Adaptive Learning in Agents Behaviour: A Framework for Electricity Markets Simulation (Case Study)

Tiago Pinto[a], Zita Vale[a, *], Tiago M. Sousa[a], Isabel Praça[a], Gabriel Santos[a], Hugo Morais[b]

[a]*GECAD - Knowledge Engineering and Decision Support Research Centre - Polytechnic of Porto (IPP)*
R. Dr. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal
[b]*AUTomation and Control Group – Technical University of Denmark (DTU)*
Richard Petersens Plads, Building 326, 118

Contents

1. Market Negotiations' Specification

The spot or day-ahead market is a daily basis functioning market [1], where players negotiate electric power for each hour, or half hour of the following day. Such markets are structured to consider production fluctuations as well as differences in production costs of distinct units.

In this market, each participating entity must present their selling or buying proposals for each of the 24 hourly periods of a day. These proposals or bids are typically composed by a tuple (power, price), with different meanings, whether they come from buyers or sellers, respectively: power stands for amount of power to be bought or sold, and price is the maximum accepted price or minimum selling price.

*Corresponding author: Zita Vale is with GECAD, Dr. António Bernardino de Almeida, 431, 4200-072 Porto, Portugal; Tel.: +351 22 8340500; Fax: +351 22 8321159, WebSite: http://www.gecad.isep.ipp.pt /; E-mail: zav@isep.ipp.pt

When the negotiation is finished, an economic dispatch for each period is set by the market operator. At the end of each period the market operator uses a market-clearing tool establishing the market price – a unique price that will be applied to all transactions of this period.

In market pools, the most common type of negotiation is a standard uniform auction. MIBEL day-ahead spot market works as a symmetric market, where both suppliers and consumers both submit bids. The market operator orders the selling and demand offers: selling bids start with the lowest price and move up, and demand bids start with the highest price and move down. Then, the proposed bids form the supply and demand step curves, and the point at which both curves intersect determines the market price, paid to all accepted supplier and consumers. The bids of every supplier offering prices lower than the established market price and every consumer offering prices higher than the market price will be accepted. Figure 1 shows the symmetric market prices definition.

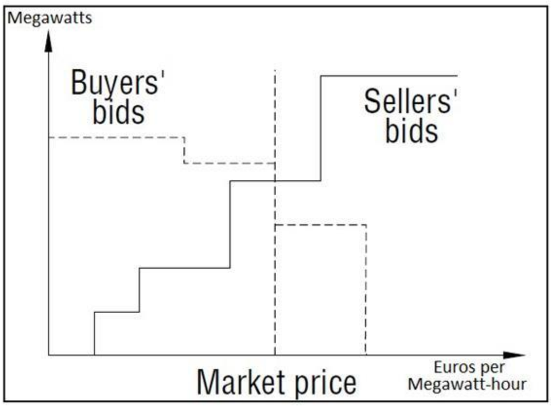


Fig. 1. Symmetric market price establishment

The profits can be improved by submitting bids that are advantageous for the player in the bidding process; i.e for a seller player, a bid price below the established market price, but still as high as possible, in order to assist in increasing the market price (origination of higher profits, through a higher market price). In the case of a buyer agent, the bid price should be above the established market price, but as low as possible, in order to reduce the cost that is paid for the bought energy.

2. Test Scenario

The test scenario involves 7 buyers and 5 sellers (3 regular sellers and 2 VPPs). This group of agents has been created with the intention of representing the Spanish reality, reduced to a smaller group, containing the essential aspects of different parts of the market, allowing a better individual analysis and study of the interactions and potentiality of each of those actors [1]. The data used in this case study has been based on real data extracted from the Iberian market operator - MIBEL [3], using an automatic data extraction that has been presented in [4]. Figure 2 presents the test scenario structure.

