtained with the baseline approach and the PSO algorithms. All performed simulations are made considering historical data for solar irradiation, but in order to make the simulated scenario closer to real case studies also uncertain forecasted data have been considered. In this case the uncertainty of the data affects the final results and reduce the attainable saving; anyway the effectiveness of the approach is still ensured.

Moreover it must be said that the proposed algorithm is much more general of the aforementioned recent ADP based techniques for smart home energy management. Indeed, on one side it allows to apply a continuous control to the battery activity, rather than the discrete control in [17], and, on the other, it is able to work also in presence of grid-connected system, taking the electricity price into account, which is not the case of the algorithms in [18–20].

Here the outline of the paper follows. Section 2 describes the considered energy system in home environment, Section 3 proposes the PSO scheme with a brief theoretical preface. In Section 4 the optimal-control ADHDP scheme is shown, the used neural networks are described and the training algorithm is explained. The computer simulations, carried out using both historical and forecasted data for solar irradiation and unitary energy prize, are described in Section 5 and related results reported therein. Conclusions are drawn in Section 6, where some future works ideas are also highlighted.

2. Simulated home energy system

The simulated home energy system is composed of four different parts: main electrical grid, external PV array, the battery and a Power Management Unit (PMU). The main electrical grid and the external PV array (able to convert solar energy in electrical one) can supply the load and/or charge the battery. Furthermore the energy in surplus from photovoltaic system can be sold to the grid. On the other hand the battery system can operate in one of the following modes:

- Charge from the grid and/or from the PV: the battery is charged with an energy quantity according to the battery charging rate (Table 1).
- Discharge: the battery supplies the load discharging itself of an energy quantity according to battery discharging rate (Table 1).
- Inactive: no energy quantity is exchanged with external systems.

As reported in Fig. 1, the PMU unit manages the energy flows described above, assuring the meeting of load demand over time.

Table 1
Battery parameters. Operating voltage is assumed to be equal to 48 V.

η	BL_0	BL_{MIN}	BL_{MAX}	Ch_{rate}/Dh_{rate}
95%	5 kW h	0.5 kW h	10 kW h	±1 kW

Another feature considered in the system is the possibility to sell to the main grid exceeded energy from PV, not usable from the PMU because the battery is already full and the load totally satisfied. Note that in the present scenario the energy of the battery cannot be directly sold to the grid, in compliance with the existing regulations in may western countries.

The battery and PV array sizing is a relevant issue in any PV energy system [25–28], and it has not been addressed in this work, but left to future investigations. The simulated PV array covers an area of 65 m^2, and the efficiency equal to ρ = 15%: these values will be used in the following to estimate the amount of renewable energy produced within the considered working horizons. The main parameter values of the battery model used in this work, similar to the one in [29], are reported in Table 1: η is the battery efficiency, BL_0 is the initial energy level of the battery, BL_{MAX} and BL_{MIN} are respectively the maximum and minimum energy level of the battery and Ch_{rate}/Dh_{rate} refers to the maximum charge/discharge rate value. In addition the battery must satisfy the following constraints:

(1) The charge and discharge rate cannot be exceeded.
(2) Battery level must be always between BL_{MIN} and BL_{MAX}.

3. PSO algorithms

In this section the battery management implemented with PSO method [30,31] is introduced. PSO is a technique developed by Eberhart and Kennedy [32,33] and inspired by certain social behaviors exhibited in bird and fish groups that is used to explore a solutions space for finding parameters that are required to optimize a specific aspect of the problem. PSO is a computational intelligence-based technique that is not largely affected by the size and nonlinearity of the problem, and it can converge to the optimal solution in many kinds of problems where most analytical methods fail to converge.

The PSO algorithm works by maintaining simultaneously various candidate solutions (particles in the swarm) in the admitted solutions space. An attractive feature of the PSO approach is its simplicity as it involves only two model equations. In PSO, the coordinates of each particle represent a possible solution associated with two vectors: the position and velocity vectors in N-dimensional solutions space. A swarm consists of a number of particles or possible solutions that fly through the feasible solutions space to find the optimal one. Each particle updates its position based on its own best exploration, best swarm overall experience, and its previous velocity vector according to (1) and (2).

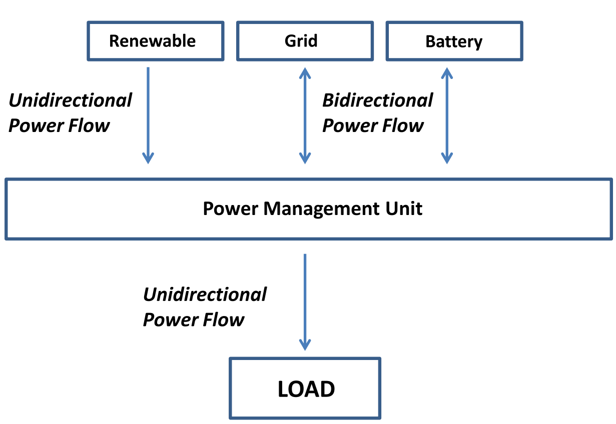

Fig. 1. Power flows within the home energy system.

The movement of each particle naturally evolves to an optimal or near-optimal solution. The position of each particle i in the solutions space is represented by the $x_i(t)$ vector and its movement by the velocity vector $v_i(t)$.

$$x_i(t) = x_i(t-1) + v_i(t) \tag{1}$$

$$v_i(t) = w * v_i(t-1) + \rho_1 * \; rand_1 * (p_i - x_i(t-1)) + \rho_2 * \; rand_2 * (p_g - x_i(t-1)) \tag{2}$$

where t is the time step, i refers to $particle - i$ in the swarm, w is the "inertia" factor, ρ_1, ρ_2 are two positive numbers (correction factors) and $rand_1$, $rand_2$ are two random numbers with uniform distribution in the range of [0.0, 1.0]. The position of each particle is determined by the vector and its movement by the velocity of the particle. The velocity update equation in (2) has three components:

(1) The first component is sometimes referred as "inertia", "momentum" or "habit": it models the tendency of the particle to continue along the same direction it has been traveling.
(2) The second component is a linear attraction towards the best position ever found by the given particle p_i (the best of particle), scaled by a random weight $\rho_1 * rand1$: this component is related to "memory", "self-knowledge", "nostalgia" or "remembrance".
(3) The third component is a linear attraction towards the best position found by any particle: p_g (global best), scaled by another random weight $\rho_2 * rand2$: this component is related to "cooperation", "social knowledge", "group knowledge" or "shared information".

The PSO algorithm can be described in general as follows:

(1) For each particle, randomly initialize the position and velocity vectors with the same size as the problem dimension N.
(2) Measure the fitness of each particle and store the particle with the best fitness value.
(3) Update velocity and position vectors according to (1) and (2) for each particle.
(4) Repeat steps 2 and 3 until a termination criterion is satisfied (maximum number of iterations or a good fitness value is reached).

Every time that step 3 is done the algorithm proceeds by adjusting the velocity with whom each particle moves in every dimension of the problem hyperspace.

The velocity of the particle is a stochastic variable, so it is subject to create an uncontrolled trajectory, making the particle not useful anymore. In order to damp these oscillations, upper and lower limits can be defined for the velocity:

$$\text{if} \quad v_i(t) > v_{max} \quad \text{then} \quad v_i(t) = v_{max} \tag{3}$$

$$\text{if} \quad v_i(t) < -v_{max} \quad \text{then} \quad v_i(t) = -v_{max} \tag{4}$$

Most of the time, v_{max} value is selected empirically, according to the characteristics of the problem. It is important to note that if the value of this parameter is too high, then the particles may move erratically, going beyond a good solution; on the other hand, if it is too small, then the movement of the particle is too much limited and the optimal solution may not be reached. Correction factors ρ_1, ρ_2 in (2) control the movement of each particle towards its individual and global best position, respectively. Small values limit the movement of the particles, while large numbers may cause the particles to diverge.

