Strategic bidding in Colombian Electricity market using a multi-agent learning approach

L. Gallego, *Member, IEEE,* O. Duarte, *Member, IEEE,* A. Delgadillo, *Member, IEEE,*

***Abstract*—In this paper, a multi-agent model of an electricity market is proposed using the Agent-based Computational Economics (ACE) methodology. The proposed methodology for modeling the bidding price behavior of Generation Companies (GENCOs) is based on a reinforcement learning algorithm (Q-Learning) that uses some soft computing techniques to face the discovery of a complex function among bidding prices, states and profits. The proposed model also comprise the power system operation of a large-scale system by simulating Optimal DC Power Flows (DCOPF) in order to obtain real dispatches of agents and a mapping from action space (*bidding strategies*) to *quantities dispatched*. In this model, agents are provided with learning capabilities so that they learn to bid depending on market prices and their risk perception so that profits are maximized. The proposed methodology is applied on colombian power market and some results about bidding strategies dynamics are shown. In addition, a new index defined as rate of market exploitation is introduced in order to characterize the agents bidding behavior.**

***Index Terms*—Agent-based Computational Economics, Bidding prices, Electricity Market, Reinforcement learning.**

I. Introduction

REINFORCEMENT learning techniques has been extensively used in artificial intelligence research and recently, this technique has gained a great importance in modeling several dynamics of social and economic interactions. As an example of such interactions, power markets dynamics have been modeled by means of a recent technique known as Agent Computational Economics (ACE) models. Broadly speaking, Agent Computational Economics models represent the dynamics of the strategic behavior through a "learning" capability provided to agents when they are exposed to an environment of iterative interaction [6]. To put it in other words, agents are able to learn from their past experience, improving their decision-making process and adapting their behavior to a changing environment (strategies of their competitors, entering of new players, policy changes). This adaptation capabilities allows to implement a wide range of strategic behavior and to study the evolution of market behavior under controlled conditions, just like an experimental economic lab.

The first step in the construction of these type of models consist in describing the economic system by a set of agents (producers, consumers, brokers, regulators), specifying its initial attributes. The set of attributes of any agent may include behavioral rules, communication protocols, learning techniques, market information and agent goals [12]. In recent years, the ACE methodology has been used extensively to model the electricity markets, as these are defined as a complex adaptive systems made up of multiple agents (generators, consumers and regulators), where each participant has its own strategy, its risk preference and a making-decision model under particular market rules [11].

In [5] a simulation model is proposed based on agents that build optimal supply curves to analyze market power in the England and Wales market. Reference [3] presents a simulation model where generators agents are represented as adaptive agents involved in a repetitive daily market. These agents looks for strategies to maximize their profit based on the improvement of the profits obtained in the last iteration. Reference [10] develops the construction of an agent-based economic model of the UK electricity market to explore the possible effects of the new market mechanisms, known as NETA, in 2001. Reference [9] reports a market electricity model in which prices are set by an auction process where generators and consumers submit their bids with a discriminatory payment mechanism. Each buyer and seller determines its bid price in an adaptive way by means of a reinforcement learning algorithm. Reference [1] develops a model where the agents use learning classifier systems to improve their strategies. This model shows that agents can learn their behavior by observing historical data and research the influence of market structures in the offer prices. In reference [2] an agent-based simulation is employed to study the power market operation under two alternative pricing systems:uniform and discriminatory(pay as bid). Reference [8] uses a modified reinforcement learning algorithm based on temperature variation to obtain optimal bidding strategies. Reference [15] implements a Q-learning algorithm to maximize long term supplier's profit in a perfectively competitive market.

In this paper, a multi-agent model of an electricity market is proposed for modeling the bidding price behavior of Generation Companies (GENCOs) based on a reinforcement learning algorithm (Q-Learning). The proposed model also comprise the power system operation of a large-scale system by simulating Optimal DC Power Flows (DCOPF) in order to obtain real dispatches of agents and a mapping from *actions* (*bidding strategies*) to *quantities dispatched*. These bidding prices are obtained from a pattern recognition methodology over the bidding strategies space proposed in [7]. In this model, agents are provided with learning capabilities so that they learn to bid depending on market prices and their risk perception so that profits are maximized.

L. Gallego and O. Duarte are professors of the National University of Colombia - and active researcher of group PAAS-UN (e-mail://lgallegov@unal.edu.co) http://www.paas.unal.edu.co

978-1-4244-2218-0/08/$25.00 ©2008 IEEE.

