

Simulation of the operation of hydro plants in an electricity market using Agent Based Models – introducing a Q Learning approach

José Carlos Sousa
FEUP/DEEC and EDP Produção
Rua Dr. Roberto Frias,
4200-465 Porto, Portugal
jose.sousa@edp.pt

João Tomé Saraiva
FEUP/DEEC and INESC TEC
Rua Dr. Roberto Frias,
4200-465 Porto, Portugal
jsaraiva@fe.up.pt

Abstract—The restructuring of power systems with the introduction of electricity markets and decentralized structures increased the number of participating entities. This is particularly true in generation and retailing which are now provided under competition. Accordingly, it is important to develop models to simulate the behavior of these agents and to optimize their participation in electricity markets. Among them, it is essential to adequately model generation agents namely in countries having a large share of hydro stations. This paper describes an agent-based approach to model the day-ahead electricity market having a particular emphasis on hydro generation. Apart from the characterization of the agents, the paper details the introduction of the Q-Learning algorithm in the model as a way to enhance the performance of generation agents. This paper also presents some preliminary results taking the Portuguese generation system as an example.

Index Terms—hydro stations, electricity markets, operation planning, agent-based models. Q learning.

I. INTRODUCTION

The continuous development of power systems and the challenges and opportunities created by these changes have radically modified the simulation and optimization of the operation of power systems. Specifically, the optimization of the operation of the power systems with a large share of hydro generation has been regaining interest both in the research community and the electricity industry due to the impact of these units not only from the technical point of view but also regarding the financial results of generation companies. In fact, the characteristics of hydro power plants such as reliability, availability, storage capability and reduced response time turn this type of assets very important for power system operation. Currently, the existence of pumping capabilities in an increasing number of hydro plants turns the management of these assets very important for generation companies as a way to increase the overall revenues. On the other hand, the mentioned characteristics turn hydro power plants very appealing as a very efficient way to provide reserve services so that they are becoming more and more

important from the point of view of the TSO's. Additionally, their dynamic characteristics combined with their storage capability turn hydro power plants an important asset to help the management of power systems having a large share of renewable generation associated to volatile primary resources as wind and solar. These concerns are particularly relevant in the Iberia Peninsula given the important share of wind and solar generation in the global generation mix.

Taking all these concerns into account, it is easily understood the importance of developing new and more specific models so that generation companies can adequately respond to competition. The role of modeling and simulation to support decision-making in complex systems has been widely established as a valid technique. Recently, agent-based models were reported as a complement to equilibrium models when the problems are too complex to be analyzed by traditional approaches. Agent-based simulation follows the metaphor of autonomous agents and multi-agent systems as the basis to conceptualize complex systems. That is, a model is built taking advantage of the interaction between agents acting in a simulation environment.

There are several approaches in the literature to optimize and simulate generation systems operating in market environment. However, the presence of a large share of hydro generation, specially pumping hydro, is not adequately addressed [1]. Accordingly, this paper presents a model using an agent-based environment that was originally described in [2-3] and in which we are now introducing a Q learning procedure. This enhanced Agent-Based Model is then used to simulate the Portuguese generation system. In this scope, we considered four types of hydro plants: run of river stations, storage stations, pumping storage stations and pure pumping stations. Hydro plants are modeled as agents that can produce and also consume (in the pumping case) meaning that they have to negotiate energy in the market as introduced in [2-3]. To support the hydro-pumping decisions different optimization models were already developed as described in [4, 5] namely using nonlinear programming and Genetic Algorithms. In this paper, we introduced a Q learning

methodology to provide the agents with learning capabilities in stand of the optimization models reported in [4, 5].

Following these ideas, this paper is structured as follows. After this Introduction, Section II overviews the Iberian Electricity Market, given that Portugal participates in this market together with Spain. Then, Section III gives an overview on the existing approaches to deal with the hydro scheduling problem with particular emphasis on agent-based models. Section IV details the developed agent-based model and Section V presents the results obtained so far. Finally Section VI draws the most relevant conclusions.

II. ELECTRICITY MARKETS REVIEW

A. *New Structures and the Unbundling Model*

During the last 25 years, several countries restructured their power systems with the main goal of introducing competitive mechanisms in some parts of the value chain. These changes included the segmentation of the traditional vertically integrated utilities in several activities, namely generation, transmission, distribution and retailing. At the same time, the advent of independent regulation, the increase of the number of agents namely in the generation and retailing and the decoupling between market functions, assigned to Market Operators, and technical operation issues assigned to System Operators brought more complexity to this area.