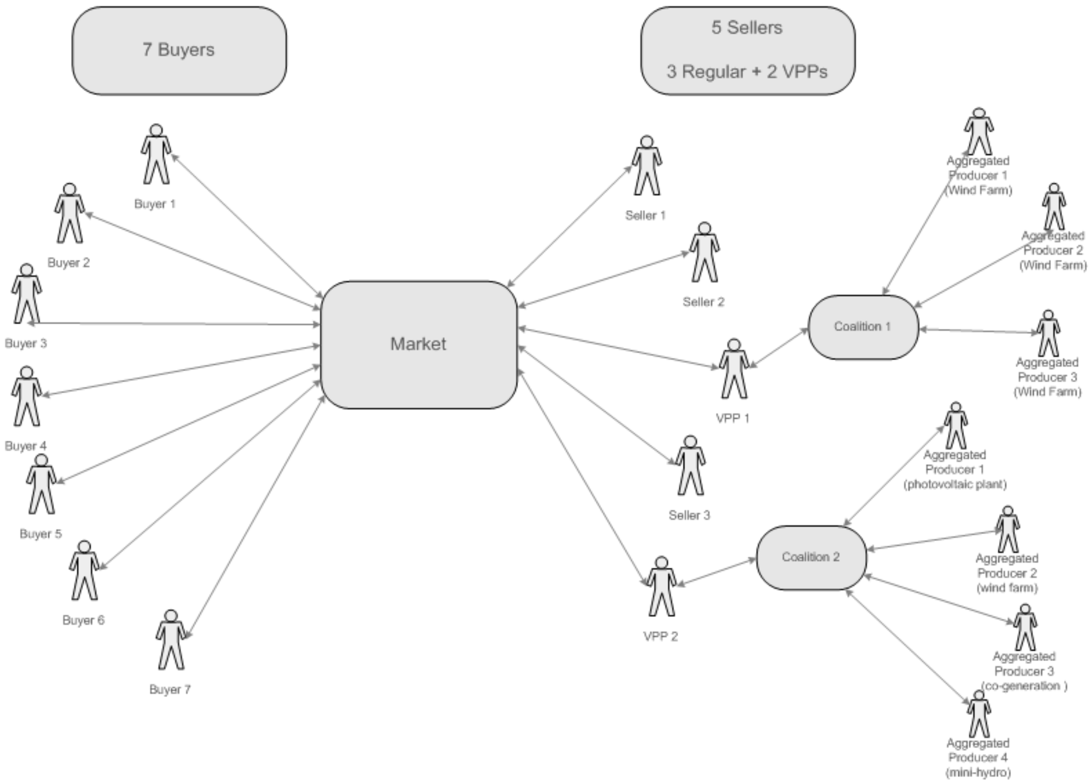


Fig. 2. Test scenario structure

The simulations consider different biddings for each agent. Seller 2, which is used as the test reference, will use each presented strategy with different parameters, depending on what it is important to be shown. Its bid power is kept constant at 50 MW for all periods of all simulated days, in order to make it easier to compare the results.

The competitor players' bids are defined as follows:

- Buyer 1 – This buyer buys power independently of the market price. The offer price is 18.30 c€/kWh (this value is much higher than average market price);
- Buyer 2 – This buyer bid price varies between two fixed prices, depending on the periods when it really needs to buy, and the ones in which the need is lower. The two variations are 10.00 and 8.00 c€/kWh;
- Buyer 3 – This buyer bid price is fixed at 4.90 c€/kWh;
- Buyer 4 – This buyer bid considers the average prices of the last four Wednesdays;
- Buyer 5 – This buyer bid considers the average prices of the last four months;
- Buyer 6 – This buyer bid considers the average prices of the last week (considering only business days);
- Buyer 7 – This buyer only buys power if market prices are lower than the usually verified market price (around 4.0 to 8.0 c€/kWh), by bidding a much lower value: 2.0 or 3.0 c€/kWh, depending on whether the current negotiation period is at a peak time of the day;
- Seller 1 – This seller needs to sell all the power that he produces. The offer price is 0.00 c€/kWh;

- Seller 3 – This seller bid considers the average prices of the last four months with an increment of 0.5 c€/kWh;
- VPP 1 – Includes four wind farms and offers a fixed value along the day. The offer price is 3.50 c€/kWh;
- VPP 2 – Includes one photovoltaic, one co-generation and one mini-hydro plants; the offer price is based on the costs of co-generation and the amount to sell is based on the total forecasted production.

In order to facilitate the understanding of the results, all the considered market negotiating entities will be attributed a constant total cost. This way, this value will not affect the results display in the analysis of each test.

All tests were performed on a computer with two Intel® Xeon® X5450 3.0GHz processors, each one with 2 cores, 4GB of random-access-memory and Windows Server 2008 32 bits operating system.

3. Specifications

This section presents the results of a set of simulations undertaken using MASCEM, with the objective of assessing the performance of ALBidS, by comparing its performance to that of all the individual strategies that have been mentioned in this paper.

The metric for comparing the performance of the methods is the profits that each is able to originate for an electricity market participant player – a seller; since the goal of ALBidS and of all strategies is to maximize the profits of a market player. The costs of production are kept constant throughout all hours of all considered days, in order to facilitate the comparison of the achieved profits.

In order to provide a suitable comparison, the same market scenario, with the exact same players, under the same circumstances, is executed repeatedly. The only variation is the behaviour of the test subject player, Seller 2. In each simulation Seller 2 uses the decision support of each of the 20 mentioned strategies, and finally of ALBidS, in order to compare the performance of all.

The simulations refer to the same 62 consecutive days (two months), starting from Saturday, 1st December, 2012, until Thursday, 31st January, 2013. The data used in this case study has been based on real data extracted from the Iberian market operator - MIBEL [3], using an automatic data extraction that has been presented in [4].

This scenario was created with the intention of representing the Iberian reality, reduced to a smaller summarized group, containing the essential aspects of different parts of the market, in order to allow a better individual analysis and study of the interactions and potentiality of each of those actors.

All comparisons are performed for three distinct cases: (i) 100% preference for the effectiveness of the strategies, *i.e*. all strategies, and consequently ALBidS, perform at their full potential; (ii) 50% preference for effectiveness, *i.e.* the execution times of the most demanding strategies are reduced, which means a reduction in their quality of results; (iii) 0% preference for the effectiveness, *i.e.*, most of the strategies are excluded from the system, while only the faster to execute are maintained.