II. THE MODEL

In this paper, an agent-based model of colombian electricity market is developed. This market is characterized by a portfolio of hydro, gas and coal generation power plants, with a strong dependence on hydrological conditions. In the proposed model, the agents are provided with a learning mechanism based on a Q-learning approach to improve their bidding price strategies. This section presents a description of agents and their learning capabilities.

A. Agents

The representation of the involved agents follows the structure shown in Fig. 1. In general terms, agents are modeled as software representations of a decision-making process that may interact in specific scenarios. These scenarios are determined by some variables that influence the actions taken by these agents.

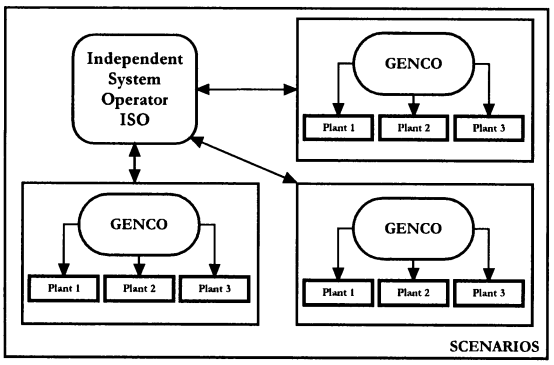

Figure 1. General Scheme of Agents

1) Generation Companies-GENCOs: In our model, the most important agent is the software representation of Generation Companies. These agents were considered asymmetric in the sense that each agent may own several power plants of different size and technology, implying that generation costs are completely different for each plant. This situation directly impact on the profits obtained by each agent.

2) Independent System Operator-ISO: The Market Operator is the software representation of the agency responsible for determining the *constrained* and *unconstrained* dispatch. In each period, this agent receives the offers from GENCOs and taking into account the power system constraints, this agent determines spot price, constrained and unconstrained generations, and some prices to compensate the existing differences between dispatches (constrained and unconstrained). Finally, these values are returned to GENCOs with the purpose of calculating profits.

On the other hand, the considered constrained dispatch consist in a DCOPF formulation. In short, DCOPF formulation assume a lossless transmission network, an homogeneous voltage profile and small enough angular differences to consider that:

$$sen(\theta_i - \theta_j) \approx \theta_i - \theta_j \tag{1}$$

Consequently, reactive power flow is neglected and active power flow in each node depends on angular differences. Thus, DCOPF problem is formulated as a mathematical optimization problem as follows:

$$min \sum_{i=1}^{n} C_i(P_{gi}) \tag{2}$$

s.a

$$P_{min} < P_{gi} < P_{max} \quad \text{Generating units capacity}$$

$$y_{ij}(\theta_i - \theta_j) < F_{ij}^{max} \quad \text{Transmission lines capacity}$$

$$\sum P_d = \sum P_g \quad \text{Active Power balance}$$

In particular, $C_i(P_{gi})$ is assumed as a linear curve with a slope equal to the bidding price of generating units. This assumption is based on a Price-Based Unit Commitment (PBUC), which is the current scheme in colombian power market.

A large-scale power system was completely modeled (*colombian power system*), including a model of economic dispatch based on a DCOPF formulation. The model was simulated using MATPOWER 3.2 which is a power system simulation tool developed in MATLAB [4]. Particularly, this tool was adapted to the implemented software that uses a class structure to represent GENCOs, Generating units and the Market Operator (ISO).

B. Learning capabilities

In general terms, the *learning* task consists in a process where agents map from the space of possible *actions* to the space of *states*, in such a way that agent goals are achieved. As a result of this mapping process, agents learn to choose an specific *action* for an specific *state* so that *rewards* are *maximized*. Broadly speaking, at each period t, agents do the following:

- Perceives the *state* x_t
- Takes an *action* a_t based on perception of state x_t and past experience
- Its behavior is reinforced by a scalar *reward* r_t, which quantifies how "good" was to choose action a_t in the state x_t
- Use a *reinforcement learning algorithm* to minimize an error raised from taking non-optimal actions given a particular state.

On the other hand, there are several types of reinforcement learning algorithms. In particular, *Q-learning* algorithm belongs to the family of *Temporal Difference Learning Methods*. This algorithm was developed by Watkins [14] to estimate the *long-term expected reward* for a given state-action pair with the appealing property that is model free[1] and can be used for on-line learning. In this algorithm each agent is provided with a matrix called *matrix Q*, which is updated on each period or iteration. Consequently, the value in a position $Q(s, a)$ of this matrix, quantifies how good was to take a particular action a

[1]Do not need an specific model (matrix of transition probabilities between states) of the markovian environment.

in the particular state s. This reinforcement learning algorithm follows the following updating rule for the values of $Q(s,a)$:

$$Q_{t+1}(s,a) \leftarrow Q_t(s,a) + \alpha[r + \gamma max_a Q_t(s_{t+1},a) - Q_t(s,a)] \quad (3)$$

where α is a learning rate reflecting the degree to which estimated Q values are updated by *new* data, and γ is a discount factor representing the weight given to future rewards.