Regarding the activities mentioned above, usually generation and retailing are liberalized and operated under competition while transmission and distribution network services are provided by regulated monopolies. In addition, in order to match the demand and the supply new mechanisms were created, namely the day-ahead pool markets, the bilateral forward contracts and ancillary services (as for instance secondary and tertiary reserve markets in several European countries). The day-ahead market that exists in several European countries (as for example NordPool for the Nordic European countries and the MIBEL involving Portugal and Spain) corresponds to a short term day ahead mechanism based on the matching of the submitted selling and buying bids for each trading hour of the next day. The market clearing prices are obtained under a marginal basis and can be very volatile. In order to deal with this volatility, long-term contracts are also possible in most market implementations under different horizons and conditions.

Another important issue to understand the recent evolution of power systems is related with the increase of dispersed and volatile generation. Several countries were very successful in increasing the installed capacity in wind parks, photovoltaic and other thermal renewable stations because of the adoption of subsidized feed-in tariffs. Because of this movement, in countries as Portugal and Spain the share of feed-in generation in the total installed capacity is above 35% (22% for wind parks) and the share of renewable generation is above 50% admitting average hydro years. About 60% of these units are connected to distribution networks which is forcing changing the operation paradigm of these grids. Considering this issue, hydro power plants play an important role in systems having a large share of renewables due to their dynamic characteristics and storing capabilities.

B. *The Iberian Electricity Market*

In line with what was mentioned above, the Portuguese and Spanish power systems went through several changes since the late 1990's. In Portugal, the power industry was nationalized in the 1970s with the creation of a vertically integrated utility. This structure started to change in 1995 when a new electricity law was passed admitting the coexistence of a public and a market driven sector. Later, in 2006, a new electricity law was passed organizing the industry in generation, transmission, distribution and retailing activities. The Regulatory Agency was created in 1995 and is responsible for the publication of several codes and for setting the tariffs. Since 2007, all clients are eligible and by the end of 2015 the free market represented 89% of the total demand.

The Spanish power system was also organized in terms of vertically integrated utilities having a regional distribution. A new law was also passed in 1995 in a first attempt to introduce competitive mechanisms. Later on, by the end of 1997 a new law was approved enabling the launch of the electricity day-ahead market on the 1st of January 1998. Since then, a fast transition of regulated captive clients to the free market was implemented so that full eligibility was achieved in 2003. The implementation of the common electricity market, MIBEL, started with the signature of a memorandum by the Portuguese and the Spanish governments in 2001. After several delays, a common bilateral contract trading mechanism was set in place in 2006 and the joint day-ahead market started on the 1st of July 2007 as an extension of the already existing Spanish day-ahead market. In the first operation years the electricity prices in the two areas were different in a large number of hours due to the application of market splitting to solve congestion in the interconnections. Nowadays, due to the increase of the interconnection capacity and the increasing share of generation connected to distribution networks, transmission grids are less loaded so that the number of congested hours declined. As a result, the prices converged to common values in almost 85% of the hours in 2013 and 2014.

Regarding the generation mix, both countries have a large share of hydro plants with a huge variation in their annual output. In terms of the renewable share, both countries were very successful in increasing the amount of renewables. This corresponded to a strategic policy adopted by successive governments to use more intensively endogenous resources, to enlarge the energetic independency and also to develop new industrial activities thus creating new jobs. By the end of 2015, wind power reached an installed capacity of 4634 MW out of 18553MW in Portugal (25%) and of 22854 MW out of 102613 MW in Spain (22 %) with a contribution to demand supply of 24% in Portugal and 19% in Spain.

III. LITERATURE REVIEW ON HYDRO UNITS IN MARKETS

A. *Hydro Scheduling Problem*

One of the main problems that generation companies having hydro power plants in their portfolio have to face is to build the most profitable operation strategy in order to maximize their revenues. In a competitive environment, they have to prepare selling and buying bids, when they have pumping, and submit them to the day-ahead Market Operator.

In addition, the nonlinear relation between the power, the flow and net head, and the uncertainty associated to the hydro conditions turn the optimization of hydro power plants in a complex and nonlinear problem. There are several approaches available in the literature to deal with this kind of problems. In [6] the author's use dynamic programming to solve the hydro scheduling problem but this technique usually leads to the well-known "curse of dimensionality". Other authors use mixed integer linear programming [7] and meta-heuristics, as Simulated Annealing [8], Neural Networks [9] and Genetic Algorithms [5]. The mentioned nonlinear relation can also be addressed using an iterative procedure as described in [4].

