

A new electricity market design for power systems with large share of hydro: Improving flexibility and ensuring efficiency and security in the Brazilian case

Felipe Alves Calabria

CNPq Brazil and FEUP

Fac. de Engenharia da Univ. do Porto

Porto, Portugal

felipecalabria@gmail.com

João Tomé Saraiva

INESC TEC and FEUP/DEEC

Fac. de Engenharia da Univ. do Porto

Porto, Portugal

jsaraiva@fe.up.pt

Ana Paula Rocha

LIACC/DEI/FEUP

Fac. de Engenharia da Univ. do Porto

Porto, Portugal

arocha@fe.up.pt

Abstract— This paper discusses some problems related with the current Brazilian electricity market, brings out some dilemmas that should be examined in order to implement a more market oriented approach, and proposes a new market design to overcome these issues. The proposed market design is based on the concept of energy right accounts as virtual reservoirs and aims at enhancing the flexibility to enable market participants to comply with their contracts, while still ensuring the efficient use of the energy resources and maintaining the current security supply level. In addition, in order to simulate the behavior of the market participants in this new framework, an Agent-Based Model - ABM where agents use reinforcement Q-Learning is developed for the study case. The results show that this new market design is suitable to be applied to hydrothermal systems having a large share of hydros.

Index Terms - Electricity supply industry deregulation, Hydroelectric power generation, Multi-agent systems.

I. INTRODUCTION

Several electricity market structures have evolved to ensure free access, guarantee fair competition, foster higher efficiency, and decrease price, while maintaining or improving reliability of operation and security of supply. Although Brazil made significantly progresses towards liberalization, its electricity market has certain particularities that contribute to distinguish it from other markets. With a continental interconnected transmission system, a large and growing demand, and a total installed generation capacity around 137 GW [1], from which around 70% comes from hydropower plants with multiple owners coexisting in hydro cascades, this electricity market has gone through two large institutional and regulatory reforms in the last twenty years. Currently, it contains special processes, mechanisms and instruments to handle the mentioned peculiarities.

Hydro companies usually participate in national public auctions with descending price in order to sign long-term contracts (usually up to 35 years) to build and operate power

plants. This electricity purchase occurs in terms of a single buyer model, in the scope of which PPAs - Power Purchase Agreements are signed between generators and distribution companies. After closing a contract and when the power plant becomes commissioned, market participants can perform a seasonal adjustment of the contracted energy and of the physical guarantee. The physical guarantee corresponds to the maximum energy production that can be maintained almost continuously over the years. It is defined by the Ministry of Mines and Energy and it corresponds to a certificate that determines the total amount of energy that can be traded in the Brazilian electricity market.

Regarding the seasonal adjustment, generation companies are allowed once a year to perform the “seasonalization” process, i.e. to set the monthly amount of the annual contracted energy and the monthly amount of the physical guarantee. Herewith, there is an annual “window” to define these amounts. However, the dispatch of the power plants in Brazil is carried out in a centralized way by ONS, the Brazilian Independent System Operator - ISO, which doesn’t consider the mentioned signed contracts when dispatching generation units. The ISO’s dispatch goal is to minimize the total operating cost, and this is done considering, among other issues, future scenarios of water inflows and using a software package driven by a minimal cost dispatch approach.

In this scheme, generators are not allowed to decide their own physical generation in order to comply with their contracts since there is not a bid based market, and the ISO decides their outputs without considering their commercial commitments. Thus, if they are not dispatched, they can be exposed to the risk of having to buy electricity at volatile short-term market prices to complete the energy committed in their signed contracts. Because of that, it was implemented a mechanism to share the aforementioned risk among hydros. This mechanism is known as *MRE* – Mechanism for Energy Reallocation. Shortly, the *MRE* reallocates energy, transferring the surplus generated from those that produce beyond their

monthly allocated physical guarantee to those that produce below. The *MRE* is automatically performed, and there is no flexibility for hydros to address their exposition risk according to their own risk perception and strategy.

The short-term market (known as *MCP*) then takes place after the ISO's dispatch. Unlike other markets, the Brazilian short-term market is not a marketplace where generators are active through a self-dispatch procedure, or generators influence, through their bids, the dispatch schedule. Ultimately, it is not an environment where market participants can engage in short-term trades on behalf of their own account, as there is no short-term declaration of intent.