Two different methods for energy management have been developed by using the PSO paradigm: the PSO Online and the PSO Offline. The former is an adapted version of the algorithm in [21], whereas the latter is originally proposed in this work.

3.1. PSO Online

The PSO Online is an online optimization method, it computes the optimal control every time step t knowing only the current state. Similar to the work done in [21] we introduce the utility function that must be minimized in (5):

$$U(t) = \sqrt{\psi_1^2 + \psi_2^2} \tag{5}$$

where

$$\begin{aligned} &\psi_1 = (L(t) - R(t) + u(t)) * C(t)/min(C(t)) \\ &\psi_2 = BL_{cap} - (BL(t) + u(t)) \\ &with \qquad 1 < t < T \end{aligned}$$

and where $L(t)$, $R(t)$, $C(t)$, $min(C(t))$, BL_{cap}, $BL(t)$ are respectively the current load, renewable energy, grid electricity price, minimum grid energy prize, battery capacity and battery energy level; while $u(t)$ is the value of battery charge ($u(t) > 0$) or discharge ($u(t) < 0$) at each time t and T is the considered time horizon.

Since the Utility Function used in [21] does not consider the battery charging from main grid, the equation is transformed to (5): in this way the battery can charge itself from the grid and/or the renewable energy and discharge itself to supply all the load demand or a part of it considering the energy unitary price from the grid. The utility function is composed by two terms, ψ_1 refers to discharging the battery while ψ_2 refers to charging the battery. Minimizing $U(t)$ means charging the battery when renewable is high and/or when energy grid price is low, while discharging the battery when renewable is lower than the load and/or the energy grid price is high.

It must be noticed that $U(t)$ considers only current values at time t and it does not need forecasted data to operate, so it represents a good trade-off between computational cost and performance. In PSO Online the following parameters are used:

- Search space dimension $N = 1$.
- Swarm size = 25.
- Iterations = 30.
- Inertia factor $w = 0.7$.
- First correction factor $\rho_1 = 1.0$.
- Second correction factor $\rho_2 = 1.0$.
- Penalty factor = 10^6.
- Maximum velocity of each particle $v_{max} = (Ch_{rate} - Dh_{rate})/10$.

3.2. PSO Offline

PSO Offline is an optimization method that computes the optimal control for every time step t knowing the current state and also the forecasted data. So the operation is not online anymore, but offline in order to give an optimal solution on an extended period, for which all scenario profiles are considered in the work horizon: forecasted data about renewable energy, load profile and electricity price. The utility function adopted in this case is the following:

$$U(t) = \sum_{t=1}^{T} \sqrt{\{[L(t) - R(t) + u(t)] * C(t)\}^2} \tag{6}$$

It can be observed that, differently from (5), it considers the load, renewable energy and electricity price terms over the entire work horizon T. Minimizing the PSO Offline utility function means better performance with respect to the PSO Online one, because the former works over an extended time horizon, at the cost of an higher computational complexity.

In the PSO Offline algorithm the following parameters are used:

- Search space dimension $N = T$ (horizon).
- Swarm size = 25.
- Iterations = 100.
- Inertia factor $w = 0.7$.
- First correction factor $\rho_1 = 1.0$.
- Second correction factor $\rho_2 = 1.0$.
- Penalty factor = 10^6.
- Maximum velocity of each particle $v_{max} = (Ch_{rate} - Dh_{rate})/10$.

Obviously, both in PSO Online and Offline, $u(t)$ must satisfy battery constraints previously defined in Section 2 and reported in (7) and (8).

$$Dh_{rate} \leqslant \; u(t) \leqslant \; Ch_{rate} \tag{7}$$

$$BL_{min} \leqslant (BL(t) + u(t)) \leqslant \; BL_{max} \tag{8}$$

If one of these constraints is not satisfied, the obtained solution $u(t)$ is not valid and it must be discarded: so the function is multiplied with the penalty factor which is typically set to an high value as above.

4. ADHDP algorithm

In discrete-time nonlinear environments ACDs methods are able to optimize over time in noisy and uncertainty conditions using neural networks. Combining approximate dynamic programming and reinforcement learning, Werbos proposed a new optimization technique [23].

The goal of this technique is to design an optimal control policy, which can be able to minimize a given cost function. This optimal control is obtained adapting two neural networks: the Action Network and the Critic Network. The Action Network, taking the current state, has to drive the system to a desired state, providing a control to the latter. The Critic Network, knowing the state and the control given by the Action Network, has to check its performance and return to the Action Network a feedback signal to reach the optimal state over time. This feedback is used by the Action Network to adapt its parameters in order to improve its performance. To check the Action behavior, the Critic Network approximates the associated Hamilton–Jacobi–Bellman equation.

At the beginning of this adaptive process, the control policy cannot be optimal, but driven by the Critic feedback, the performance of the Action Network improves thus yielding an optimal control policy at the end. One of the main advantage of this method is that it just relies on the minimization of the cost function and thus the networks do not need information about the optimal "trajectory".

Therefore, starting from a given initial state and from initial conditions and constraints, ACDs determine an optimal control policy, without the help of other external training, like other neural-controllers [34]. In [35] the relationship between the main members of ACDs family is known and defined: Heuristic Dynamic Programming (HDP), Dual Heuristic Programming (DHP) [36] and Globalized Dual Heuristic Programming (GDHP), cited in order of complexity.

In this work, load demand and renewable energy are not fixed *a priori*, but they are stochastic and variable: for these reasons, an Action dependent HDP (ADHDP) model-free approach is adopted (Fig. 2) for the design of an optimal battery management system [37], according to the PMU considered. The goal of the optimal system is to manage the battery charging/discharging in order to save costs during the overall time-horizon.

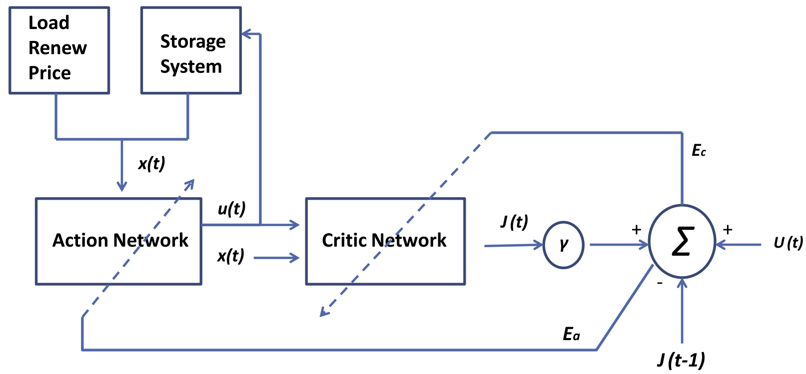

Fig. 2. ADHDP scheme.

As mentioned above, Venayagamoorthy and Welch propose an optimal controller based on the same ADHDP scheme used in this paper, but they consider an isolated scenario in which the load (splitted in critical and not critical) has to be supplied only from a PV system. Connection with main grid and electricity price are not considered in their algorithmic framework, and the goal is to optimize the control policy over time to ensure that primarily the critical load demand and secondarily the non-critical load demand are met [18–20].