As it can be seen, a detailed specification of *states*, *actions* and *rewards* must be fulfilled in order to model the behavior of agents by a reinforcement learning algorithm. In the case of the proposed model in this paper, these *entities* are shown in Fig. 2 and are described next.

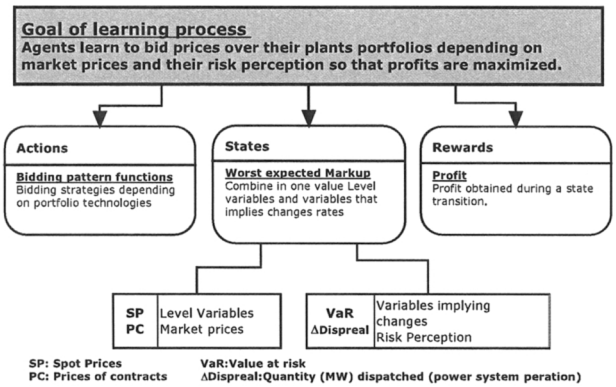

Figure 2. General scheme of the multi-agent power market proposed model

1) Actions: In the proposed model, actions consist in bidding prices of GENCOs over their plants portfolios. More specifically, actions are represented by *bidding patterns strategies* which were obtained by applying a clustering algorithm over a bidding function space. This methodology was proposed by authors in reference [7]. Therefore, each agent has a collection of *bidding pattern functions*, which indicates them how to bid each one of the generating units they own. On the other hand, unlike the model developed by authors in reference [6], this pattern recognition achieved over the continuous domain of bidding prices, represents a big reduction on the space search, which makes possible a more effective learning process.

2) Rewards: In this model, rewards are represented by *profits* obtained from bidding an specific *bidding function* in a particular *state*. The agent profits are determined by the difference between incomes and costs, which are completely different for each agent. In particular, generation costs for each plant were built as linear curves whose slope is the average generation cost. This data were obtained from reference [13]. In this manner, daily agent profits are calculated as follows:

$$\pi = SI + CI + DI - TC \quad (4)$$

where,
SI: Spot Market Income
CI: Contract Market Income
DI: Income earned by the differences between constrained and unconstrained dispatch
TC: Total Average Cost

3) States: In general, *states* are representations of the perception about the environment. In this manner, two possible kinds of *states* may be established depending on the effect of the actions taken by agents: *exogenous* and *endogenous*.

The former refers to states that are *independent* of *actions*, that is, actions do not modify states. For instance, hydrology may be considered as a variable defining a particular state in the sense that hydrology modify the actions to be selected by agent, however, bidding prices do not change hydrology. This type of states has been considered by authors in other works [6].

On the contrary, *endogenous states* are *modified by actions*, for instance, spot prices are modified by bidding prices. In particular, states in this model are considered as *endogenous*.

The *state* in this model is represented by the *worst expected markup*. In short, this *worst expected markup* summarize two different types of variables: *level* and *change* variables. *Level* variables are represented by market prices and in particular, *spot prices* and *contract prices*. On the other hand, *change* variables are represented by *reductions* in *earned incomes* because of market prices and dispatched quantities *variations*, which may be also modeled as *risk of price* and *risk of quantity*, respectively.