Moreover, the short-term market price (known as Price for Settlement of Differences or *PLD* for short) is not a result from the market participant's interaction, but it is one of the outputs of the software package that is run by the ISO. So, the *MCP*, rather than a market, is a mechanism to settle differences (i.e., exposed positions) between the amounts of electricity committed by contracts and those amounts of electricity that each agent ends up providing / receiving. These differences must be automatically purchased or sold in the *MCP* and are valued using the *PLD* price. Lastly, having in mind these characteristics, the following problems of this market design are recognized and should be addressed:

- The conciliation between the commercial commitments and the physical dispatch is not smooth: there is a lack of "trading opportunities" to encourage participants to comply with their contracts (namely related with the unique annual "window" of the seasonalization process) and there is no flexibility for hydros to address their exposition risk according to their own risk perception and market strategy (since the *MRE* mechanism is automatically activated);
- The Brazilian short-term market acts as a mechanism to settle differences rather than a true market, and neither short-term price nor the dispatch schedule is influenced by the market; and
- Inconsistencies regarding the software package code, a centerpiece of this market design, may cause a huge impact on the electricity sector, affecting the confidence of the market agents.

As can be observed, the current market design poses relevant issues concerning the proper functioning of an electricity market. In order to overcome these problems it is proposed in this paper a new market design aiming at enhancing the flexibility of market agents to comply with their commercial commitments while maintaining the optimized operation of the entire system. According to these ideas, this paper is organized as follows. Section II details the dilemmas associated with the current market structure in Brazil, Section III explains the proposed virtual reservoir based design, Section IV describes the implemented ABM (Agent Based Model) simulation approach, Section V presents the results obtained and Section VI outlines the main conclusions.

II. TRACKING A SOLUTION

A solution typically adopted in other electricity markets is the employment of a more market-oriented approach. This

market approach would enable generators to bid, in the short-term, quantities and prices, which could be used to set the market positions and to substitute the seasonalization process and the *MRE*. As a result, the short-term market price would be based on the interaction between market participants. Nevertheless, there are some dilemmas to be faced, especially if considering a power system with a large share of hydros. These dilemmas are briefly addressed in the next paragraphs.

A. Dilemma 1: Efficiency of Energy Resources

Putting into perspective the dichotomy between a centralized dispatch (based on the cost merit order) and a decentralized one (based on a market approach), it becomes clear how important it is to coordinate the use of the water stored in the reservoirs in order to safeguard the efficiency of using the energy resources while taking advantage of the all potential stored energy as discussed in [2].

A decentralized dispatch, e.g. a scheme of bids in a market pool, can be essentially an instantaneous process and the inter-temporal features of river chain operations will not be entirely captured if a pure single-period market clearing mechanism is adopted. Moreover, the presence of several owners in the hydro cascades, which is the case of Brazil, endorses a market design based on a centralized dispatch.

B. Dilemma 2: Security of Supply

The Brazilian short-term market exists for about fifteen years. For the time being, the *PLD* has an average value of 37.72 €/MWh, and the hydro average successful bid price from the national public auctions is 38.16 €/MWh [3]. The proximity of both values suggests that the short-term market can provide enough incentives to induce adequate investment to build new power plants. Nevertheless, the *PLD* standard deviation of the entire set of data is around 55.36 €/MWh. With an average of 37.72 €/MWh, this standard deviation imposes a huge risk to the health of the business, especially regarding cash flow stability.

Accordingly, it is advocated the need to have a capacity mechanism to provide enough incentives to ensure the security of supply. Nowadays, this concern is addressed both via the contracting scheme where the demand must be fully ex-ante contracted and contracts physically backed, and via the ISO dispatch, either through the mechanism of risk aversion implemented into the software package or through a dispatch out of the merit order authorized by *CMSE*, the federal entity that monitors the supply adequacy in the country.

C. Dilemma 3: Flexibility to Comply with Contracts

If in one hand a market design having an ISO central dispatch is to be kept, on the other hand:

- there is one unique "window" to define the amount of the physical guarantee that will uphold the contracts and generators operate their power plants just following the amount of production defined by the ISO;
- the mechanism to share the short-term risk exposition (*MRE*) is automatically performed, which imposes a kind of "strait jacket" on the market participants.