The optimal scheme proposed here uses two networks (Action and Critic networks) as previously mentioned. The input to the Action network is the system state, and the output $u(t)$ is the amount of energy used to charge or discharge the battery. This quantity is not a discrete value, used only for describing battery behavior (*Charge*, *Discharge*, *Idle*) like proposed in [17,38], but it is a continuous value that represents the real energy to dispatch, improving the system accuracy. The input of the Critic Network consists of the current system state and the current control provided by the Action Network. In order to extend the temporal horizon and minimize the costs in a longer period, it is possible to consider also the previous states and controls as Critic inputs.

Increasing the number of past inputs the costs are reduced but the computational complexity of the neural network increases, due to an higher number of hidden neurons is needed. In this way, like in [18], a reasonable trade-off consists in inserting only two past states and controls in the Critic. For this reason the Critic Network takes the state and the control at time t, $t-1$, $t-2$ as inputs: this information is used by the network to compute the cost function at time t, used to create the feedback signal for the Action network.

4.1. Critic neural network

As previously mentioned, the inputs of the Critic Network are the system state and the output of the Action Network in three different time-steps ($t, t-1, t-2$). The network structure is here detailed:

- 15 linear input neurons;
- 40 sigmoidal hidden neurons;
- 1 linear output neuron.

The network is trained using standard backpropagation (BP) algorithm. Its output is the estimated cost function at time t, given by Bellman's equation as follows:

$$J(t) = \sum_{i=t}^{\infty} \gamma^{i-t} U(i) \tag{9}$$

The discount factor γ is used for non-infinite horizon problems and it can assume continuous values in the range [0, 1]: in this case a value equal to 0.8 is considered. The Utility Function $U(t)$ is very important for driving the Critic Network to improve the Action Network performance.

When the Utility Function is minimized, the control policy is optimal and the cost is the lowest. The proposed $U(t)$ is the following:

$$U(t) = [(L(t) - R(t) + u(t)) * C(t)]^2 \tag{10}$$

According to [39] the Critic Network weight update is expressed as follows:

$$\Delta W_c(t) = \alpha_c E_c(t) \frac{\partial J(t)}{\partial W_c} \tag{11}$$

$$W_c(t+1) = W_c(t) + \Delta W_c(t) \tag{12}$$

where α_c is the learning rate, W_c are the critic weights and $E_c(t)$ is the critic network error given by (13).

$$E_c(t) = U(t) + \gamma J(t) - J(t-1) \tag{13}$$

4.2. Action neural network

The network structure is the following:

- 4 linear input neurons;
- 40 sigmoidal hidden neurons;
- 1 linear output neuron.

This network is trained using the backpropagation (BP) algorithm. The input of the Action Network is the current state of the system, composed of four components:

- Load ($L(t)$).
- Produced Renewable Energy ($R(t)$).
- Unitary Electricity Price ($C(t)$).
- Battery Level (BL).

The output of the network is the control, i.e. the amount of energy charging/discharging the battery. The current control, $u(t)$, is used to adjust the battery energy level in the subsequent state. The found Action output has to be checked to ensure that the battery bounds (maximum discharge/charge rate, maximum and minimum battery level) are satisfied for each time step. This is obtained by forcing the control to respect the same limits defined for PSO algorithms in (7) and (8).

Similar to (11) and (12) the weights refresh in the Action Network is given by:

$$\Delta W_a(t) = \alpha_a E_a(t) \frac{\partial u(t)}{\partial W_a} \tag{14}$$

$$W_a(t+1) = W_a(t) + \Delta W_a(t) \tag{15}$$

where α_a is the learning rate, W_a are the critic weights and $E_a(t)$ is the action network error expressed as:

$$E_a(t) = U(t) + \gamma J(t) - J(t-1) \quad (16)$$

4.3. Training phase

In this section the iterative training used for both neural networks is explained step by step and represented in the flowchart in Fig. 3.

Step1 : the Action and Critic weights can be initialized before the training in two different ways:

- **Step1.1** : initialize random weights for both the networks (range of values $[-1, 1]$)
- **Step1.2** : initialize weights with a pre-training made with PSO Online or Offline algorithms described in Sections 3.1 and 3.2.

Step2 : train the Critic Network and update the weights using (11) and (12). Then refresh Action Network using (14) and (15).

Step3 : evaluate the system performance computing the total cost that must be minimized in the time horizon. If the cost decreases, the control policy improves, and the new action weights are the best; otherwise, revert to old action weights and add a small random perturbation. Then restart the training from Step 2.

The weights initialization explained in the first step can be carried out in two different ways: with the random initialization (Step 1.1) the size of the values must be within a specific range in order to avoid convergence problems. In fact if the size is too big or too small the number of epochs, for a given learning rate, to reach the optimal weight values can be higher than the maximum chosen one. After several simulations the optimal range found is $[-1, 1]$. If a pretraining mode is chosen (Step 1.2), the initial weights are given after a pretraining made with PSO Online method (SubSection 3.1), optimal for one time slot, or with PSO Offline method (SubSection 3.2), more accurate than the former also in a longer time horizon.

In this way (starting from pretrained and not random weights) is possible to reach optimal battery management in large horizons with ADHDP method, using a lower number of epochs and a lower execution time. So the main difference between the two initializations is the execution time needed to reach optimal results.

The weight update is accomplished as follows:

- For a given time step and the relative state, the non optimal control is computed, using random or pretrained Action Network weights. According to (12) it is possible to calculate the error of the Critic Network and then update its weights.
- Considering the updated critic weights, a new cost function is obtained and according to (15) it is possible to compute the error of the Action Network and then to update its weights.

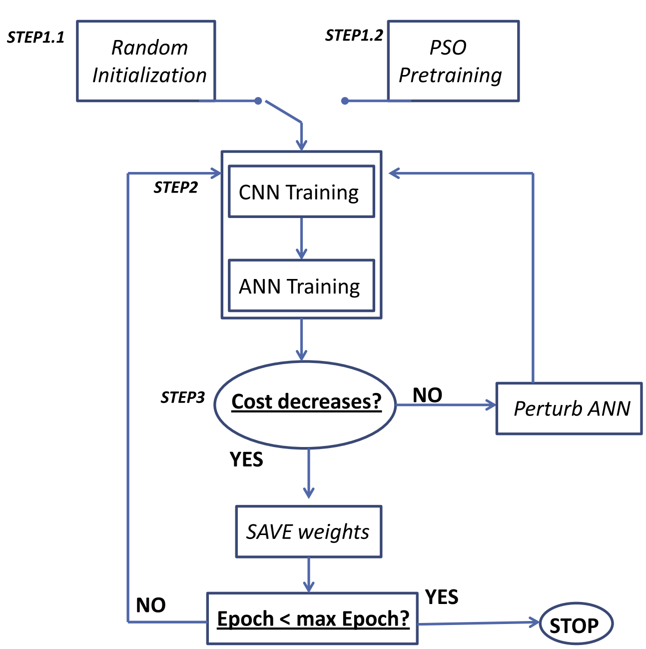

Fig. 3. The ADHDP training algorithm.

These two actions are accomplished for each time step of the total chosen horizon, and when the last step is completed a series of controls is provided. With these controls it is possible to evaluate the performance of the Action Network using several different metrics (Step 3); in this study the total cost is expressed in dollar, and it is calculated in correspondence of those time instants when the system imports energy from the main grid (17):

$$\begin{aligned} &TotalCost = \sum_t (L(t) - R(t) + u(t)) * C(t) \\ &when \qquad (L(t) - R(t) + u(t)) > 0 \end{aligned} \quad (17)$$

After the cost computation the system restarts from Step 2, trying to find a better sequence control for all the time horizon. If the new cost is lower than the smallest one previously obtained, the performance of the Action Network improve and the relative weights are saved as the best; if not, the Action Network performance gets worse and the system reverts to old best weights and adds a small random perturbation.