For all three cases, the context analysis mechanism receives the input of 4, as the number of alternative contexts to be used, so that it is possible to compare the performance of the strategies when acting in different contexts. For the considered scenario, the four different contexts are easy to understand: separation between business days and weekends plus holidays; and separation from peak and off-peak consumption hours of the day. From the 62 considered days, 42 are business days and 20 are not (9 weekends, which equals 18 days, plus two holidays verified in both countries of MIBEL (Portugal and Spain): 25th December, and 1st January). From the 24 hours of the day, 5 are grouped as peak hours of consumption: from 19h to 23h; and the remaining 19 are clustered as off-peak. Therefore, the four different contexts are as follows:

- Context 1: peak hours of business days (total of 210 periods during the 62 days);
- Context 2: off-peak hours of business days (total of 798 hourly periods);
- Context 3: peak hours of non-business days (total of 100 periods);
- Context 4: off-peak hours of non-business days (total of 380 periods).

Besides the comparison of the profits that each strategy originates in each context, the strategies' confidence weights evolution throughout the time is also compared, as well as the rate each strategy is chosen as the final output of ALBidS. The choice process is undertaken using the Roth-Erev RLA (equation 2 of the paper), with a weight value W for past events of 0,4; a low value to allow a faster adaptation to new observed events.

4. Parameterizations

These parameters are the ones defined initially (for the first iteration). Most of these are updated and changed as the iterations progress. The most obvious example is the day and period that strategies receive as indication for executing their strategies accordingly.

- Context Analysis Mechanism:
    - *NC* (number of contexts) = 4.
- Efficiency/Effectiveness Balance Mechanism:
    - *Pref* (preference for effectiveness) = 100% in the first case;
    - *Pref* = 50% in the second case;
    - *Pref* = 0% in the third case.
- Roth-Erev Reinforcement Learning Algorithm:
    - *W* (weight for past experience) = 0,4.
- Average Agents:
    - *Day* = 2012-12-01;

- *Period* = 1.
- Regression Agents:
    - *Day* = 2012-12-01;
    - *Period* = 1.
- Composed Goal Directed Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1.
- Adaptive Derivative Following Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1.
- Market Price Following Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1.
- ANN Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1;
    - *Training_Limit* (amount of data used in the training process) = 240;
    - *Nodes* (number of nodes in the intermediate layer) = 3.
- AMES Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1;
    - Roth-Erev RLA
        - $q_i(0)$ (initial propensity) = 0;
        - C_i (smoothing parameter) = 0.9;
        - e (experimentation parameter) = 0.1;
        - r (recency parameter) = 0.2;
    - Simulated Annealing
        - $\mathrm{T_i}(0)$ (initial temperature) = 50;
        - λ (temperature decrease factor) = 0.9;
    - Action Domain Construction
        - $\mathrm{M1_i}$ (density control parameter) = 5;
        - $\mathrm{M2_i}$ (density control parameter) = 3;
        - $\mathrm{M3_i}$ (density control parameter) = 1;
        - $\mathrm{RIMax_i^L}$ (range-index parameter) = 0.4;
        - $\mathrm{RIMax_i^U}$ (range-index parameter) = 0.4;
        - $\mathrm{RIMin_i^C}$ (range-index parameter) = 1;
        - $\mathrm{SS_i}$ (slope-start parameter) = 1;
    - Internal Producers' Characteristics
        - Cap_i^L (minimum capacity) = 0 MW;
        - Cap_i^U (maximum capacity) = 50 MW;

        - a_i (fixed production costs) = 10.00;
        - b_i (variable production costs) = 0.
- SA-QL Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1;
    - Action Domain Definition
        - *numb(0)* (number of possible actions) = 10;
        - *int* (interval of possible bids space) = 10;
    - Simulated Annealing
        - $T_i(0)$ (initial temperature) = 50;
        - λ (temperature decrease factor) = 0.9;
    - Q-Learning
        - *r* (learning rate) = 0,5;
        - *e* (admissible error) = 0,8.
- Game Theory Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1;
    - Scenarios Definition
        - *SN* (number of considered scenarios) = 15;
        - λ (scaling factor) = for each considered scenario assumes a different value, from 0 to 1;
        - φ (scaling factor) = for each considered scenario assumes a different value, from 0 to 1;
    - Bid Definition
        - *nb* (number of considered bids) = 20;
        - *int* (bids interval) = 10.
- Error Theory Agent:
    - *Day* = 2012-12-01;
    - *Period* = 1;
    - *Type* (data analysis type) = C;
    - *Training_Limit* (amount of data used in the training process) = max;
    - *Nodes* (number of nodes in the intermediate layer) = 3.
- Economic Analysis Agent:
    - *Day* = 2012-12-01;
    - *Period = 1;*
    - *P* (total production) = 50 MW;
    - a_i (fixed production costs) = 10.00;
    - b_i (variable production costs) = 0;
    - *r* (risk factor) = 0,8.