Every time there is a water shortage, the *PLD* increases, there are more dispatched thermal stations and less hydro.

Depending on the amount of thermal dispatch, hydros can be displaced in such a way that *MRE* will not have the extra energy to be shared among its participants. When this occurs, an adjustment factor is applied to withdraw a fraction from their seasonalized physical guarantee. Extending this reasoning to long periods of water scarcity, this represents a significant negative exposed position for hydros. In this case, the *MRE* is not able to cover the risk of generators that have to buy electricity in the short-term market to complete the energy committed in their contracts.

If each hydro would decide its own production, the reservoir management would be on their own. However, in the Brazilian case, the decisions regarding each production level, and consequently the amount of energy to be stored in the reservoirs are taken by a third party, the ISO. Nonetheless, water shortage risk is assumed by hydros, once the *MRE* doesn't cover it. It would therefore be more consistent to follow [4]: to structure the framework of incentives such that risks are allocated to those who take the decisions and who hold the responsibility for taking them into account.

III. A NEW MARKET DESIGN

Focused on improving the flexibility to enable market participants to comply with their contracts, while still ensuring the efficient use of the energy resources and maintaining the current security of supply level, a new market design will now be proposed. Based on the concept of energy right accounts as virtual reservoirs, in this new design each hydro is modeled as an agent that has a virtual reservoir representing how much energy is virtually stored in his hydro plant. Hereafter, for each accounting period, each account is fed by the fraction of the total natural affluent energy that flowed to the hydro cascade proportional to the hydro's physical guarantee. Then, the following sequence of events is adopted:

- 1) The system operator continues his work as it is currently done (running the software package, as well as other procedures and schemes), and defines the amount of generation for each power plant. This means that it maintains its responsibility in terms of defining the physical dispatch optimizing the use of the resources, dispatching the hydro and thermal units;
- 2) The "remaining demand" is obtained for each dispatch period. This remaining demand is equal to the total demand minus the total dispatch allocated to the thermal power plants. This difference is the total demand to be supplied by hydros;
- 3) A bid based hydro short-term market is established for the remaining demand. In this market, hydros have the opportunity to withstand their bilateral commercial commitments while trying to make successful bids. The result of this market is a virtual dispatch with financial settlement purposes. To do that, hydro agents can offer bids considering a price between zero and a ceiling price defined by the regulatory agency and a quantity available within his account;
- 4) The final short-term price is calculated as a weighted average considering the most expensive successful hydro price bid and the variable cost of the last non-hydro resource dispatched by the ISO.

In this new market design two worlds would coexist: the real one, associated with the power system considering physical effects, and where the ISO runs the dispatch in a centralized way; and the virtual one, related to the settlement system and with commercial effects, and where hydro agents participate in a short-term market. Both worlds simultaneously operate, and the link between them is the total affluent natural energy that flowed along the hydro cascade in each accounting period. At last, the settlement process will occur considering the successful quantity bid of each participant, and the exposed position will be valued by the new final short-term price, which comes from the combination of thermal costs based on the ISO dispatch and the short-term price arising from the hydro short-term market.

IV. SIMULATION OF THE NEW MARKET

To simulate the behavior of the hydropower plants in this market design, an Agent-Based Model - ABM was developed using reinforcement Q-Learning. According to [5], among the modeling alternatives for bidding strategy analysis in the electricity short-term markets, ABM is highlighted because it allows designing complex electricity markets as collections of rule-based agents interacting with one another dynamically and intelligently, simulating human beings' behavior to build optimal bidding strategies. Generally, the ABM procedure can be described as follows [6]: (1) define the research questions to be addressed; (2) construct a model comprising an initial population of agents; (3) specify the initial model state by defining the agents attributes and the structural and institutional framework of the electricity market within which the agents operate; (4) make the model evolve over time without further intervention; (5) analyze simulation results and evaluate the observed regularities.

In this research, agents represent hydro companies that prepare their bids depending on the level of their (virtual) reservoirs, the hydro short-term market ceiling price and month of the year, which is associated with the water inflows into the hydro cascade and the demand to be supplied by hydros. Their goal is to manage their reservoir level in order to maximize their profit. Agents learn how to act optimally through the Q-learning algorithm, which is one of the most adopted reinforcement machine learning techniques and focuses on the impacts of rewards and punishments on the choices of the agents seeking their goal.