This random perturbation is confined in the range $[-0.1, 0.1]$ and it is applied in order to avoid that the system persists in a local minimum. After a fixed number of epochs the system is assumed to get a minimum and thus provides its optimal control solution over the chosen time horizon.

5. Computer simulations

In this section the simulation results of ADHDP pre-trained with PSO methods are discussed and compared with the ones obtained with an heuristic baseline approach. The home energy system described in Section 2 is taken as case study for our simulations. The goal is to minimize the total cost of electricity imported from the main grid, over a variable time horizon, finding the optimal battery operational strategy of the energy system satisfying load conditions and battery constraints.

The home scenario proposed in this paper can be easily extended to a residential one, with a different load profile over time, in which all the users share the same storage and PV systems. In the scenario under study, the cost to be minimized is a function of unitary electricity price and load demand. The optimal battery operation strategy refers to the quantity of energy that charges or discharges the battery, in order to achieve the lowest electricity cost for the user. The computer simulations are made using historical and forecasted data, over an horizon of 48 and 96 h in the former case (Section 5.4) and 24 h in the latter (Section 5.5).

5.1. Load profile

The normalized load and unitary price profiles are taken from [17]. In this study, the optimal scheduling problem is viewed as a discrete time problem with the time step of 1 h and it is assumed that the residential load over each time step is varying hourly with noise. In this way the daily load profile is divided into 24 h periods,

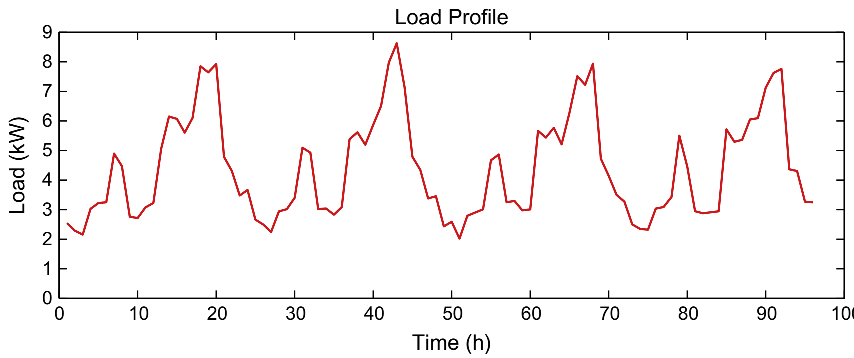

Fig. 4. The load profile used in simulations.

representing each hour of the day. It is possible to increase the time resolution, but for the sake of coherence with several papers in current literature [40–43], the slots used for every day are 24. A typical load profile is shown in Fig. 4, where $L(t)$ expresses the load dependence on time variable t over the chosen horizon of 96 h ($t = 1, 2, \ldots, 96$).

5.2. Electricity price profile

In order to improve power system efficiency and allow new power system construction projects [44], in this study a stationary price profile is not used.

In several countries, the unitary price of electricity ($/kW h) is different over time: for example in several European countries the daytime is split into *hourly price bands* in which the price is different. For instance, during the night, when the energy demand is low, the band provides a lower price than during the day, when the energy demand is higher.

In this work a real-time pricing (ordinary in US) is used to shift electricity usage from peak load hours to light load hours and the electricity price varies from hour to hour based on sale market prices. Hourly, market-based electricity prices typically change as the demand for electricity changes: higher demand usually means higher hourly prices. There is usually a small price spike in the morning and another slightly larger spike in the evening when the corresponding demand is high. Fig. 5 demonstrates a typical real-time pricing from [45] in a 96 h horizon. The varying electricity price is expressed as *Cost*, in dollar cents. Computer simulations discussed in the following Subsection have been carried out considering both historical and forecasted data for the electricity price.

To reduce the total electricity cost, the optimal battery management system in general has to charge the battery when the electrical rate is low and the PV energy is available, and discharge it when it is high, in order to avoid energy purchase from the grid.

5.3. Solar irradiation profiles

To evaluate the performances of the proposed ADHDP battery management system in different operating conditions, in this section different solar irradiation scenarios are introduced, maintaining the same load and unitary price profiles described above. Solar Irradiation (SI) is the amount of direct and diffused solar energy received on a horizontal surface during a 60-min period. SI is expressed in kW h/m^2.

For computer simulations carried out in this work, four United States cities, with different climatic conditions and solar irradiation profiles, have been chosen and related data taken from [46]. The cities are San Diego (California), Columbia (South Carolina), Austin (Texas) and Seattle (Washington). Employed data are related to the first four days of April.

According to the size and efficiency values of the PV system mentioned in Section 2, it follows that the available energy at the output of the PV system $R(t)$ can be expressed as:

$$R(t) = A * \rho * SI(t) \tag{18}$$

where $SI(t)$ is the solar irradiation at time t and A is the area of the PV system. Fig. 6 thus reports the produced energy profiles for the four case studies, according to the PV size and efficiency characteristics mentioned above. It can be observed that Seattle presents the lowest maximum value of renewable energy produced by the PV array, likely due to its geographical position.

5.4. Experimental tests

In this subsection the performed tests are described. The following illustrations have been obtained by means of the ADHDP energy management system pre-trained with the PSO Online and Offline techniques, considering *Load* as the quantity of energy required, *Renew* as the renewable energy, *Cost* as grid unitary electricity price and *BL* as the battery energy level.

Considering the battery model described in Table 1, the first simulations are those where the PSO methods are used to pre-train the ADHDP neural networks: the simulation results are shown in Fig. 7 and they refer to cities with low produced energy profiles in the observation period (like Austin, as shown in Fig. 6). In the figure the PSO performance is reported to evaluate the improvements obtained with the ADHDP optimization scheme, shown by the differences between the *BL* graphs. The chosen horizon is 48 h, but a similar behavior can be registered also in longer time scenarios.

In this scenario there is a big difference between the battery management provided by ADHDP pre-trained with PSO Online if compared with the one obtained with PSO Online considered as standalone method, as evident in the upper illustrations of Fig. 7. The latter charges the battery when the price is low and when renewable energy is high, but it discharges at non convenient times because it does not make any assumptions about the future system states, whereas the former does and thus is able to provide an optimal control policy. In fact, the PSO Online approach minimizes (5) step by step, instead the ADHDP manages the overall time horizon. For these reasons the ADHDP optimization is able to provide an high saving.

In Fig. 7 the main noticeable differences are in the time range 30–35 h, when the battery is not discharged with ADHDP optimization, but in this range the system supplies the load using the main grid. Only when the price is higher, in 40–45 h and when

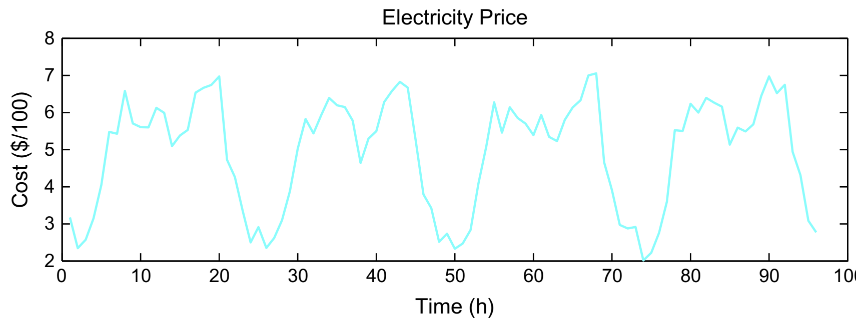

Fig. 5. The unitary electricity price profile used in simulations.