- Determinism Theory Agent:
  - *Day* = 2012-12-01;
  - *Period* = 1;
  - Explicit Enumeration
    - *int* (increment interval) = 0,005;
  - PSO
    - *P* (number of particles) = 50;
    - *I* (number of iterations) = 1500;
    - C1 (social factor) = 0,5;
    - C2 (cognitive factor) = 0,5;
    - W (inertia) = 0,3;
    - *MaxS* (maximum speed) = 1;
    - *MinS* (minimum speed) = -1;
  - Tabu Search
    - *S* (size of the Tabu List) = 4;
    - *It* (number of iterations) = 500;
    - *Type* (type of neighbourhood generation) = Range 2;
  - Simulated Annealing
    - $T_i(0)$ (initial temperature) = 50;
    - λ (temperature decrease factor) = 0.9.
- SVM Agent:
  - *Day* = 2012-12-01;
  - *Period* = 1;
  - *Kernel* (kernel function) = e-RBF;
  - *TrainingLimit* (amount of training data) = 20;
  - *σ* (angle of the kernel function)=18;
  - *ε-insensitive* (error sensitivity) = 0;
  - *C* (kernel function limit) = ∞,
  - *offset* (offset of the kernel function) = 0;
- Simple Metalearner Agent:
  - *Day* = 2012-12-01;
  - *Period* = 1;
- Weighted Metalearner Agent:
  - *Day* = 2012-12-01;
  - *Period* = 1;
- ANN-based Metalearner Agent:
  - *Day* = 2012-12-01;
  - *Period* = 1;
  - *Training_Limit* (amount of data used in the training process) = max;
  -

- STH Metalearner Agent:
  - *Day* = 2012-12-01;
  - *Period* = 1;
  - *W* (type of used strategies weights) = RLA;
  - GA:
    - *Gen* (number of generations) = 200;
    - *Dev* (deviation) = 0,10;
    - *Cr1* (first crossover point) = 10;
    - *Cr2* (second crossover point) = 23;
    - *PM* (mutation probability) = 0,025.

5. Results for 100% preference for effectiveness

Figure 3 presents the comparison of the confidence weights of the Main Agent on each of the strategies, throughout the simulation time, for each of the four contexts.

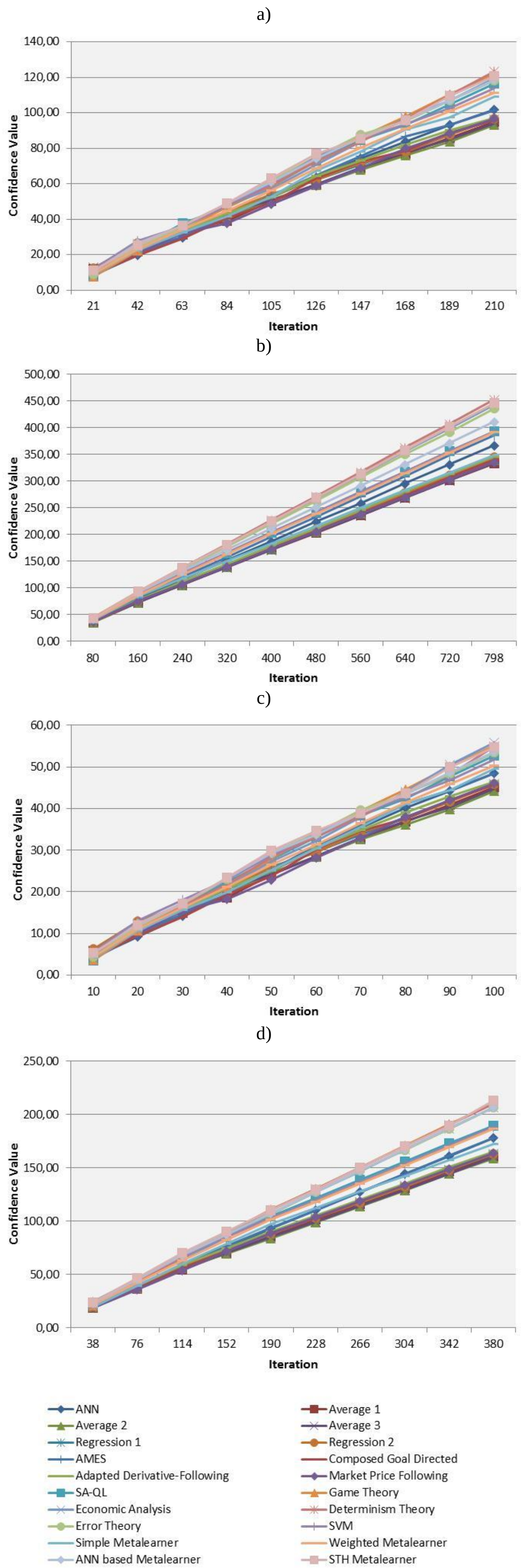


Fig. 3. Confidence weights of each of the strategies in: a) Context 1, b) Context 2, c) Context 3, d) Context 4

From Figure 3 it is visible that, by starting with the same confidence weights and with no previous learning process, strategies take a number of iterations between a significant separation can be observed. During the first iterations the same strategies have presented higher confidence weights in all contexts: the simpler strategies, with reduced or null learning capabilities, such as the strategies based on averages and regressions, and the simple metalearner. After a few iterations, the SVM starts increasing its confidence values, as this methodology requires a reduced amount of training data. As the time progresses, the strategies with more complex learning processes start improving their performance, due to the experience that they start gathering, and their learning process starts becoming more effective. In the final iterations of Context 3, in Figure 3 c), which maximum number of iterations is 100, it is visible that the more complete strategies start detaching from the simpler ones. This detachment is more evident from Figure 3 a), in Context 1, which, with its 210 iterations allows the learning process of the most complete strategies to show better performances, therefore increasing their confidence values. From the Contexts with the higher number of iterations, namely Context 2 and Context 4, in Figures 3 b) and d) respectively, not only is this detachment even more clear, as one can additionally see some intermediate sub-groups, of medium complexity strategies, such as the SA-QL, the ANN-based Metalearner, and the AMES strategy. In these contexts the simpler strategies show that their best confidence values during the first iterations are long gone, and they show the worst confidence values in the end. The group of strategies that achieves the higher confidence values in the bigger number of iterations is composed by the Game Theory strategy, the Determinism Theory, The Economic Analysis, and the STH Metalearner.