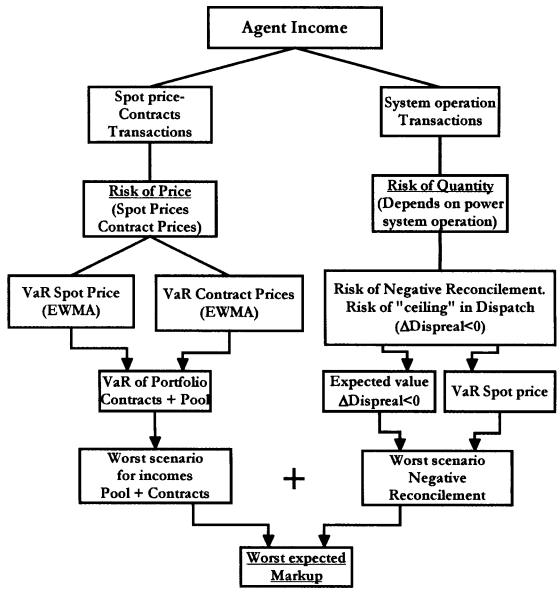

Figure 3. Risk Map of an agent in the proposed model

a) Risk of Price: Risk of price is represented by a reduction in incomes caused by natural *market price fluctuations*. A popular form of quantifying this risk is to estimate a parameter known as *Value at Risk (VaR)*. In short, *VaR* indicates the maximum expected loss with a given probability over a given period of time. For instance, an agent may want to estimate the maximum expected loss in Contracts and Pool income with a probability of 95% for the next day. Precisely, these are the

sort of estimations that can be achieved by using *Value at Risk* indicators.

In the model, VaR calculation is achieved using an analytical method based on variance-covariance matrices. In principle, variance of price returns must be calculated to obtain VaR values. However, calculation of variance may be weighted in such a way that most recent data influence in a stronger manner this calculation. Consequently, an Exponentially Weighted Moving Averages (EWMA), was used to calculate covariance matrices. To sum up, VaR values were calculated to estimate the changes in a portfolio composed by Pool and Contracts incomes for each agent. As a result, an expected loss in this income is obtained for the next day. This process is detailed in Fig. 3.

b) Risk of Quantity: In the model, risk of quantity is represented by a mapping from *bidding patterns* to *quantities dispatched* (MW). Particularly, this mapping was carried out by reproducing all the hourly dispatches in a period of 4 years. This reproduction was achieved by simulating DCOPFs over a large-scale power system (colombian transmission network) with agents bidding their bidding patterns functions. These simulations were achieved considering two different scenarios: *constrained* and *unconstrained* transmission networks. As a result of these simulations, a mapping from *bidding pattern strategies* to *changes in quantities dispatched*($\Delta Dispreal$) was obtained.

It is important to highlight the meaning of *changes in quantities dispatched*($\Delta Dispreal$). In principle, ($\Delta Dispreal$) is obtained by *deducting* quantities dispatched in a *constrained* model of the network (*real case*) from quantities dispatched in an *unconstrained* model of the same network (*ideal case*). In this manner, a *negative* value for ($\Delta Dispreal$) implies that exists a *"cap"* in generation and, on the contrary, a *positive* value implies the existence of a *"window"* in generation. As a final result, each agent estimates a *probability distribution* of *changes in quantities dispatched* for each one of their bidding strategies. In this way, each agent has a perception about the effect of transmission network constraints over bidding strategies and quantities dispatched.

In the case of this model, *negative* values for $\Delta Dispreal$ are related to risk of quantity because these negative values may cause generation caps and so, may cause a *reduction* in incomes. Therefore, risk of quantity is calculated by multiplying the expected value of ($\Delta Dispreal$) and a *reconcilement* price, which corresponds to spot prices in the case of the colombian power market. As a result, an expected income reduction is obtained. This calculation is detailed in Fig. 3.

c) Worst Expected Markup: As a consequence of estimating the risk associated to quantities and prices, two expected income reductions are obtained. Therefore, Worst Expected Markup (WEM) is obtained as follows:

$$WEM = \frac{[CI - (ER_{RP} + ER_{RQ})]/Q}{C_Q} \tag{5}$$

where CI represents current income, ER_{RP} represents the expected income reduction produced by considering risk of price, ER_{RQ} represents the expected income reduction produced by considering risk of quantity, C_Q represents average costs and Q represents current quantity dispatched.

4) Model Dynamics: Once actions, states and rewards are defined, the implemented reinforcement algorithm based on Qlearning should be described as an iterative process where matrix Q is constantly updated assuring that all the possible states are visited. In addition, a discretization over states is needed, therefore, Worst Expected Markups for each agent where divided in 8 intervals using a clustering algorithm known as *k-means*. Consequently, each agent exhibits different values of Worst Expected Markup for each state. The dynamics of the reinforcement algorithm is described in Algorithm 1 and is shown in Fig. 4.