An agent using Q-learning is a goal-oriented learner that, for a given Markov decision process, continually interacts with his environment, receives feedback (rewards or punishments) from it, and searches for the most profitable action considering the past experience. In other words, at each time step t the agent is in some state s_t , chooses any action a_t that is available in state s_t , receives a corresponding reward r_t , and moves into a new state s_{t+1} . Thus, in Q-learning it is imperative to carefully structure the set of states $S = \{s_1, s_2, \dots, s_n\}$, actions $A = \{a_1, a_2, \dots, a_n\}$ and rewards $r_t(s_t, a_t)$. Lastly, the capability of learning to act optimally in Markovian domains by experiencing the consequences of actions is given by a quality function $Q(s_t, a_t)$. This function provides the expected utility of taking a given action in a given state [7], and it is given by (1).

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha \cdot \{r_t(s_t, a_t) + \gamma \cdot \max_a Q(s_{t+1}, a) - Q(s_t, a_t)\} \quad (1)$$

So, for each admissible pair (s, a) , the quality function is defined as the Q value. The parameter α is the learning rate, which reflects the degree to which recently learned information will override the old one (α equal to 0 will make the agent not learn, while equal to 1 will induce the agent to consider only the most recent information), and the parameter γ entitles the discount factor that determines the importance of future reinforcements (γ equal to 0 will make the agent myopic by only considering current rewards, while values closer to 1 turn distant rewards more important). Additionally, the expression $\max_a Q(s_{t+1}, a)$ represents the best the agent thinks it can do in state s_{t+1} .

In order to adequately balance the capacity not to converge to local optima and the acceleration of the learning process, a Simulated Annealing (SA) mechanism is included into the developed algorithm as detailed in [8]. An overview of the entire algorithm is provided in Figure 1.

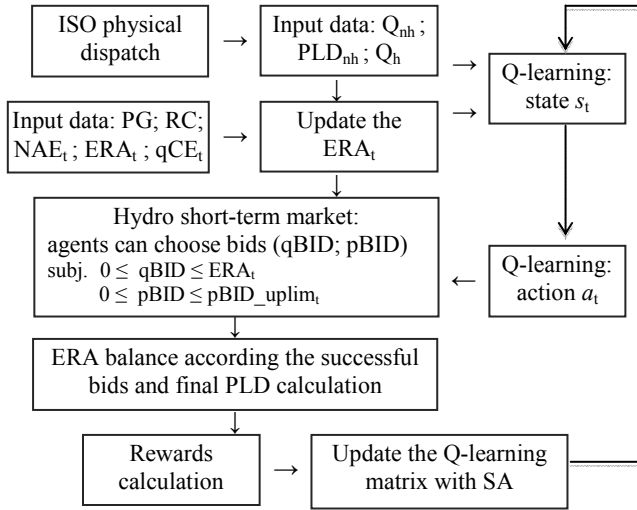


Figure 1. Overview of the developed algorithm.

In brief, to simulate the new market rules the algorithm gets from the ISO physical dispatch procedure the total demand to be supplied by non-hydros (Q_{nh}) and by hydros (Q_h), and the variable cost of the last non-hydro dispatched unit (PLD_{nh}). It also requires information about ex-ante bilateral contracts (qCE_t) for each account period t , i.e. the amount of energy that must be endured in the hydro short-term market by the bids. Other input data is as follows: the natural affluent energy that flowed in period t to the hydro cascade (NAE_t), from which a fraction is allocated to each hydro giving their physical guarantee (PG); the reservoir capacity (RC), and the virtual reservoir level or energy right account (ERA_t) in its initial stage.