Figure 4 presents the rate in which each strategy has been chosen by the Main Agent as the final output of ALBidS, in each context.

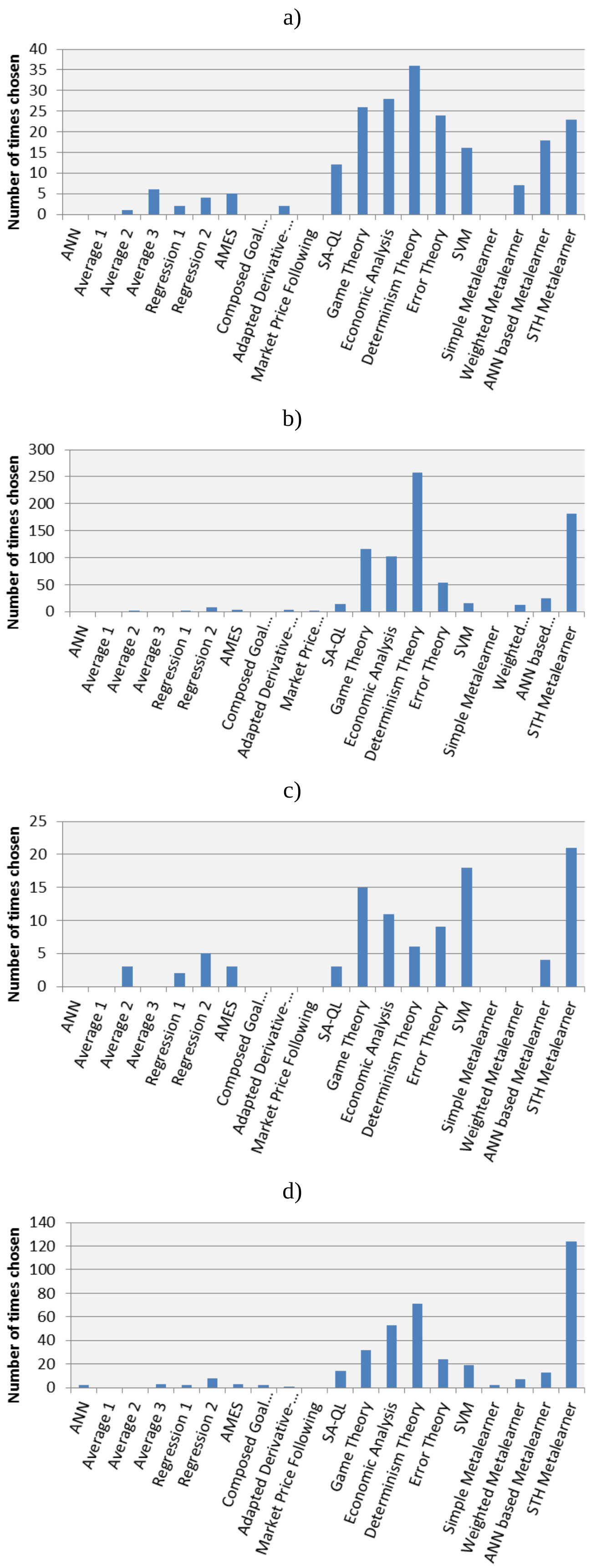
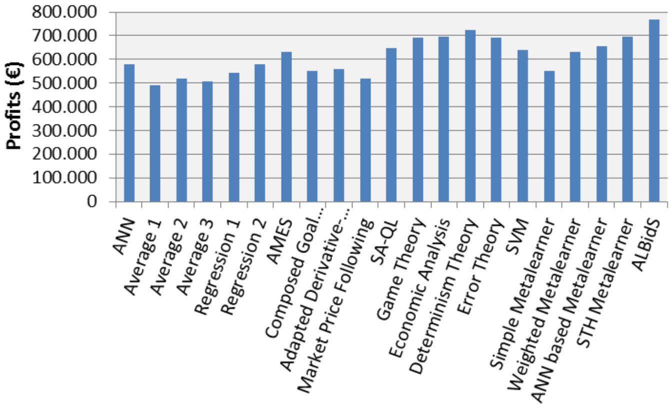


Fig. 4. Strategies choice rate for: a) Context 1, b) Context 2, c) Context 3, d) Context 4