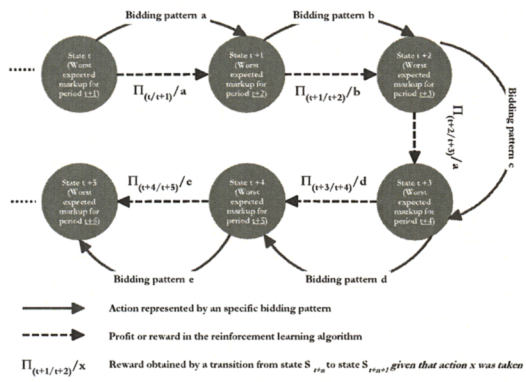

Figure 4. Learning dynamics in the proposed model

Algorithm 1 Q-learning algorithm proposed for modeling bidding behavior of agents

1: Initialize $Q(s, a)$ // Randomly initialize Q matrix
Repeat for ever day:
2: Check state s_t // (Estimate worst expected markup)
3: Choose an action a in the state s_t // (Sel. a bidding strategy p.e. e-greedy)
4: Bid the chosen strategy p
5: Power system operation // (Optimal DC power flows hourly)
6: At the end of the current day calculate Profits (r) and estimate s_{t+1}
7: Update $Q(s, a)$ as:
8: $Q_{t+1}(s, a) \leftarrow Q_t(s, a) + \alpha\,[r + \gamma max_a Q_t(s_{t+1}, a) - Q_t(s, a)]$
9: $s_t \leftarrow s_{t+1}$

III. RESULTS

A. Dynamics of Bidding strategies

One of the most interesting results obtained with this model was a plataform to observe the dynamics of bidding strategies. In principle, the goal in learning tasks is to construct a mapping from actions to states. With the proposed model it is possible to observe the bidding behavior for each one of the possible states. In this regard, figures 5 and 6 shows the dynamics of different bidding strategies of two GENCOs for the most visited states. These strategies are bidding functions over their plants portfolio. For instance, agent illustrated in Fig. 5 has a hydro portfolio of 6 plants and a thermal portfolio of 1 plant. The prices shown in these figures (*Phyd and Pterm*) are just weighted averages of this bidding functions used for the purpose of comparing the different strategies, but they

really represent a different bidding price for each one of the generating units.

The *x-axis* represents the number of visits to the specific state and the *y-axis* represents the *probability* of choosing a particular bidding strategy as states are visited. For instance, in Fig. 5 is presented the *bidding dynamics* for an agent with 8 strategies. As it can be seen, a *prevalence* of some strategies over others *emerge* from the constant *interaction* with its competitors. For example, bidding strategy No. 8 seems to be the *dominant* strategy when state No. 3, 4 and 5 are visited. This process shows an *stability* in the chosen strategies which also indicates that a sort of *convergence* is reached. On the other hand, it is possible to distinguish that some strategies are also *extinguished* as the states are visited. Even more, a process of *exploration* it is always present in the beginning of the simulations, implying that agents search intensively for their best actions and after some periods, they begin to *exploit* the *acquired knowledge* about bidding strategies and states. As a final result it is possible to establish a probability distribution over bidding strategies in each state, which is the final goal of learning, that is, a mapping from actions to states. This probability distribution could be a great contribution to model bidding behavior with some analytical tools as non-cooperative games with mixed strategies. Furthermore, this result may be interpreted as an equilibrium constructed by the repetitive interaction of agents that may be impossible to discover by using some traditional analytical tools because of the complexity involved in a large scale power market.

Another interesting experiment consist in changing the value of some variables that may influence the agents bidding behavior. In principle, agents face a constant scenario of hydrology, contract prices and demand. These variables were also discretized using the *k-means* algorithm. In this regard, it could be interesting to investigate the adaptivity of agents behavior when a sudden change in these constant scenarios is produced. Following this approach, an experiment was designed to expose the adaptability of agents when contract prices changes from low to high values. The results are shown in Fig. 7. This Fig. clearly reveals a change in the *dominant* strategies caused by a sudden change in contracting prices values. This particular behavior is indicated in this figure by a red circle that coincide with the moment when this change was intentionally produced. In addition, no inertia is exhibited in the bidding behavior which confirms the idea that agents are provided with a great capability of adaptation to different scenarios.