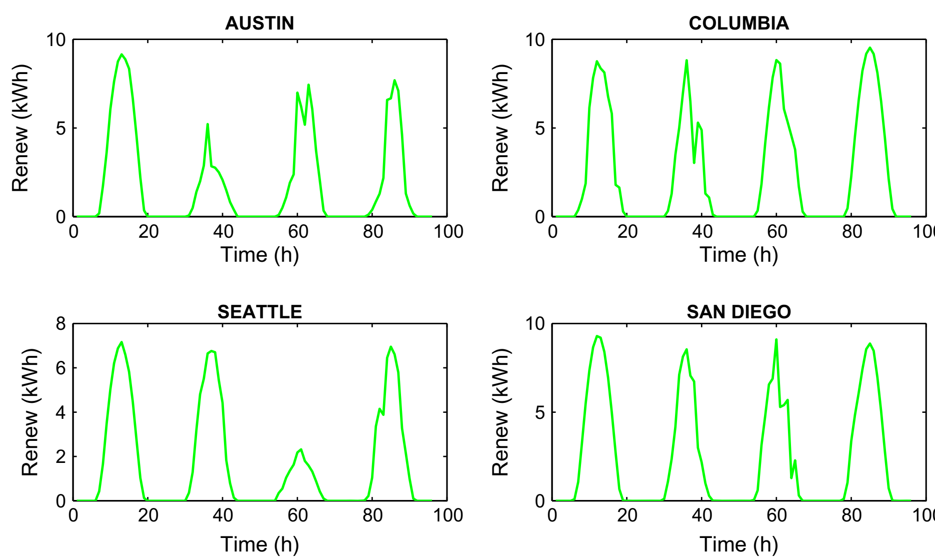

Fig. 6. Renewable energy profiles ($R(t)$) relative to the four different case studies.

there are no renewable resources the battery is discharged, in order to increase the saving.

Regarding the offline approach, the proposed simulation shows that the ADHDP optimization is able to outperform the PSO Offline performance only in correspondence of those hours when the energy unitary price is low. Indeed in these cases there is and high battery charging and the saving increases.

The next simulations involve a city with high solar irradiation, like San Diego, considering a time horizon of 96 h. In this case the optimization obtained with the ADHDP method cannot provide remarkable differences in the battery management in both online and offline approaches. This is due to the fact that in environments characterized by high solar irradiation, the differences in the battery management are attenuated and the saving provided by ADHDP method is limited. In fact a quite good optimization can be obtained also considering only one time step, like made by PSO Online technique. Finally, starting from a pre-trained with PSO weights, the ADHDP training phase is able to decrease the total cost by increasing the training epochs.

Fig. 8 shows that if the PSO online approach is considered, the ADHDP optimization provides a battery management very closed to the PSO one: the only remarkable difference is visible in the range 45–50 h, when the ADHDP charges the battery due to the low energy rate. Another slight improvement occurs in the range 55–60 h, when the ADHDP uses the main grid to supply the load, and discharge the battery when the electricity unitary price is higher, like in the range 90–96 h. This behavior is maintained if an offline approach for ADHDP Networks pre-training is employed.

From the lower illustrations of Fig. 8 it is possible to note that the battery management provided by ADHDP pre-trained with PSO Offline outperforms it only in the range 0–10 h and 70–80 h, in which the PSO in not able to discharge the battery when the unitary electricity price is low and performs it instead in correspondence of less convenient time steps.

After these simulations, and according with Table 2, it is possible to observe that the offline approach for the ADHDP optimization provides better results than the online one. As mentioned, increasing the number of past states as critic inputs, the ADHDP method can improve its performances, at the cost of an increasing computational complexity. These results are better explained in Table 2, where all costs are reported and compared with the ones obtained by means of the baseline algorithm. This baseline algorithm, applied in the same scenario described above, consists of the following simply rules (applied at each time step):

- If the load is greater than the available renewable energy, the energy demand is supplied discharging the battery (according with the discharge rate Dh_{rate} constraint). If the battery help is not enough, the needed energy to supply totally the load is imported from the main grid.

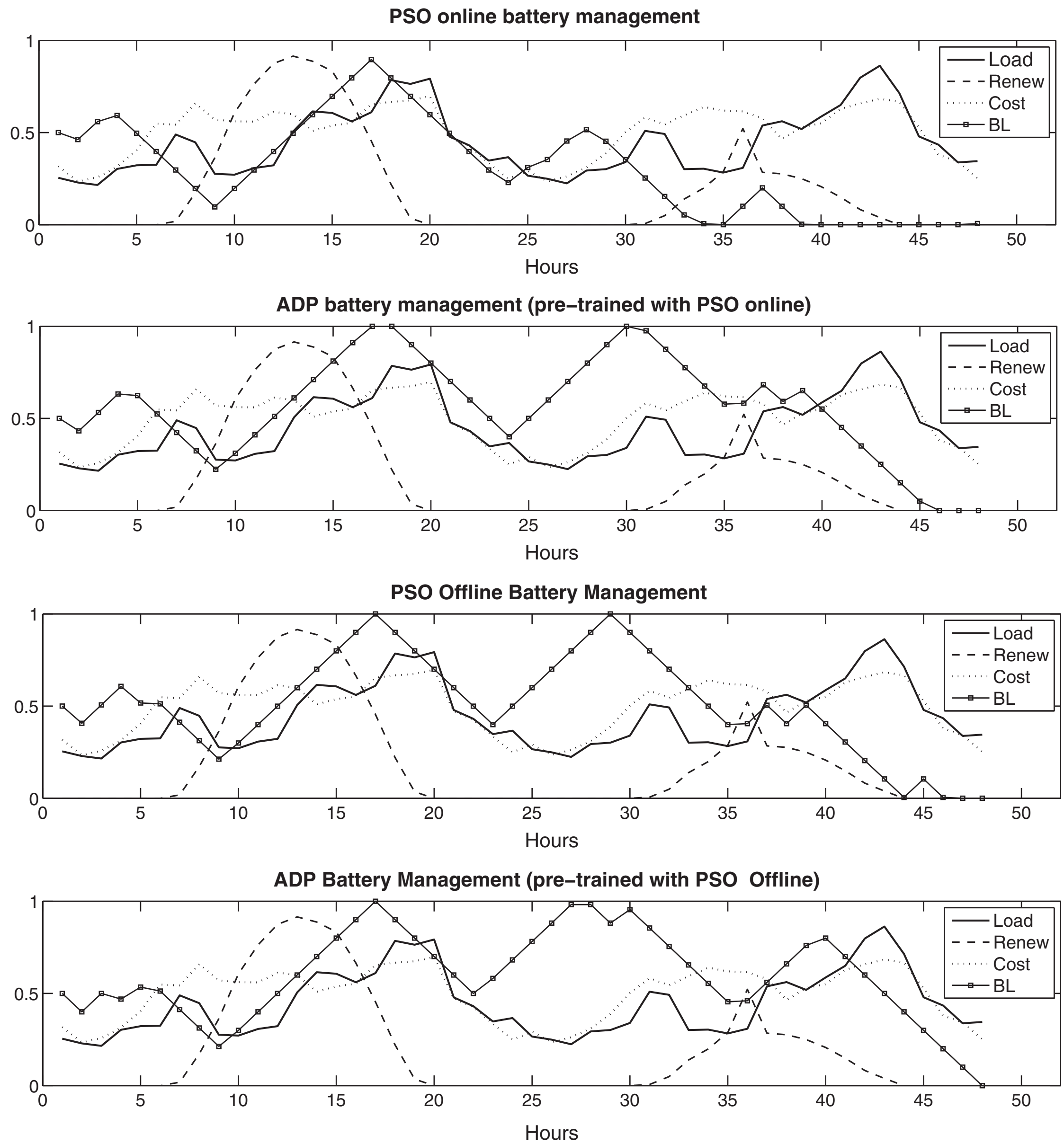

Fig. 7. Computer simulations in the Austin case study. The horizon is 48-h. Each plot is relative to a different optimization algorithm and reports the normalized values (in the range [0, 1]) of load, battery energy, produced renewable energy and cost profiles.