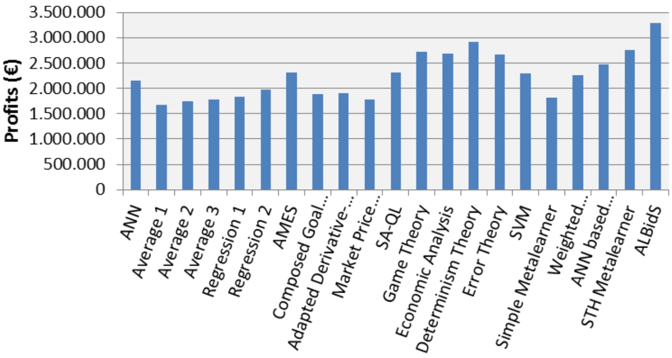
From Figure 4 it is visible that the simpler strategies, such as the ones based on averages and regressions of the market prices, have been chosen a number of times in all 4 contexts. They have been chosen during the first iterations, while the other, more complex strategies do not reach an adequate learning maturity which enables them to achieve the most advantage results. The SVM has also been chosen a considerable amount of times in all four contexts (The most evident is in Context 3 – Figure 4 c), which by presenting a smaller number of iterations, does not provide enough time for the most complex strategies to evidence themselves, therefore the SVM still manages to end up as the second more chosen strategy). This is due to the low amount of training data that this approach requires, which enables it to achieve good results from an early point. Intermediate complexity strategies such as the SA-QL, AMES, and the Weighted Metalearner, also present some amount of selections. However, it is evident that the strategies with the more complex learning capabilities, such as the Determinism Theory, the STH Metalearner, Game Theory, Economic Analysis, are the ones that end up being chosen more often, mainly in the contexts with the higher number of iterations, due to the best performance that is achieved after the learning process matures.

Figure 5 presents the comparison of the profits that each strategy has provided for the supported market player in the total of the iterations of each context.

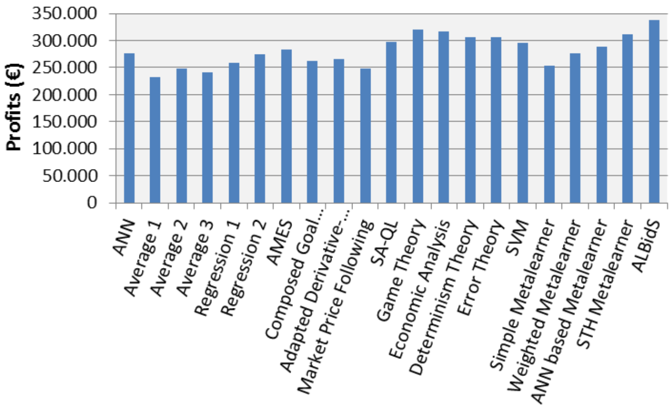
a)



b)



c)



d)

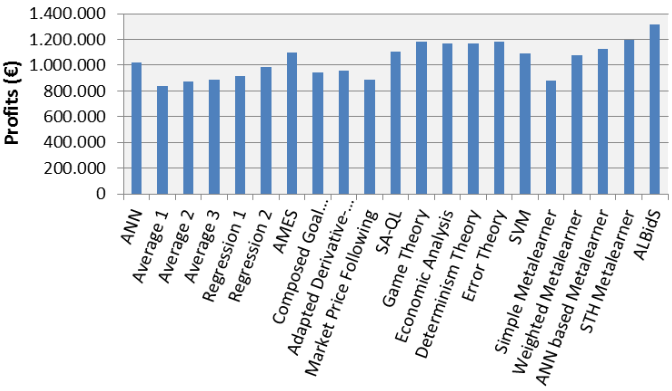


Fig. 5. Profits provided by all strategies, and by ALBidS, in: a) Context 1, b) Context 2, c) Context 3, d) Context 4

From Figure 5 it is visible that ALBidS is able to achieve higher profits than all strategies, in all four contexts. The difference between ALBidS and the individual strategies is more visible in the contexts with larger number of iterations, namely Context 2 and 4, Figures 5 b) and d) respectively. The larger number of iterations gives more time for all strategies to refine their independent learning process, and ALBidS benefits from that, as the quality of choices improves. The contexts with the least number of iterations, namely Context 3, Figure 5 c) represent a more balanced outcome between all strategies, although the difference between the quality of the strategies can still be noticed. Nevertheless, ALBidS is able to achieve higher profits than all. The good response of ALBidS in all contexts is supported by Figure 6, which shows the total profits that have been achieved by each strategy in the total of the 24 periods of the 62 considered days (total of the four contexts).

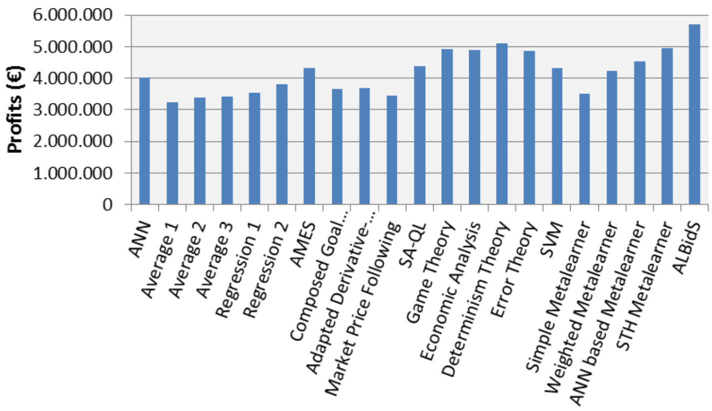


Fig. 6. Profits provided by all strategies, and by ALBidS in the total of the 62 considered days

From Figure 6 it is visible that ALBidS has been able to provide higher profits for the supported player than all the other strategies, in the total of the 62 considered days. The simpler strategies show that their capability of outperforming the more complex strategies in the first iterations is not nearly enough to compete with those, as the differences are evident in the total profits. The strategies that show the best performances, and that obtain the higher profits are the Determinism Theory, followed very closely by the STH Metalearner, and by the Game Theory, Economic Analysis, and Error Theory.