B. Electricity Markets Modeling

There are several works that were developed to model electricity markets using different techniques. These approaches can be organized in four main areas [10]:

- Optimization problems, addressing a single company also known as single firm optimization models;
- Equilibrium Models based on Game Theory, considering a larger number of competitors;
- Agent-Based Models, ABM, that simulate the behavior of the companies and the interactions between agents;
- Hybrid solutions.

Optimization models typically address the maximization of the revenues of a single company, often considered as a price taker. Some examples were described in Section III-A. Equilibrium Models represent the market behavior considering the competition between all participants. More recently, Agent-Based Models are becoming an interesting alternative when the complex level of the problem prevents using traditional equilibrium framework. Agent-based computational economics (ACE) corresponds to the computational study of economic dynamic systems modelled as virtual worlds of interacting autonomous agents in an environment.

C. Agent-Based Models in Electricity Markets

Currently there are several models, most of them having commercial nature, addressing this issue. AMES (Agent-based Modeling of Electricity Systems) is an open source platform to simulate the strategic trading behaviors in restructured markets considering AC grids [11]. EMCAS (Electricity Market Complex Adaptive Systems) is a commercial ABM software developed by the Argonne National Lab having the capability of taking decentralized decision-making along with learning and adaptation for agents. EMCAS is linked to the VALORAGUA model [12] that provides longue term operation planning strategies for hydro plants. With this information, EMCAS uses the price forecasts and weekly hydro schedules given by VALORAGUA to do intra-week hydro plant optimization for hourly supply offers. Finally, MASCEM (Multi Agent based Electricity Market) is a simulation platform based on a multi-agent framework [13]. It includes day-ahead, and balancing markets and considers both simple and complex bids turning it both in a short and a medium term model.

Nevertheless, hydro generation, specially pumping hydro stations, is not adequately characterized taking into account the increase of renewable volatile sources. For instance, EMCAS includes the VALORAGUA model turning it very dependent on the performance of this model. This also means that EMCAS does not include the definition of bidding strategies to hydro power plants. Taking this into account, the main objective of this research is to simulate hydro generation in a market environment using an ABM platform, especially regarding hydro with pumping given the extra flexibility these stations have in terms of buying electricity in off peak hours when eventually extra wind generation is available and selling it in peak hours. This will allow us to study their impact on systems having a large penetration of renewable sources, especially wind.

D. Q learning

The characteristics of electricity markets contribute to create a complex dynamic and adaptive system. Each market player faces an uncertain environment mainly due to the inherent uncertainty of power system conditions and the lack of complete knowledge about the competitor's strategic behavior. In these circumstances, learning and constructing the model of the economic system is a very complicated task for market participants, and a model free learning can be an appropriate alternative to build a desired bidding strategy [14]. Q learning is a reinforcement learning methodology [15] in which agents can learn a task by interacting with the environment through a trial and error search. The Q learning algorithm was initially proposed in [16] and it can be classified as a free model because it doesn't need an explicit knowledge about its environment. Instead, the knowledge of the optimal strategy increases while the historic interaction with the environment is being built by trial and error.

Q learning is a useful algorithm to solve Markov decision problems, and this is done by evaluating the payoff for a given state-action pair. So, the Q learning matrix is composed by cells known as Q values. These Q-values are calculated for each pair of state (s) and action (a), and therefore they can also be described as Q(s, a). As the Q learning focuses on the impacts of rewards (R) on the choices of actions in each state, the Q values are obtained by a function that provides the expected utility of taking a given action in a given state [16]. The Q(s, a) function is typically given by (1).

$$Q(s_t, a_t) = (1 - \lambda)Q(s_t, a_t) + \lambda[R(s_t, a_t) + \gamma \max_{a'} Q(s_{t+1}, a')] \quad (1)$$

In (1) $\lambda \in (0,1)$ is the learning rate, which controls the degree to which recently learned information will override the old one (λ equal to 0 makes the agent not learn, while equal to 1 induces the agent to consider only the most recent information). The parameter γ is the discount factor that determines the importance of future reinforcements (γ equal to 0 makes the agent myopic by only considering current rewards, while values closer to 1 turn distant rewards more important). The expression $\max_{a'} Q(s_{t+1}, a')$ represents the best the agent thinks it can do in state s_{t+1} [16].