- If the available renewable energy is greater than the load demand, the surplus is used to charge the battery (according with the charging rate Ch_{rate} constraint). If the battery is already full, or the surplus is greater than the charging rate, the remained energy, not usable in other ways, is sold to the main grid.

The baseline method performance, expressed in terms of energy cost in dollars, is reported in Table 2 and used as element of comparison to evaluate the effectiveness of the ADHDP based techniques. Also the costs relative to the PSO based methods are provided. Table 3 reports the savings in percentage obtained with the ADHDP and PSO algorithms with respect to the baseline approach. In Tables 2 and 3 *PSO On* refers to PSO Online algorithm, *PSO Off* to PSO Offline algorithm, *ADP On* is related to ADHDP version pre-trained with PSO Online and *ADP Off* to the version pre-trained with PSO Offline.

It is therefore evident that the offline techniques outperform the sample-by-sample ones, because the optimization is accomplished on a wider temporal range. Moreover the ADHDP scheme allows achieving a significant improvement with respect to the PSO based algorithms, thus justifying its usage and effectiveness for optimal home energy management.

Finally, similar conclusions can be drawn if different battery sizes and characteristics are considered. Indeed the ADHDP algorithmic framework always shows its superiority with respect to the other techniques in terms of percentage savings also by varying the battery operating conditions. Related results are not reported for the sake of conciseness and future works will address the important battery and PV array dimensioning issue, taking the presence of the Energy Resource Scheduling algorithm into account.

5.5. ADHDP performances with forecasted data

The ADHDP proposed scheme is able to work over an extended time horizon, using historical irradiation data taken from [46] and historical unitary prices taken from [17]. A more realistic scenario

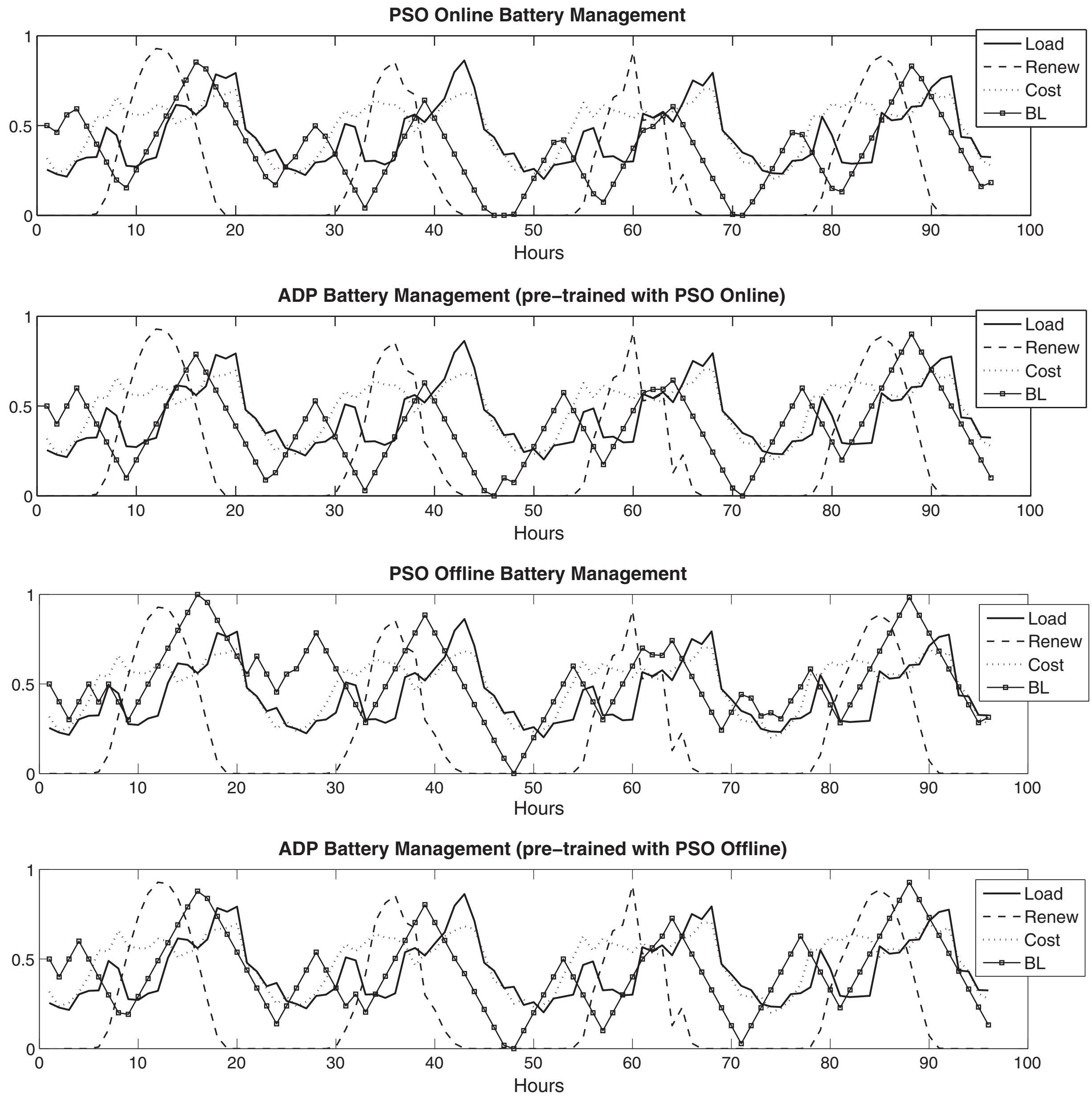

Fig. 8. Computer simulations in the San Diego case study. The horizon is 96-h. Each plot is relative to a different optimization algorithm and reports the normalized values (in the range [0, 1]) of load, battery energy, produced renewable energy and cost profiles.

where the system must work with forecasted data could be considered with the usual aim to minimize the total cost of the imported energy. This is what effectively done in this section, where forecasted irradiation profiles and unitary electricity prices are inserted within the ADHDP optimization scheme and new computer simulations performed to assess the effectiveness of the approach. As preliminary investigation, the forecast is 24-h ahead, but it is possible to consider also shorter horizons to employ more accurate forecasts within the optimization scheme.

According to the literature [47–49] the error that affects the irradiation forecast is not uniform over the year. It depends indeed on the temperature, sky coverage, wind and other factors. The most important one is surely the sky coverage: in sunny days, without clouds and perturbations, the forecast is good, and very closed to the real case. Also in covered days, in which there is lack of sun and the sky is always full of clouds, the error is quite low since the amount of diffuse irradiation can be reasonably predicted. More problematic is the case when the sky is partially covered and direct irradiation is not always present, with sudden changes which are obviously very hard to predict and the forecast error is inevitably higher than the previous two cases. According to these assumptions, the historical data considered above (Fig. 6) are splitted into three categories as follows:

- High SI (Solar Irradiation): for sunny days like San Diego (0–24 h) and Columbia (72–96 h). The prediction error adopted in this case is 5%.
- Medium SI: for uncertain days like Austin (48–72 h) and Seattle (24–48 h). The prediction error adopted in this case is 25%.
- Low SI: for cloudy days like Austin (24–48 h) and Seattle (48–72 h). The prediction error adopted in this case is 10%.

Similar considerations can be made for the forecast of unitary electricity price. According with literature [50,51] the prediction

Table 2
Energy costs obtained by using the baseline procedure and the PSO and ADHDP based algorithms.