6. Results for 50% preference for effectiveness

The execution with 50% preference for effectiveness results in the reduction of the execution time of the most time demanding strategies. The ones that suffer the most from this reduction, which is verified by their decrease in quality of results, are the Game Theory strategy, the Determinism Theory, STH Metalearner. From the strategies that have achieved the best results in the case with 100% preference for effectiveness, the one that is required to reduce the execution time by a smaller amount, due to its relatively faster execution time when compared to the other more complex strategies, is the Economic Analysis. Table I presents a summary of the results that have been verified in this case.

Table I – Summary of the case with 50% for effectiveness

	Chosen Rate				Final Confidence Value			
Context	1	2	3	4	1	2	3	4
ANN	0	0	0	2	101,92	367,56	48,44	178,01
Average 1	0	0	0	0	94,98	333,49	45,15	161,10
Average 2	1	2	3	0	93,45	340,48	44,27	159,05
Average 3	6	0	0	3	94,27	333,95	44,74	160,46
Regression 1	2	2	2	2	96,27	339,91	45,92	163,18
Regression 2	4	8	5	8	96,85	343,86	45,88	162,82
AMES	5	4	3	3	101,65	385,64	48,50	177,80
Composed Goal Directed	0	0	0	2	96,00	339,66	45,62	165,08
Adapted Derivative-Following	2	3	0	1	97,31	346,03	46,46	164,86
Market Price Following	0	1	0	0	96,70	335,27	46,02	164,02
SA-QL	22	14	6	14	116,57	393,19	52,70	189,52
Game Theory	12	48	8	18	112,26	409,26	50,75	193,49
Economic Analysis	36	375	28	168	119,53	443,14	55,91	206,88
Determinism Theory	51	102	6	31	121,06	430,19	52,16	199,77
Error Theory	24	53	9	24	108,63	400,91	49,24	189,97
SVM	16	26	18	38	113,91	393,21	51,84	187,75
Simple Metalearner	0	0	0	0	97,43	340,72	46,95	165,02
Weighted Metalearner	7	12	0	7	103,58	363,65	47,80	173,17
ANN based Metalearner	6	25	4	13	110,34	383,25	49,91	192,18
STH Metalearner	16	123	8	46	112,45	414,25	50,94	198,05

From Table I it is visible that, despite the decrease in execution time, and consequent degradation in execution time, the Determinism Theory strategy has still been able to be the strategy with the higher confidence value and the most chosen strategy in Context 1. In the other contexts, the Economic Analysis has been the strategy that achieves the best results. It is also visible that, while the simpler strategies maintain their performance, the most complex ones, present a decrease in their confidence values, when compared to the case with 100% preference for effectiveness. Figure 7 presents the profits that each strategy, and ALBidS have achieved in this case, in the total of the 62 days, for the four contexts.

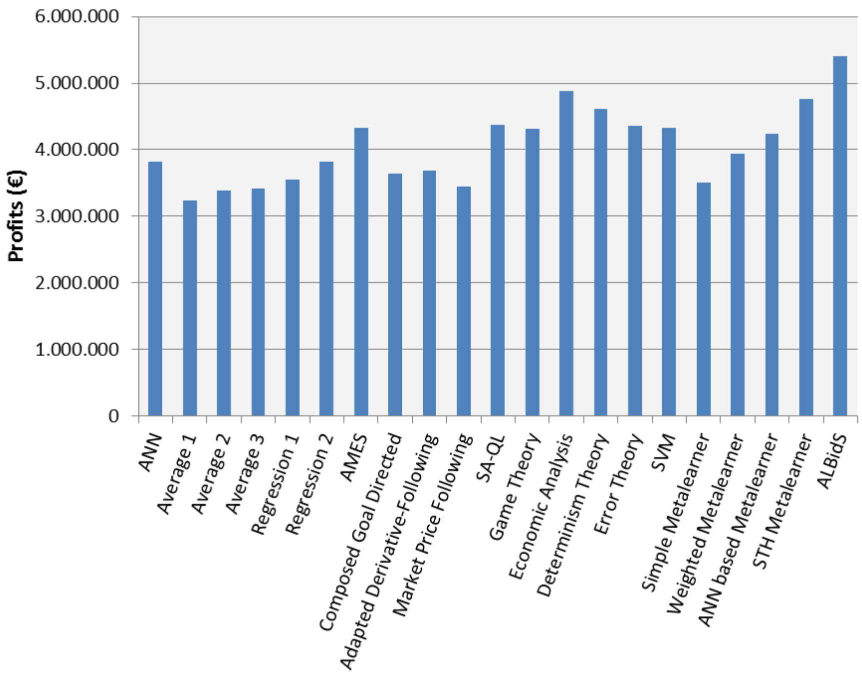


Fig. 7. Profits provided by all strategies, and by ALBidS in the total of the 62 considered days

From Figure 7 it is visible that the strategies that presented the best results in the case with 100% preference for effectiveness have decreased their achieved profits. The exception is the Economic Analysis strategy, which is now the strategy that achieves the higher profits. ALBidS is once again able to achieve higher profits than all strategies, by choosing the most appropriate strategy as the time progresses.