In addition to λ and γ , an agent can use the ϵ parameter, known as ϵ -greedy strategy, to make a tradeoff between exploitation and exploration [17]. It means that the agent

selects the action that has the maximum Q value with high probability $(1-\epsilon)$ and an arbitrary action from all admissible actions with small probability ϵ , regardless of the Q values.

IV. DEVELOPED AGENT-BASED MODEL

As mentioned before the main goal of this paper is to introduce a Q learning procedure in the Agent-Based Hydro plants model detailed in [1], [2] and [3].

A. Agent-Based Model hydro modeling

Hydro agents have to bid their energy in the market and their strategy depends on the type of hydro. In our work, the bid strategy just depends on the bid price since we admitted that the bided quantity corresponds to the power associated with the water that is available in the unit. Depending on the type of hydro unit, the bidding price strategy is determined by the water value on the reservoir, by a learning parameter α and by a decision supporting tool, all of them originally described in [2, 3] and modeled by (2). The water value function $f(\text{water value})$ provides each plant with a reference bid price that depends on the reservoir level, as illustrated in Figure 1. This curve indicates that if the reservoir level is larger, then the value of the stored water is more reduced and so a more reduced bidding price can also be used. This water value function is calculated for each week according to the procedure detailed in [2]. In [2] we use an optimization model to compute the shadows prices of each reservoir (water values), for several hydro conditions.

$$\text{Bid price strategy} = f(\text{water value}) + \text{bid up/down}(\alpha) \quad (2)$$

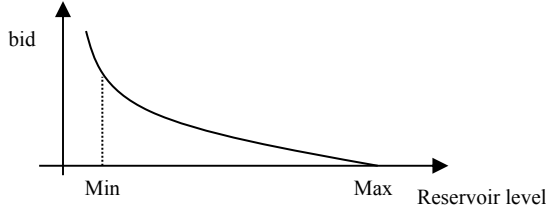


Figure 1. Base bid price in function of the reservoir level.

The *bid up/down* parameter α used in (2) models an agent strategy to increase or decrease its bid price as a way to increase the profit. This parameter is given by the Q learning procedure and it is modeled using a sigmoid function that reflects the risk profile of each agent. If an agent has a higher risk profile, the bid range is larger. On the other hand, a low risk profile will lead to a smaller bid range as illustrated in Figure 2. This strategy is an adaptation of the *derivative-following* strategy presented in [18] and discussed in [13] and also used in [12].

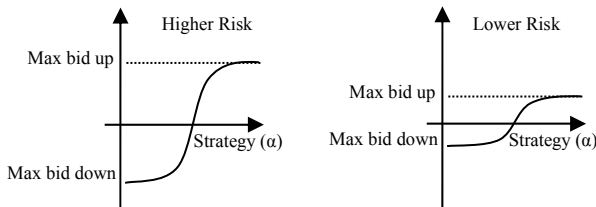


Figure 2. Bidding strategy taking into account the risk profile of each agent (higher risk profile on the left and lower on the right).

The developed ABM model considers four types of hydro agents having different bidding strategies [3] as briefly outlined below:

- Run of river – these agents typically have a water value function near 0, so they will have more focus on their bid up/down strategy;
- Storage - these agents will have a bid value directly related to their water value function as well as to their bid up/down strategy;
- Storage with pumping - the bid price is linked to their water value function and to the bid up/down strategy. They also have the possibility of buying energy to pump water to their reservoir, taking advantage of low prices;
- Pure pumping - these agents are assigned a zero water value because these reservoirs are usually small. They will use decision support tools to forecast the day-ahead electricity prices so that they can define an arbitrage strategy based on price differential between peak and off peak hours [3].

The ABM model also includes thermal and renewable generation agents, that have a similar strategy of the hydro power plant agents, but in which the water value function is substituted by their marginal cost in the case of thermal units and by 0 €/MW.h in renewable generation agents in order to model their dispatch priority according to the Portuguese legislation.

The Market Operator agent is an artifact agent, because it doesn't have a decision making process [1]. It performs the market clearing operation determining the market price and communicating the market results to all market agents. Regarding demand agents, we considered two types of agents: inelastic agents that buy energy at the maximum value allowed in the MIBEL rules (180 €/MW.h), and elastic agents that are designed to model the behaviour of consumers that can directly participate in the market, typically large industries or hydro pumping stations. Elastic consumers will display some demand response regarding price variations in their buying curves. Finally, a Regulator agent is also used. Its main goal is to monitor the generator bids and penalize the generation agents if the bid prices are very different of the marginal cost regarding thermal stations or of the water value for hydro stations.