B. Rate of market exploitation

Another interesting result was to observe general trends in bidding behavior over some aggregated market variables. In particular, spot prices may summarize the aggregated bidding behavior of agents. An example of this tendencies is shown in Fig. 8 for the daily spot price. As it can be seen, an incremental tendency is exhibited by the market as periods increase. This result may be interpreted as an exploitation of the acquired knowledge about the market, allowing to increase the daily spot price. The most important aspect of this observation is

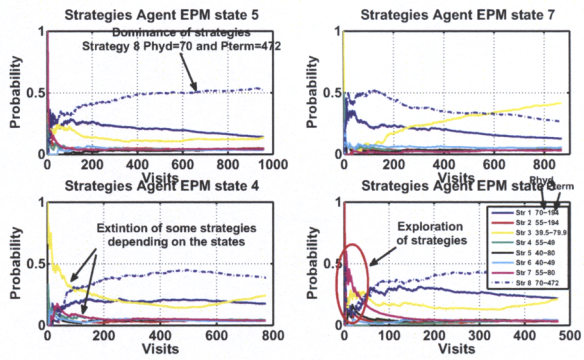

Figure 5. Dynamics of bidding strategies for a particular agent over the four most visited states

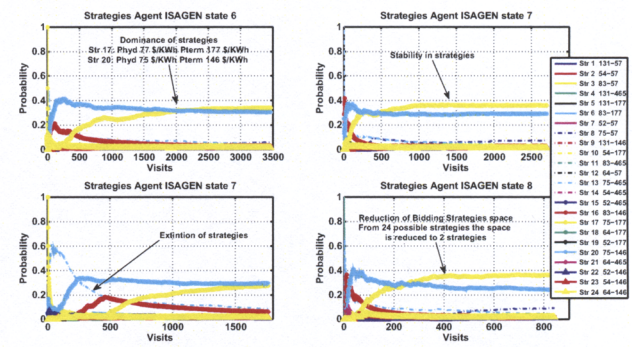

Figure 6. Dynamics of bidding strategies for a particular agent over the four most visited states

that this behavior *emerge from the repetitive and competitive interaction* among agents, which may reveal a natural and tacit collusion among GENCOs, reflecting the acquired knowledge about the market. In this sense, the slope of this tendency may be interpreted as a *rate of market exploitation* indicating how much the agents exploit the acquired knowledge to push market prices beyond competitive margins. This new parameter may be used to identify some failures in market design and also to test new market structures.

Furthermore, it is also possible to compare this *rates of exploitation* between different scenarios. Fig. 9 shows a

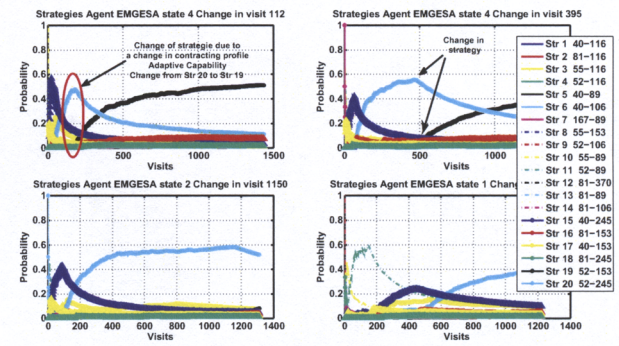

Figure 7. Adaptive capability of agents . Adaptation to a change in contract prices profiles

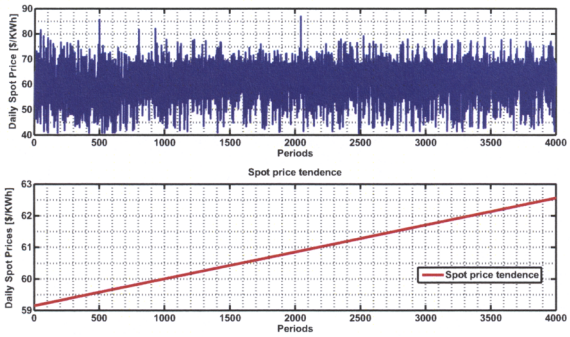

Figure 8. The Concept of Rate of exploitation

comparison between scenarios with different contract prices, revealing that under some conditions agents may have more chances to exploit their acquired knowledge about the market.

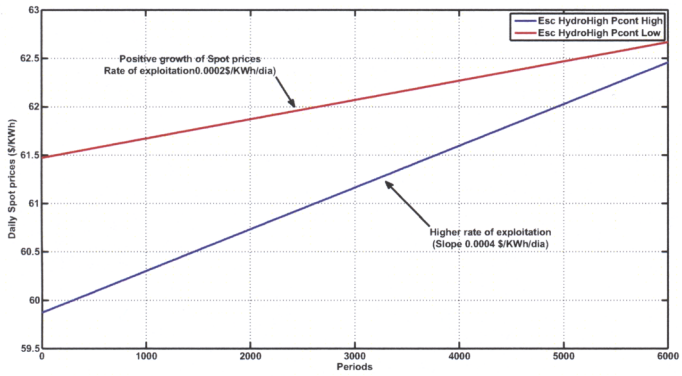

Figure 9. Comparison of Rates of exploitation between scenarios with different prices of contracts

C. *Analysis of scenarios*

As it was mentioned before, some constant scenarios were considered under different experiments. The designed experiments are shown in Fig. 10 and as it can be seen, each scenario is properly named. Of course, many different analysis may be achieved given the great amount of information available for each one of the ten GENCOs considered. However, for the purpose of this paper, only spot price and rate of exploration results are presented in Table I.