7. Results for 0% preference for effectiveness

Using 0% preference for the effectiveness of ALBidS means that all strategies that need more time than the execution time of MASCEM for running the market simulation, are excluded. This results in the utilization of a reduced number of strategies, i.e. only the faster to execute. Table II shows the summary of the results for the case with o% preference for effectiveness.

Table II – Summary of the case with 0% for effectiveness

	Chosen Rate				Final Confidence Value			
Context	1	2	3	4	1	2	3	4
Average 1	0	4	0	6	94,98	333,49	45,15	161,10
Average 2	1	34	8	0	93,45	340,48	44,27	159,05
Average 3	6	8	0	3	94,27	333,95	44,74	160,46
Regression 1	14	29	6	6	96,27	339,91	45,92	163,18
Regression 2	31	82	34	8	96,85	343,86	45,88	162,82
Composed Goal Directed	36	24	11	195	96,00	339,66	45,62	165,08
Adapted Derivative-Following	18	286	14	39	97,31	346,03	46,46	164,86
Market Price Following	11	8	4	6	96,70	335,27	46,02	164,02
Simple Metalearner	0	0	0	0	94,99	337,09	44,92	159,49
Weighted Metalearner	93	323	23	117	99,25	348,20	46,29	168,55

From Table II it is visible that using a reduced number of strategies leads to higher competitiveness between them. The faster response time and minor learning capabilities from the strategies supports this fact. In the first two contexts, the strategy that was chosen more often was the Weighted Metalearner. In Context 3 the most chosen strategy was a regression approach, and in Context 4, the Composed Goal Directed strategy. However, the Weighted Metalearner was the strategy that achieved the best confidence weight in the final of all iterations in all contexts except for one. Note that finishing the simulation with the higher confidence values does not necessarily mean that the strategy is chosen more often than others. It means that it is the one being chosen in the final iterations, but it may be chosen much less often during previous iterations where its confidence value was not still as high. Figure 9 presents the comparison of the achieved profits from the used strategies in the case with 0% preference for effectiveness.

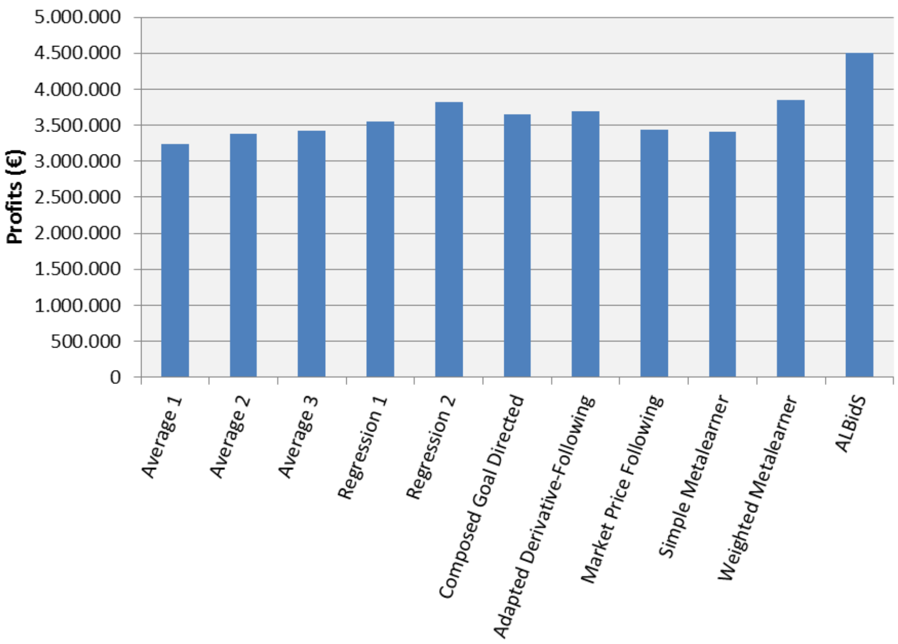


Fig. 8. Profits provided by all strategies, and by ALBidS in the total of the 62 considered days

From Figure 8 it is visible that ALBidS has achieved higher profits than all the strategies, even when using a limited number of approaches. The Weighted Metalearner was the strategy that achieved the higher profits, followed by Regression 2.

References

1. MIBEL - *Operador del Mercado Ibérico de Energia*, homepage. Available: http://www.omel.es/. Last accessed January 2014

2. Vale, Z., Pinto, T., Praça, I., Morais, H., "MASCEM - Electricity markets simulation with strategic players", IEEE Intelligent Systems, vol. 26, no. 2, pp. 54-60 Special Issue on AI in Power Systems and Energy Markets, 2011
3. MIBEL data files. Available: http://www.omie.es/aplicaciones/datosftp/datosftp.jsp?path=/. Last accessed July 2014
4. Praça, I., Sousa, T. M., Freitas, A., Pinto, T., Vale, Z., Silva, M. "Adaptive tool for automatic data collection of real electricity markets", Intelligent Agent Technology in Electricity Markets and Power Systems (IATEM) workshop of the 23rd International Conference on Database and Expert Systems Applications – DEXA 2012, 3-6 September, 2012