The hydro short-term market takes place considering the remaining demand Q_h , and bids can be done once respected the following constraints: the quantity bid (qBID) is limited by the ERA balance in each period t , and the price bid (pBID) is limited by a maximum regulatory price ($pBID_{uplim_t}$). The bids are sorted considering the pBIDs, the successful bids are then identified, and the ERA updated. The clearing price rule

is adopted to define the final price of the hydro short-term market (PLD_h), and the final short-term price (PLD_{final}) is determined by (2) as the weighted average of variable cost of the last thermal dispatched unit and the price of the short-term hydro based market.

$$PLD_{final} = (PLD_h \cdot Q_h + PLD_{nh} \cdot Q_{nh}) / (Q_h + Q_{nh}) \quad (2)$$

The agents then make two bids. The first bid (BID1) addresses the need to comply with bilateral contracts. So, in this simulation and regarding BID1 agents are required to offer either the exact amount in the bilateral contract (if $ERA_t \geq qCE_t$) or the maximum available amount in order to mitigate the exposed position (if $ERA_t < qCE_t$). The second bid (BID2) aims at allowing the agents to manage their reservoirs in order to optimize the leftover stored energy. That is, agents would be allowed to get extra profit in the short-term market when there is more energy in the reservoir than the one need to comply with the bilateral ex-ante contracts. The Q-learning is then structured focusing on the hydros' flexibility to manage this leftover energy (i.e. BID2).

The state space of the Q-learning is organized in twelve states, one for each month of the year. Moreover, it is possible to select among twenty-seven different bids (actions), each one formed by a pair (qBID, pBID). The action space is formed by qBIDs equal to 0% (a1 and a2), 20% (from a3 to a7), 40% (from a8 to a12), 60% (from a13 to a17), 80% (from a18 to a22) and 100% (from a23 to a27) of the ERAaBID1. ERAaBID1 is the available energy in the virtual reservoir after the first bid (BID1). The pBID can be 0%, 25%, 50%, 75% or 100% of the maximum allowed price bid for the account period ($pBID_{uplim_t}$). Both the state and action space are illustrated in Table I, which corresponds to the Q-learning matrix.

TABLE I. ILLUSTRATION OF THE Q-LEARNING MATRIX

States	January	$s1$													
													
	December	$s12$													
			a1	a2	a3	a4	a5	a6	a7	...	a23	a24	a25	a26	a27
			0	100	0	25	50	75	100	...	0	25	50	75	100
			%	%	%	%	%	%	%	...	%	%	%	%	%
			pBID (% of the pBID uplim _t) =												
			qBID =												
			0	20%	...	ERAaBID1	...	100%	...	ERAaBID1					
			Actions												

Finally, the reward is calculated using equation (3).

$$\text{reward}_t = \text{factor}_{\text{goalt}} * (\text{qBID2suc}_t - \text{ERAaBID2}_t) * \text{PLD}_{\text{final}} \quad (3)$$

The $\text{factor}_{\text{goalt}}$ is the adjustment factor concerning the reservoir level goals, qBID2suc_t is the successful quantity BID2, and ERAaBID2_t is the leftover energy in the reservoir after BID2 in period t . The value of the $\text{factor}_{\text{goalt}}$ is obtained by a linear programming problem that runs the same two years in a row and aims at avoiding spillage and maximizing profits on the hydro short-term market, taking into account the qCE_t and $pBID_{uplim_t}$. For the first year of data, this optimization outputs the goal level of each reservoir in each month. The $\text{factor}_{\text{goalt}}$ is equal to 1 if the actual level of the reservoir is within a range where the median is the goal level, and it

gradually reduces to 0.9, 0.7, 0.4 and zero as the current level of the reservoir moves away from the goal range.

Regarding the reward equation, the following should be detailed. There are periods (close to the start period of low water flows) when it is better to save water. In this case, it is expected low $q_{\text{BID}2\text{suc}_t}$, which results in high $\text{ERAaBID}2_t$. Therefore, agents should value the $\text{ERAaBID}2_t$ more than $q_{\text{BID}2\text{suc}_t}$. Nevertheless, in certain occasions (e.g. when there is water scarcity and prices are high) the best action is to offer high $q_{\text{BID}2\text{suc}_t}$ (at a high price), decreasing the level of the $\text{ERAaBID}2_t$. The decision whether to price either the $q_{\text{BID}2\text{suc}_t}$ or $\text{ERAaBID}2_t$ in a given period is guided by the factor $\text{factor}_{\text{goal}t}$. If the agent offers a high q_{BID} when it should be low, the current level of the reservoir will be away from the goal level, and the reward will penalize the agent through the factor $\text{factor}_{\text{goal}t}$. Ultimately, the $\text{PLD}_{\text{final}t}$ in equation (3) will boost the learning agent, reinforcing the periods of high or low price (i.e. low or high water inflow).