As it is expected, daily spot prices strongly depends on demand given that the estimation of this price is the result of an unconstrained dispatch to supply national demand, therefore, the higher the demand, the higher the spot prices. In addition, the effect of hydrology is more evident when hydrological conditions are high, causing that daily spot prices are a little bit lower. Anyway, apparently the effect of hydrology on spot prices is not so strong when hydrological conditions are low. A similar result was obtained by the authors in reference [7].

On the other hand, rates of exploitation do not exhibit a clear dependence on scenarios. However, the most interesting results

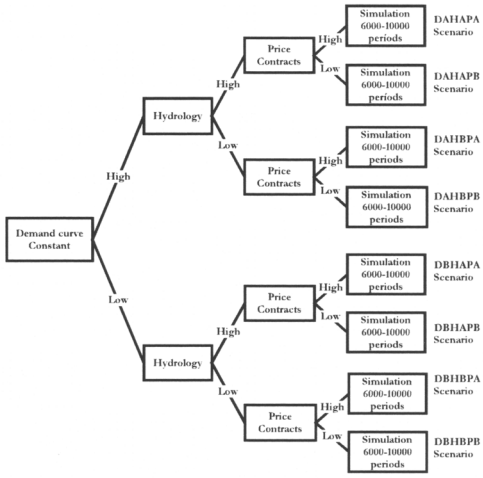

Figure 10. Experiments achieved with the proposed multi-agent model

Table I
COMPARISON OF RATES OF EXPLOITATION FOR THE DIFFERENT SIMULATED SCENARIOS

Scenario	Spot price [$/KWh]	Deviation Standard[$/KWh]	Rate of exploitation [($/KWh)/day]
DAHAPA	61.2	8.0	0.0004
DAHAPB	62.1	8.1	0.0002
DAHBPA	62.4	7.0	0.001
DAHBPB	62.8	7.7	0.0007
DBHAPA	57.5	7.4	0.0005
DBHAPB	56.5	7.0	-0.0004
DBHBPA	60.9	7.3	0.0009
DBHBPB	62.3	7.6	0.0004

are highlighted in red color. It is interesting to observe that in almost all scenarios there is a positive rate of exploitation, revealing that agents find the way of exploiting the acquired knowledge most of the time. In addition, in only one scenario this rate of exploitation is negative. This case takes place under high hydrological conditions and low prices of contracts, possibly explained by an increasing competence among agents when low cost technologies are predominant and low prices in contracts focus the competence on the pool market. On the contrary, when there is no enough water in reservoirs and prices of contracts are high, agents find an opportunity to take advantage of the knowledge acquired by pushing up daily spot prices.

IV. CONCLUSIONS

The results of this paper are focused on the advantages of ACE methodologies that allows to discover global trends in the behavior of bidding prices (price paths) under specific structural market conditions and repeated interactions. In this particular model, agents are provided with *learning* capabilities so that they learn to bid depending on market prices and their risk perception in order that profits are maximized.

On the other hand, the proposed multi-agent model also comprise the power system operation of a large-scale system

by simulating Optimal DC Power Flows (DCOPF) in order to obtain real dispatches of agents and a mapping from *actions* (*bidding strategies*) to *quantities dispatched*. As a final result, each agent estimates a *probability distribution* of *changes in quantities dispatched* for each one of their bidding strategies. In this way, each agent has a perception about the effect of transmission network constraints over bidding strategies and quantities dispatched.

In addition, the dynamics of bidding strategies reveals a constant trade-off between exploitation and exploration. Some strategies are completely extinguished and some others seem to be dominant depending on the state. Furthermore, agents are provided with an adaptive capability which allows them to response against sudden changes in scenarios. As a final result, some probability distribution over bidding strategies in each state can be established, that is, a mapping from actions to states. Furthermore, this result may be interpreted as an equilibrium constructed by the repetitive interaction of agents that may be impossible to discover by using some traditional analytical tools because of the complexity involved in a large scale power market.