B. Q learning methodology

In this work we used a *bid up/down* (α) parameter to model the strategy of each agent by increasing or decreasing its bid price as a way to increase the profit. This behavior is modeled by the sigmoid function in Figure 2 to reflect the risk profile of each agent. As mentioned, Q learning is a useful algorithm to solve Markov decision problems, and this is done by evaluating the payoff for a given state-action pair $Q(s,a)$. In our work, and in order to simplify the problem we used 7 states (s_1 to s_7) as illustrated in Figure 3 to discretize this sigmoid function.

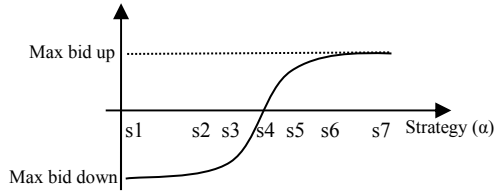


Figure 3. States (s1 to s7) used in Q learning procedure.

State s1 indicates a maximum bid down, s4 means that neither bid up nor bid down is used and s7 represents a maximum bid up. The actions (a) represent the choice of a different state, as for example a12 is the action of passing from s1 to s2. The reward function corresponds to the profit that each agent obtains in the market due to the use of an action in a given state. In our simulation, the agents will learn through their experience or training. In an initial phase, the agents will explore randomly state to state until they reach the end of exploration period. In this case, the end of the exploitation period occurs when the Q values are no longer increased more than 5% (convergence) regarding the values in the Q matrix in the previous iteration. Then, with the Q values defined, the agents start their bidding offers taking into account the learned experience. In our work, the selection of the best action that has the maximum Q value has a probability of $(1-\epsilon)$ and an arbitrary action from all admissible actions is possible with small probability ϵ , regardless of the Q values. In this case ϵ was set at 10%.

Using these ideas the Q learning algorithm evolves as follows.

1. Initialize the matrix Q as zero matrix;
2. During the exploitation period, for each bid:
 - A. Randomly select the initial state;
 - B. Do while not reach the end of exploitation:
 - a. Select one among all possible actions for the current state;
 - b. Using this possible action, consider to go to the next state;
 - c. Get the maximum Q value of the next state based on all possible actions;
 - d. Compute the pay-off given by equation (1);
 - e. Set the next state as the current state;
 - End Do.
- End For.
3. Select the bids with the best Q values with probability of $(1-\epsilon)$.

V. PRELIMINARY RESULTS

In this paper the test case was based on a simplified version of the Portuguese generation system, to allow a better analysis of the results from the Q Learning algorithm. We considered 22 hydro power plants having constant inflows and 11 thermal (coal and CCGT) units. The generation mix also includes 5 reservoir pumping plants. In the simulations we used the 2013 historical generation profile for renewable units. The demand is assumed totally inelastic and prepared to pay the maximum price admitted in MIBEL, 180 €/MWh. The used demand profile corresponds to the 2013 demand data. For simplification and to better understand the results it was

considered that all generators have the same risk profile and that the maximum bid up and bid down were set at 5 €/MWh and at -5 €/MWh. For hydro power plants, we used the hydro condition of 2013 (average hydro year).

A. Definition of the learning parameters

As indicated in (1), there are 2 parameters that have to be defined. Their values are typically case dependent, and in our work we tested several combinations and evaluated the corresponding global average market price. The criterion to select the final combination was to choose the learning parameters that lead to a higher market price, which means that the agents were more effective in maximizing their profit. In the first place, we set γ at 1 and tested different values for λ . Then for the λ associated with the largest average price that was obtained, we changed the value of γ to get the best combination. The results are presented in Table 1.

Table 1. Testing different learning parameters.

λ	Average annual Market Price (€/MWh)	γ	Average annual Market Price (€/MWh)
1	49,13	1	49,21
0.75	49,15	0.75	49,41
0.5	49,25	0.5	49,23
0.25	48,95	0.25	49,24

Although the differences among the annual average market prices are very small we can conclude that the best values for λ and for γ are 0,5 and 0.75. These values were used in the subsequent simulations.