V. RESULTS

The described market design and its flexibility to address the physical dispatch and commercial commitments was tested considering a case study that includes four hydro power plants with the characteristics in Table II. In this example, hydros compete to supply a demand Q_h that varies along the simulation period and their aim is to extra profit by selling their leftover energy in the short-term. Since in Brazil the settlement process occurs monthly, the current version of the hydro short-term market also operates on a monthly basis.

TABLE II. CHARACTERISTIC OF THE HYDROS (MW AVERAGE)

Hydro	H1	H2	H3	H4
Physical guarantee (PG)	125	100	50	225
Reservoir capacity (RC)	500	400	200	900
ERA in $t=1$	0	0	0	0

The simulation of the system considers three different weather patterns (year with *high*, *medium* and *low* water inflows) and market conditions (level of ex-ante contract – q_{CE} equal to 100%, 50% and 0% of the physical guarantee). Two types of ceiling price rules are performed: one where the ceiling price is equal to the variable cost of the last non-hydro resources dispatched by the ISO in each month (PLD_{nh}); and another where the ceiling price is *Fixed* at 827 \$/MW average. In Table III we can observe the monthly demand and the dispatch data. As typically more than 80% of the demand in Brazil is supplied by generation coming from hydro power plants, the data in this table illustrates this pattern. Furthermore, the seasonalization of the bilateral contracts is modeled bearing in mind that the annual amount of the contracted energy is allocated considering weights determined by the natural affluent energy that flowed to the cascade.

With the aim of covering all conditions described above, we considered 18 scenarios. The algorithm converges in average after simulating 7,700 years and considering a tolerance of 0.5% on the values of the Q-learning matrix built in a given iteration regarding the matrix of the previous one. Table IV details the scenarios and presents the results

regarding the factor $\text{factor}_{\text{goal}t}$ and the annual average PLD for Hydro 4 (the largest reservoir).

TABLE III. DEMAND AND DISPATCH DATA THROUGHOUT THE YEAR (MW AVERAGE OR \$/MW AVERAGE)

	Month	1	2	3	4	5	6	7	8	9	10	11	12
		Q_{total}	500	500	500	500	500	500	500	500	500	500	500
	PLD_{nh}	217	117	117	57	57	57	117	117	317	417	517	617
High water flow	Q_{nh}	167	5	5	5	5	5	5	5	5	5	5	5
	Q_h	333	495	495	495	495	495	495	495	495	495	495	495
	$\text{NAE}_{\text{total}}$	333	666	1.166	1.833	1.666	833	666	167	167	0	0	0
Medium water flow	Q_{nh}	250	5	5	5	5	5	5	5	5	5	5	250
	Q_h	250	495	495	495	495	495	495	495	495	495	495	250
	$\text{NAE}_{\text{total}}$	267	533	933	1.466	1.333	667	533	133	133	0	0	0
Low water flow	Q_{nh}	300	100	5	5	5	5	5	5	200	300	300	300
	Q_h	200	400	495	495	495	495	495	495	300	200	200	200
	$\text{NAE}_{\text{total}}$	200	400	700	1.100	1.000	500	400	100	100	0	0	0