Finally, the concept of *rate of exploitation* is introduced to describe the degree to which agents exploits the acquired knowledge about the market, allowing to increase market prices. This behavior *emerge from the repetitive and competitive interaction* among agents, possibly revealing a *natural* and *tacit collusion*. In this sense, this new parameter may indicate how much the agents exploit the acquired knowledge to push market prices beyond competitive margins. This new tool may be used to identify some failures in market design and also to test new market structures, just as a virtual economic lab.

Acknowledgment

The authors would like to thank to the research group PAAS UN of National University of Colombia and National Science Agency of Colombia COLCIENCIAS for the supporting of this work, as well as XM SA, the Colombian Independent Operator, and UPME for providing the historical market data and the colombian power system data, respectively.

References

[1] A. Bagnall and G. Smith, "A multi-agent model of the uk market in electricity generation," *IEEE Transactions on Evolutionary Computation*, vol. 9, no. 5, pp. 522–536, Oct 2005.

[2] A. Bakistzis and A. Tellidou, "Agent-based simulation of power markets under uniform and pay-as-bid pricing rules using reinforcement learning," in *2006 Power Systems Conference and Exposition*, 2006.

[3] J. Bower and D. Bunn, "A model-based comparison of pool and bilateral market mechanisms for electricity trading," *Energy Journal*, vol. 21, no. 3, pp. –, 2000.

[4] R. Z. Carlos Murillo and D. Gan, "Matpower 3.2," http://www.pserc.cornell.edu, 2007.

[5] J. Day and D. Bunn, "Divestiture of generation assets in the electricity pool of england and wales: A computational approach to analyzing market power," *Journal of Regulatory Economics*, vol. 19, no. 2, pp. 123–141, 2001.

[6] L. Gallego and A. Delgadillo, "Agent learning methodology for generators in an electricity market," IEEE Power Engineering Society. IEEE PES General Meeting, July 2008.

[7] L. Gallego and O. Duarte, "Modeling of bidding prices using soft computing techniques," *2008 IEEE PES Transmission and Distribution Conference and Exposition*, 2008.

[8] M. Naghibi, "Application of q-learning with temperature variation for bidding strategies in market based power systems," *Energy Conversion and Management*, no. 47, pp. 1529–1538, 2006.

[9] J. Nicolaisen, V. Petrov, and L. Tesfatsion, "Market power and efficiency in a computational electricity market with discriminatory double-auction pricing," *IEEE Transactions on Evolutionary Computation*, vol. 5, pp. 504–523, 2001.

[10] F. Oliveira and D. Bunn, "Agent-based simulation: An application to the new electricity trading arrangements of england and wales," *IEEE Transactions on Evolutionary Computation*, vol. 5, pp. 493–503, 2001.

[11] Y. Shun-kun and Y. Jia-hai, "Agent-based computational economics: Methodology and its application in electricity market research," in *IEEE PES General Meeting*, July 2005.

[12] L. Tesfatsion, "Agent-based computational economics: Growing economies from the bottom up," *Artficial Life*, vol. 8, no. 1, pp. 55–82, 2002.

[13] UPME, "Costos indicativos de generación eléctrica en colombia," Unidad de Planeación Minero Energetica, Tech. Rep., 2005.

[14] C. J. . Watkins, "Learning from delayed rewards," Ph.D. dissertation, University of Cambridge, Psychology Department, 1989.

[15] G. Xiong and T. Hashiyama, "An electricity supplier bidding strategy through q-learning," . I. Power Engineering Society Summer Meeting, Ed.

Luis Gallego PhD c. MSc Eng. Luis Eduardo Gallego Vega was born in Bogota, Colombia, in 1976. He finished his undergraduate and Master studies in the Department of Electrical Engineering of the National University of Colombia, actually is candidate to obtain a PhD in Electrical Engineering in the same institution. Eng. Gallego has been researcher in the Research Program of Acquisition and Analysis of Electromagnetic Signals of the National University of Colombia - PAAS -UN since 2000, working in research projects mainly related to power quality, power markets and lightning protection. Eng. Gallego has been enrolled in teaching activities related to power quality and power markets as professor of National University of Colombia. His research interests are power quality analysis, power markets and computational intelligence applied to power systems modeling.

Andres Delgadillo Eng. Andres Delgadillo was born in Bogota, Colombia, in 1983. He finished his undergraduate studies in the Department of Electrical Engineering of the National University of Colombia. He finished the master program in Economics in the same institution. Eng. Delgadillo has been research assistant in the Research Program of Acquisition and Analysis of Electromagnetic Signals of the National University of Colombia - PAAS - UN Group since 2004. His research interests are electricity markets, soft computing techniques and power quality analysis.