City	Horizon (h)	Baseline ($)	PSO On ($)	PSO Off ($)	ADP On ($)	ADP Off ($)
Columbia	1–48	6.39	6.21	6.09	6.11	6.07
	1–96	11.82	11.59	11.48	11.47	11.44
San Diego	1–48	6.17	6.01	5.96	5.96	5.94
	1–96	12.18	11.89	11.78	11.77	11.75
Austin	1–48	7.12	6.97	6.49	6.49	6.48
	1–96	14.09	13.74	12.86	12.96	12.84
Seattle	1–48	6.38	6.25	6.20	6.18	6.14
	1–96	14.52	14.21	13.19	13.22	13.17

Table 3
PSO and ADHDP savings in percentage with respect to the energy cost of the baseline algorithm.

City	Horizon (h)	PSO On (%)	PSO Off (%)	ADP On (%)	ADP Off (%)
Columbia	1–48	2.82	4.69	4.38	5.01
	1–96	1.96	2.87	2.97	3.21
San Diego	1–48	2.59	3.40	3.40	3.89
	1–96	2.37	3.28	3.37	3.53
Austin	1–48	2.11	8.85	8.85	8.99
	1–96	2.49	8.72	8.02	8.87
Seattle	1–48	2.04	2.82	3.14	3.76
	1–96	2.13	9.16	8.96	9.30

Table 4
Energy costs of the baseline, PSO Online and ADHDP over a 24-h horizon. The ADHDP scheme uses forecasted data for training. Savings percentages of the ADHDP with respect to the Baseline algorithm are also provided.

City	Horizon (h)	SI level	Baseline ($)	PSO On ($)	ADHDP ($)	Old Sav. (%)	New Sav. (%)
San Diego	1–24	High	2.74	2.74	2.69	3.50	1.82
Columbia	73–96	High	2.55	2.55	2.53	3.12	0.79
Austin	49–72	Medium	3.18	3.16	3.06	8.23	3.77
Seattle	25–48	Medium	3.10	3.08	3.03	5.39	2.26
Austin	49–72	Low	4.15	4.12	4.04	4.90	2.65
Seattle	25–48	Low	4.64	4.61	4.53	5.11	2.37

error that affects the price in a day-ahead scenario is in the worst case around 12%.

The goal of the following simulations is to evaluate the robustness of the ADHDP technique when the training is not performed by using historical data but the forecasted ones. In order to suitably evaluate this issue, the following total energy costs, over a 24-h horizon, have been computed:

(1) the one provided by the baseline method, which operates step-by-step.
(2) the one provided by the PSO Online algorithm, which again operates step-by-step.
(3) the one obtained by employing the ADHDP scheme for which forecasted data are used for the training phase (Fig. 3) whereas historical data are obviously considered to evaluate the ADHDP performance over the same horizon.

Forecasted data have been obtained by adding to the historical data a uniformly distributed noise, whose variance is proportional to the aforementioned error percentages, both for solar irradiation (which thus changes with the three levels above) and for the unitary electricity price (fixed to 12%). Of course, due to the data mismatch among the training and testing data, an overall decrease of ADHDP performance is expected with respect to what highleghted in previous subsection (Table 3).

In Table 4 the simulatios results under these operating conditions are reported. The ADHDP cost refers to the total cost in dollars obtained with the training and testing explained in this section. In a short term scenario like this, the PSO Online provides very close results to the baseline approach, as reported. The *Old Saving* refers to the percentage of saving (with respect to the Baseline approach) obtained by using the ADHDP techniques considering the training mode involving the knowledge of historical data over the all horizon, as done in previous Subsection. The *New Saving* instead refers to the saving obtained using the new training mode for ADHDP, in which the optimal controls are obtained by using forecasted data for solar irradiation and unitary electricity price.

As expected though, results reported in Table 4 confirm that savings in this case study decrease, due to the uncertainty given by forecasted data, but we are still able to get a significant improvement with respect to the baseline and PSO Online approaches. This allows us concluding that the employment of the ADHDP paradigm is successful for home energy management purposes even when forecasted data are used for optimization within the work horizon.

6. Conclusions and future works

In this paper, a new kind of energy management system based on ADHDP and pre-trained with PSO Online and Offline algoritms for battery management in smart home, connected to the grid and to a photovoltaic system, have been presented. The obtained results show that ADHDP algorithm has an optimal control policy in battery management if compared with PSO one, and the former outperforms the latter in terms of economic benefits and battery control policy, especially if the external environment provides limited renewable resources. Furthermore, performances of a new version of PSO algorithm have been proved, and the best results are given by the ADHDP technique based on PSO Offline training. The proposed methods represent an interesting way to integrate economic savings and renewable energy sources in micro-grids. In order to make the simulations more closed to the real case also forecasted data are considered instead of the historical one for the solar irradiation and the unitary electricity price. The results shows that the performances remain good also in this case, and that this method can be effectively used in a realistic scenarios.

Future work tends to improve the performances of the proposed method, considering a more complex scenario with others energy renewable sources and more complex system storage. Research field involves the ADHDP scheme with forecasted data, considering shorter predict scenarios, more accurate and closed to the real case. The ADHDP method also can be efficiently involved as support for static controllers used as task scheduler in home environment, providing with its self-learning scheme, an accurate Energy Resource Scheduling strategy to correct mismatches that can raise between the forecasted and the real data.