TABLE IV. ACHIEVEMENT OF RESERVOIR LEVEL TARGETS FOR HYDRO 4 AND THE ANNUAL AVERAGE PLD

Scenarios		Factor _{goal}												PLD _h average			
Water flow	Ceiling price rule	qCE level	n°	State													
				1	2	3	4	5	6	7	8	9	10		11	12	
high	PLD _{nh}	100%	1	1	0.9	0	0	0	0.4	0.4	0.9	0.9	0.7	0	1	142	
medium			7	1	1	0.7	0	0	0	0	0	0	0.7	0.9	1	162	
low			13	1	1	1	0	0	0	0	0.4	0.7	0.7	0.9	1	172	
high		50%	2	1	1	0	0	0.7	1	0.9	0.4	0	0	0	1	151	
medium			8	1	1	0.4	0	0.7	0.9	0.9	0.4	0.4	0	0	1	171	
low			14	1	1	1	0.9	0.7	1	1	0.7	0	0	0	0.7	185	
high		0%	3	0.9	0.9	0	0	0	0.9	0.7	0	0	0	0	1	185	
medium			9	1	1	0.9	0.4	0.7	0.7	0	0	0	0	0	0.9	192	
low			15	1	1	0	0	0.9	1	0.7	0.4	0	0	0	0.7	201	
high	Fixed	100%	4	1	1	0	0	0	0	0	0	0.4	0.7	0.7	0.7	374	
medium			10	1	1	0.7	0	0	0	0	0	0	0.7	0.9	0.9	0.9	258
low			16	1	1	1	0.4	0	0	0	0	0.4	0.9	1	1	345	
high		50%	5	0.9	0.9	0.4	0	0	0	0	0.7	1	0.9	0.7	0.7	500	
medium			11	1	1	0.7	0	0	0	0	0.7	0.9	0.9	0.9	0.9	500	
low			17	1	1	0.9	0.4	0	0	0.4	0.9	0.9	0.9	0.7	0.7	517	
high		0%	6	1	0.7	0	0	0	0.4	0.7	1	1	0.9	0.9	1	551	
medium			12	0.9	1	0.4	0	0	0.7	0.7	1	0.7	0.7	0.9	1	655	
low			18	0.9	0.9	0.9	0	0	0.4	0.9	0.7	0.7	0.4	0.7	0.7	672	

A. The quantity bid behavior

As mentioned before, the factor $\text{factor}_{\text{goal}t}$ is obtained by a linear optimization and it reflects the achievement of reservoir level targets. This linear programming establishes this targets taking into account that agents are free to define their successful quantity bid ($q_{\text{BID}2\text{suc}_t}$). So, as it can be seen in Figure 2, the target scheme from the linear programming output guide the agent to: (i) decrease his reservoir level (offering high $q_{\text{BID}2\text{s}}$) until month 3 in order to minimize a possible spillage once the beginning of the year is a period with high rainfall (i.e. high deposits – DEP into the ERA); (ii) recover the reservoir level (offering very low $q_{\text{BID}2\text{s}}$) in the following months (4, 5, 6 and 7); (iii) sustain the reservoir at a high level during months 7, 8, 9, 10 and 11 ($q_{\text{BID}2\text{s}}$ equal to zero); and

(iv) offer all the stored energy in the month of higher price (i.e. month 12).

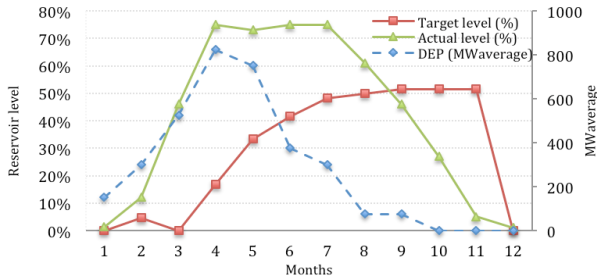


Figure 2. Hydro 4 in scenario 1: DEP, reservoir levels target and actual reservoir levels in each month

Nevertheless, as indicate by the shaded cells in Table IV and in Figure 2, the monthly reservoir level does not end exactly that way. Because all the agents operate over the same goal-oriented scheme, in the beginning of the learning process they tend to offer everything they can in month 12. But once there is competition, the bid quantities are larger than the total demand in this month, leading to non-accepted bids. As a result, agents adapt to competition and improve their behavior offering qBID2 values previously in month 11, 10, 9 and 8, where the ceiling price is also at a good level. In the early months of the year something similar occurs: all hydros try to get rid of the water; competition grows, which decreases the value of the planned qBID2suc. Then, the current reservoir level ends up larger than the target.

Thereby, we observe that, while the linear programming target scheme suggests the agent a preliminary strategy, the Q-learning algorithm teaches the agent how to, at the same time, follow the targets and deal with the dynamic of the market.

B. The price bid behavior

By comparing the annual $PLD_{average}$ of each scenario we can notice that the price increases either (i) when the annual water inflow decreases, (ii) when there is more leftover stored energy available for BID2 due to less commitment by bilateral ex-ante contracts, or (iii) when ceiling price rule moves from the PLD_{nh} pattern to the *Fixed* one. For scenario no. 1 (the one with the lowest $PLD_{average}$), Figure 3 presents the bid prices (dots, each color representing a hydro), the monthly water inflows (NAE), ceiling prices (pBIDuplim) and the final hydro short-term prices.