The ϵ parameter represents the probability of choosing an arbitrary action and it was set at 0.1. Using this value means that an agent can select a decision different from the best one, but it is important to “experiment” other actions rather than the one suggested by the largest Q values as a way to enhance the learning capabilities of the agents.

B. Bidding strategies results

Figure 4 shows the evolution of the bidding strategy along the first 3 months of the simulation period. In this figure we present the results for 4 types of power plants: run of river (yellow chart), hydro with pumping (blue chart), coal (brown chart and CCGT power plant (red chart).

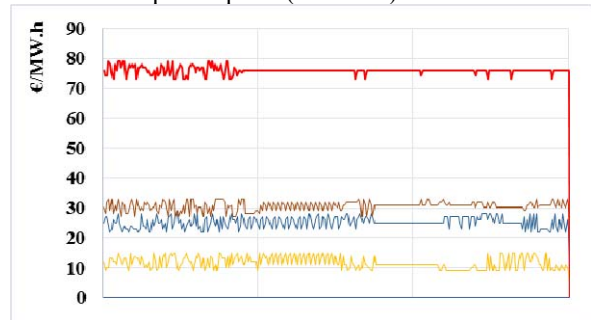


Figure 4. Example of hourly bidding strategies for different generation units.

As we can see, there is an initial phase where the agents are exploring the strategies, and bid randomly to get the Q values. After the Q values are no longer increased, the agents

tend to adopt a more “stable” strategy. In the case of hydro power plants, there is a small trend to bid down, because this bid price reduction guarantees a 100% of utilization of the available water. Apart from that, this strategy doesn’t also expose these agents to negative profits, because they have a zero generation cost. In the case of coal units, the bidding trend is to bid up, because of the cost of coal, and a bid down strategy could originate negative profits. The CCGT agent converges to bid at its marginal cost. In this case, a bidding down strategy increases the danger of getting negative profits, and a bid up decision would mean putting these units out of the dispatch taking into account that they are currently the marginal technology in the Iberian market. This is explained because feed in generation has a large installed capacity and it has priority in the dispatch thus moving some technologies, namely CCGT, to the marginal position.

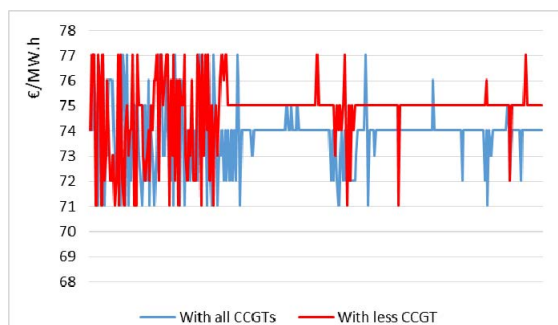


Figure 5. Bidding strategies examples for a CCGT with less competition

To further test the Q learning algorithm, we simulated the generation system again now removing 5 CCGT units. The bidding result is presented in Figure 5. As we can see, since there are less CCGT units in the bidding process, competition is more reduced and so the remaining CCGT’s are now bidding 1 €/MWh above as the initial step is finished. In other words, once some CCGT units are removed, we are increasing the market power of the agent that owns the remaining ones. To increase its profit, this agent adopts a bid up strategy that is it bids above its marginal cost thus reacting to a change of its environment. This shows that the agents learn by experience and by interacting the environment by the application of Q learning procedure. In our simulation the demand is inelastic (only hydro pumping have elasticity). The consideration of other elastic demand, as for example electrical vehicles may change this strategies.

VI. CONCLUSIONS

This paper presents the preliminary results obtained with an ABM model to simulate the electricity market focusing on hydro units and including a Q learning strategy. The results confirm that the agents have learning capabilities and are maximizing their profit using this reinforcement learning strategy. Nevertheless there are some issues to be enhanced. The Q learning algorithm needs some time to allow agents to learn by experience, that is, it takes some time to build the Q matrix by trial and error. On the other hand, the dynamic of electricity markets is large meaning that markets prices are volatile, due to the renewable variability and the variation in the hydro conditions. These changes can turn the learning process more complicated. Therefore, future developments

include the combination of Q-learning with an internal optimization procedure to reduce the learning time and to enhance the capability of reacting to environmental changes.

ACKNOWLEDGMENT

The authors are thankful to Z. Kokkinogenis, R. Rossetti and A. P. Rocha from LIAC - Artificial Intelligence and Computers Sciences Laboratory (FEUP).