References

[1] National Renewable Energy Laboratory (NREL) of U.S. Department of Energy. 2010 Renewable energy data book, energy efficiency and renewable energy (EERE), 2010. <http://www.nrel.gov/analysis/pdfs/51680.pdf>.
[2] Amin SM, Wollenberg BF. Toward a smart grid: power delivery for the 21st century. IEEE Power Energy Mag 2005;3(5):34–41.
[3] Garrity TF. Getting smart. IEEE Power Energy Mag 2008;6(2):38–45.
[4] Li F, Qiao W, Sun H, Wan H, Wang J, Xia Y, et al. Smart transmission grid: vision and framework. IEEE Trans Smart Grid 2010;1(2):168–77.
[5] Meliopoulos APS, Cokkinides G, Huang RK, Farantatos E, Choi S, Lee Y, et al. Smart grid technologies for autonomous operation and control. IEEE Trans Smart Grid 2011;2(1):1–10.
[6] Sheble GB. Smart grid millionaire. IEEE Power Energy 2008;6(1):22–8.
[7] Venayagamoorthy GK. Potentials and promises of computational intelligence for smart grids. In: IEEE Power & Energy Society general meeting; 2009. p. 1–6.
[8] Papaioannou I, Purvins A, Tzimas E. Demand shifting analysis at high penetration of distributed generation in low voltage grids. Int J Electr Power Energy Syst 2013;44(1):540–6.
[9] Zayandehroodi H, Mohamed A, Shareef H, Mohammadjafari M. A new approach to power system protection in distribution network with dg units by using radial basis function neural network. Int J Eng Intel Syst 2011;19(2):1–15.
[10] Markets and Markets. Smart grid technology market, analysis & global forecast by hardware, software & communication network technologies (2011–2016); 2012. <http://www.marketsandmarkets.com>.
[11] Castillo-Cagigal M, Gutiérrez A, Monasterio-Huelin F, Caamaño-Martín E, Masa D, Jiménez-Leube J. A semi-distributed electric demand-side management system with pv generation for self-consumption enhancement. Energy Convers Manage 2011;52(7):2659–66.
[12] Matallanas E, Castillo-Cagigal M, Gutiérrez A, Monasterio-Huelin F, Caamaño-Martín E, Masa D, et al. Neural network controller for active demand-side management with pv energy in the residential sector. Appl Energy 2012;91(1):90–7.
[13] Mohsenian-Rad A, Leon-Garcia A. Optimal residential load control with price prediction in real-time electricity pricing environments. IEEE Trans Smart Grid 2010;1(2):120–33.
[14] Riffonneau Y, Bacha S, Barruel F, Ploix S. Optimal power flow management for grid connected pv systems with batteries. IEEE Trans Sustain Energy 2011;2(3):309–20.
[15] Maly DK, Kwan KS. Optimal battery energy storage system (BESS) charge scheduling with dynamic programming. IEE Proc Sci Meas Technol 1995;142(6):453–8.
[16] Changsong C, Shanxu D, Tao C, Bangyin L, Huazhong Y. Energy trading model for optimal microgrid scheduling based on genetic algorithm. In: IEEE 6th international power electronics and motion control conference; 2009. p. 2136–9.
[17] Huang T, Liu D. Residential energy system control and management using adaptive dynamic programming. In: The 2011 international joint conference on neural networks (IJCNN). IEEE; 2011. p. 119–24.
[18] Welch RL, Venayagamoorthy GK. Optimal control of a photovoltaic solar energy with adaptive critics. In: Proceedings of the international joint conference on neural networks (IJCNN); 2007. p. 985–90.
[19] Welch RL, Venayagamoorthy GK. Hdp based optimal control of a grid independent PV system. In: IEEE Power Engineering Society general meeting; 2006. p. 1–6.
[20] Welch RL, Venayagamoorthy GK. Comparison of two optimal control strategies for a grid independent PV system. In: IEEE Power Engineering Society general meeting, vol. 3; 2006. p. 1120–7.
[21] Gudi N, Wang L, Devabhaktuni V, Depuru S. A demand-side management simulation platform incorporating optimal management of distributed renewable resources. In: Power systems conference and exposition (PSCE), 2011 IEEE/PES. IEEE; 2011. p. 1–7.
[22] Murray JJ, Cox CJ, Lendaris GG, Saeks R. Adaptive dynamic programming. IEEE Trans Syst Man Cybernet 2002;32(2):140–53.
[23] Werbos PJ. In: White D, Sofge D, editors. Approximate dynamic programming for real-time control and neural modeling. New York: Van Nostrand; 1992 [chapter 13].
[24] Fuselli D, De Angelis F, Boaro M, Liu D, Wei Q, Squartini S, et al. Optimal battery management with ADHDP in smart home environments. In: Advances in neural networks – ISNN 2012, vol. 7367, Part I, LNCS series; 2012.
[25] Lo CH, Anderson MD. Economic dispatch and optimal sizing of battery energy storage systems in utility load-leveling operations. IEEE Trans Energy Convers 1999;14(3):824–9.
[26] Ru Y, Kleissl J, Martinez S. Storage size determination for grid-connected photovoltaic systems. IEEE Trans Sustain Energy 2012:1–14.
[27] Tan X, Li Q, Wang H. Advances and trends of energy storage technology in microgrid. Int J Electr Power Energy Syst 2013;44(1):179–91.
[28] Chen T, Hsieh T, Yang N, Yang J, Liao C. Evaluation of advantages of an energy storage system using recycled EV batteries. Int J Electr Power Energy Syst 2013;45(1):264–70.
[29] Zhang D, Papageorgiou L, Samsatli N, Shah N. Optimal scheduling of smart homes energy consumption with microgrid. In: ENERGY 2011. The first international conference on smart grids, green communications and IT energy-aware technologies; 2011. p. 70–5.
[30] AlRashidi MR, El-Hawary ME. A survey of particle swarm optimization applications in electric power systems. IEEE Trans Evolut Comput 2009;13(4):913–8.
[31] Del Valle Y, Venayagamoorthy GK, Mohagheghi S, Hernandez JC, Harley RG. Particle swarm optimization: basic concepts, variants and applications in power systems. IEEE Trans Evolut Comput 2008;12(2):171–95.
[32] Kennedy J, Eberhart R. Particle swarm optimization. In: Proceedings IEEE international conference on neural networks, ICNN, vol. 4; 1995. p. 1942–8.
[33] Eberhart R, Kennedy J. A new optimizer using particle swarm theory. In: Proceedings of the sixth international symposium on micro machine and human science, 1995, MHS'95. IEEE; 1995. p. 39–43.
[34] Venayagamoorthy GK, Harley RG, Wunsch DC. Comparison of heuristic dynamic programming and dual heuristic programming adaptive critics for neurocontrol of a turbogenerator. IEEE Trans Neural Networks 2002;13(3):764–73.
[35] Prokhorov D, Wunsch DC. Adaptive critic designs. IEEE Trans Neural Networks 1997;8(5):997–1007.
[36] Lendaris GG, Paintz C. Training strategies for critic and action neural networks in dual heuristic programming method. In: Proceedings of international conference on neural network, vol. 2. 1997; p. 712–7.
[37] Liu D, Xiong X, Zhang Y. Action-dependent adaptive critic designs. In: Proceedings of the INNS–IEEE international joint conference on neural network, vol. 2; 2001. p. 990–5.
[38] Boaro M, Fuselli D, Angelis F, Liu D, Wei Q, Piazza F. Adaptive dynamic programming algorithm for renewable energy scheduling and battery management. Cognitive Computation 1–14, in press, http://dx.doi.org/10.1007/s12559-012-9191-y.
[39] Si J, Wang YT. On-line learning control by association and reinforcement. IEEE Trans Neural Networks 2000;3:221–6.
[40] Bakirtzis AG, Dokopoulos PS. Short term generation scheduling in a small autonomous system with unconventional energy system. IEEE Trans Power Syst 1988;3(3):1230–6.
[41] Chacra FA, Bastard P, Fleury G, Clavreul R. Impact of energy storage costs on economical performance in a distribution substation. IEEE Trans Power Syst 2005;20(2):684–91.
[42] Corrigan PM, Heydt GT. Optimized dispatch of a residential solar energy system. In: Proceedings of the North American power symposium; 2007. p. 183–8.
[43] Lu B, Shahidehpour M. Short-term scheduling of battery in a grid-connected PV/battery system. IEEE Trans Power Syst 2005;20(2):1053–61.
[44] Lee TY. Operating schedule of battery energy storage system in a time-of-use rate industrial user with wind turbine generators: a multipass iteration particle swarm optimization approach. IEEE Trans Energy Convers 2007;22(3):774–82.
[45] ComEd U. Residential real-time pricing; 2010. <http://www.thewattspot.com>.
[46] Wilcox S, Marion W. Users manual for TMY3 data sets. National Renewable Energy Laboratory; 2008.
[47] Marquez M, Coimbra C. Forecasting of global and direct solar irradiance using stochastic learning methods, ground experiments and the NWS database. Solar Energy 2011;85(5):746–56.

[48] Isa I, Omar S, Saad Z, Noor NM, Osman MK. Weather forecasting using photovoltaic system and neural network. In: Second international conference on computational intelligence; 2010. p. 96–100.
[49] Unsihuay-Vila C, Zambroni de Souza A, Marangon-Lima J, Balestrassi P. Electricity demand and spot price forecasting using evolutionary computation combined with chaotic nonlinear dynamic model. Int J Electr Power Energy Syst 2010;32(2):108–16.
[50] Conejo AJ, Plazas MA, Espinola R, Molina AB. Day-ahead electricity price forecasting using the wavelet transform and ARIMA models. IEEE Trans Power Syst 2005;20(2):1035–42.
[51] Li G, Liu C-C, Mattson C, Lawarree J. Day-ahead electricity price forecasting in a grid environment. IEEE Trans Power Syst 2007;22(1):266–74.