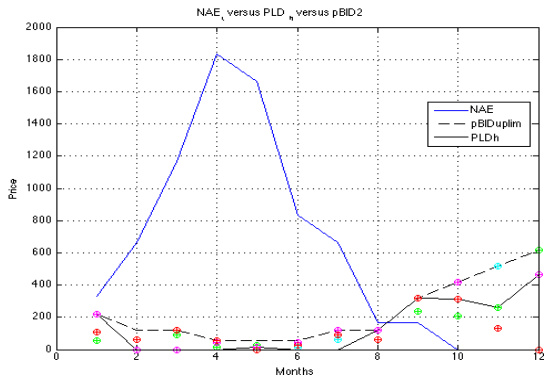


Figure 3. Price bids versus ceiling price and NAE in scenario no. 1.

In economic terms, we should expect an increase of the product price when the availability of the resource needed to produce it (water inflow) decreases. So, when moving from a good hydrological year to a medium or a bad one, the electricity short-term price level should rise. As pointed out by [9], market power in hydro markets arises from the strategic water management of reservoirs. In other words, the reallocation of a given amount of energy across periods can lead to an increase in market prices. Nevertheless, companies acting as price-makers have their market power reduced or cancelled when they are forced to be bilateral contracted. This price trend can also be seen in this case study. Moving from a higher level (100%) of bilateral contracts to a lower one (50% and 0%) originates a sharp price rise. Lastly, the used ceiling price rule significantly influences the final clearing market price, what may suggest, among other things, the exercise of market power.

VI. CONCLUSION

The reported results indicate that, after complying with the bilateral contracts, agents can manage their reservoirs in order to optimize the leftover stored energy. The entire algorithm combines a goal-oriented scheme, built through the mentioned linear programming, with a goal-oriented learner from the Q-learning. Through the performed simulation, we can notice that both the quantity and the price bid strategy are consistent with what is expected in real electricity markets.

Finally, from a broad perspective, the proposed market design maintains the current levels of efficiency and security, while increasing the level of flexibility regarding the commercial behavior of the agents. In the Brazilian case, this flexibility is achieved by replacing the *MRE* and the seasonalization processes by a virtual reservoir model. As a consequence, the management of (virtual) reservoirs is under the responsibility of each hydro, which could (virtually) save water according to their own risk perception. In doing so, the operation of the physical system by the ISO is not affected, ensuring the efficiency of the hydro cascade operation and maintaining the current level of the security of supply.

REFERENCES

- [1] ANEEL – Brazilian Electricity Regulatory Agency. Database of electricity generation [Online]. Available: <http://www.aneel.gov.br>
- [2] A. Philpott, Z. Guan, J. Khazaei, G. Zakeri, "Production inefficiency of electricity markets with hydro generation," *Utilities Policy*, vol. 18, pp. 174-185, Dec. 2010.
- [3] CCEE - Electric Power Commercialization Chamber. Prices. [Online]. Available: <http://www.ccee.org.br>
- [4] IEA - International Energy Agency, *Energy market experience: Lessons from liberalized electricity markets*. Paris: OECD/IEA, 2005.
- [5] Li G., Shi J., Qu X. Modeling methods for GenCo bidding strategy optimization in the liberalized electricity spot market: A state-of-the-art review. *Energy*, vol. 36, pp. 4686-4700, 2011.
- [6] Weidlich, A.; Veit, D. A critical survey of agent-based wholesale electricity market models. *Energy Economics*, 30, pp. 1728-1759, 2008.
- [7] Watkins, C.; Dayan, P. Technical Note: Q-Learning. *Machine Learning*, 8, pp. 279-292, 1992.
- [8] M. Guo; Y. Liu; J. Malec, "A new Q-learning algorithm based on the metropolis criterion", *IEEE Transactions on Systems, Man, and Cybernetics, Part B*, vol. 34, n. 5, pp. 2140-2143, Oct. 2004.
- [9] Rangel, L. F. Competition policy and regulation in hydro-dominated electricity markets. *Energy Policy* 36 (2008) 1292–1302.