REFERENCES

- [1] J. C. Sousa, Z. Kokkinogenis, R. Rossetti, J. T. Saraiva, “Electricity Market Modeling and Renewable Energy Integration: an Agent-Based Conceptual Model”, in *Proc.10th Int. Multidisciplinary Modeling & Simulation Multiconference, I3M/SESDE 2013*, Athens, Sept. 2013.
- [2] J. C. Sousa, J. T. Saraiva, Z. Kokkinogenis, R. Rossetti, “Operation Planning of Hydro Power Plants Using Agent Based Models”, in *Proc. of MedPower 2014*, Athens, Nov. 2014.
- [3] J. C. Sousa, J. T. Saraiva, “Simulation of the operation of hydro plants in an electricity market using Agent-Based Models”, in *Proc. of 12th International Conference on the European Energy Market, EEM2015*, Lisbon, Portugal, May 2015.
- [4] J. C. Sousa, V. T. Mendes, J. T. Saraiva, “Estimation of the Remuneration of Hydro Plants in a Market Environment Using an Iterative Under-Relaxation Approach”, in *Proc. 2009 IEEE Bucharest Power Tech*, Romania, 28 June – 2 July 2009.
- [5] Guerreiro, C.A., J. T. Saraiva, J. C. Sousa, V. T. Mendes, “Operation planning of hydro stations using genetic algorithms considering their impact on the electricity market prices”, in *Proc. 11th International European Electricity Market Conference, EEM’14*, Krakow, May 2014.
- [6] S.-C. Chang, C.-H. Chen, I.-K. Fong, P. B. Luh, “Hydroelectric Generation Scheduling With An Effective Differential Dynamic Programming Algorithm”, *IEEE Transactions on Power Systems*, vol. 5, no. 3, pp. 737-743, August 1990.
- [7] G. W. Wang, C. T. Su, “A Practical Mixed Integer Linear Programming Based Short Term Hydro Scheduling”, in *Proc. of the IEEE Asia-Pacific Trans. and Dist. Conference*, vol. 3, pp. 1606-1610, October 2002.
- [8] K. Wong, Y. Wong, “A Parallel Simulated Annealing Algorithm for Short-Term Hydro Scheduling”, in *Proc. of the IEEE Int Forum on Applications of Neural Networks Power Systems*, pp. 335-340, April 1993.
- [9] R. H. Liang, Y. Y. Hsu, “Short-Term Hydro-Scheduling using Hopfield Neural Network”, *IEE Proc. Generation, Transmission & Distribution*, vol. 143, no. 3, pp. 269-275, May 1996.
- [10] G. Li, J. Shi, X. Qu, “Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market—A state-of-the-art review”, *Energy*, vol. 36, no. 8, pp. 4686-4700, August 2011.
- [11] H. Li, L. Tesfatsion, “Development of open source software for power market research: The AMES test bed”, *Journal of Energy Markets*, vol. 2, no. 2, pp. 111-128, 2009.
- [12] P. Thimmapuram, T. D. Veselka, V. Koritarov, S. Vilela, R. Pereira, R. F. Silva, “Modeling hydro power plants in deregulated electricity markets: Integration and application of EMCAS and VALORAGUA”, in *Proc. EEM2008*, pp. 1-6, May 2008.
- [13] Z. Vale, T. Pinto, I. Praça, H. Morais, “MASCHEM: electricity markets simulation with strategic agents”, *IEEE Transactions on Intelligent Systems*, vol. 26, no. 2, pp. 9-17, March-April 2011.
- [14] M. Rahimiyan, H. R. Mashhadi, “An adaptive Q-learning algorithm developed for agent-based computational modeling of electricity market”, *IEEE Tran. on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, vol. 40, no. 5, pp. 547-556, Sept. 2010.
- [15] Sutton, R. S.; Barto, A. G. *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press, 1998.
- [16] Watkins, C.; Dayan, P. Technical Note: Q-Learning. *Machine Learning*, 8, pages 279-292, 1992.
- [17] Felipe Alves Calabria, Phd Thesis: Enhancing Flexibility and Ensuring Efficiency and Security: “Improving the Electricity Market in Brazil Using a Virtual Reservoir Model”, FEUP, 2015.
- [18] A. Greenwald, J. Kephart, and G. Tesauro, “Strategic Pricebots Dynamics,” in *Proceedings of the 1st ACM Conf. Electronic Commerce*, ACM Press, pp. 58–67, 1999.