Article

Decision Support for Energy Contracts Negotiation with Game Theory and Adaptive Learning

Tiago Pinto [1,2,*], Zita Vale [1], Isabel Praça [1], E. J. Solteiro Pires [2] and Fernando Lopes [3]

[1] Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development (GECAD), Institute of Engineering of the Polytechnic of Porto (ISEP/IPP),
Rua Dr. António Bernardino de Almeida, 431, Porto 4200-072, Portugal;
E-Mails: zav@isep.ipp.pt (Z.V.); icp@isep.ipp.pt (I.P.)
[2] Universidade de Trás-os-Montes e Alto Douro (UTAD), Quinta de Prados, Vila Real 5000-801, Portugal; E-Mail: epires@utad.pt
[3] National Research Institute (LNEG), Estrada do Paco do Lumiar, 22, Lisbon 1649-038, Portugal;
E-Mail: fernando.lopes@lneg.pt

* Author to whom correspondence should be addressed; E-Mail: tmcfp@isep.ipp.pt;
Tel.: +351-22-8340500; Fax: +351-22-8321159.

Abstract: This paper presents a decision support methodology for electricity market players' bilateral contract negotiations. The proposed model is based on the application of game theory, using artificial intelligence to enhance decision support method's adaptive features. This model is integrated in AiD-EM (Adaptive Decision Support for Electricity Markets Negotiations), a multi-agent system that provides electricity market players with strategic behavior capabilities to improve their outcomes from energy contracts' negotiations. Although a diversity of tools that enable the study and simulation of electricity markets has emerged during the past few years, these are mostly directed to the analysis of market models and power systems' technical constraints, making them suitable tools to support decisions of market operators and regulators. However, the equally important support of market negotiating players' decisions is being highly neglected. The proposed model contributes to overcome the existing gap concerning effective and realistic decision support for electricity market negotiating entities. The proposed method is validated by realistic electricity market simulations using real data from the Iberian market operator—MIBEL. Results show that

the proposed adaptive decision support features enable electricity market players to improve their outcomes from bilateral contracts' negotiations.

Keywords: adaptive learning; bilateral contracts; decision support; electricity markets; game theory; multi-agent simulation

1. Introduction

The increasing use of renewable based generation has led to an intensive electricity markets (EM) restructuring process, which has been completely changing the EM paradigm. The privatization, liberalization and international integration of previously nationally owned systems are some examples of the transformations that have been applied [1]. These changes have progressively refined the used EM models, which began to operate using more reliable and complex models. However, EMs are still restricted to the participation of large players [2], which hardens the massive integration of renewable energy sources in the power system. This problem is being addressed in different ways in different parts of the globe [3] but some common solutions are also being globally adopted. Worldwide EMs are evolving into regional markets and some into continental scales, supporting transactions of huge amounts of electrical energy and enabling the efficient use of renewable based generation in places where it exceeds the local needs.

A reference case of this evolution is the European EM where the majority of European countries have joined together into common market operators, resulting in joint regional EM composed of several countries [4]. Additionally, in early 2015, several of these regional European EM have been coupled in a common market platform, performing in a day-ahead basis [5]. The transformation of National EM into regional and continental EM is evidenced by other examples, such as the US EM, which operates using several regional markets, e.g., California Independent System Operator (CAISO) [6] and Midcontinent Independent System Operator (MISO) [7]. In Latin-American, Brazil has also integrated all the regions in a joint EM [8]. These markets, although not representing a Continent as a whole, can be considered as continental EM due to these countries' size.

Each EM has its own rules and clearing price mechanisms, taking into account the power systems reality and the available energy mix. Some markets have the clearing mechanism based on the optimization of offers, such as most EM in the US [7] and other based on symmetric or asymmetric auctions, as is the case of most European countries. In essentially all energy markets worldwide, energy trade by means of bilateral contracts is also supported [9]. Despite the differences, market mechanisms are tending to become more and more alike in order to ease the transition towards markets unification.

Due to the constant evolution of the EM environment, including the introduction of new players [10] and changes in EM operation, it becomes essential for professionals in this area to completely understand the markets' principles and how to evaluate their investments under such a competitive environment [11]. The use of simulation tools has grown with the need for understanding those mechanisms and how the involved players' interaction affects the outcomes of the markets [12]. Artificial Intelligence (AI) plays an important role in this field, as multi-agent based simulation is particularly well fitted to analyze dynamic and adaptive systems with complex interactions among constituents, such as the EM [13].

This is supported by the several multi-agent modeling tools that can be fruitfully applied to the study of restructured wholesale power markets. Some relevant tools in this domain are AMES (Agent-based Modeling of Electricity Systems) [14], EMCAS (Electricity Market Complex Adaptive System) [15], GAPEX (Genoa Artificial Power Exchange) [16], and MASCEM (Multi-Agent Simulator of Competitive Electricity Markets) [13,17]. EM simulation platforms provide good solutions to test, validate and experiment new alternatives for market operation and players' interactions. However, these tools are usually focused on the market perspective, being valuable for market operators and regulators, while almost completely disregarding the market negotiation players' side. In fact, the decision support for electricity market negotiating players is a rather unexplored area, which should be properly addressed in order to provide the means for market players to adapt to the constantly changing EM environment, and learn how to take the most advantages out of market participation.

This paper approaches the problem of lack of decision support solutions for EM players by proposing an innovative model, based on game theory [18], to support EM players' actions when participating in bilateral contract negotiations. The problematic of bilateral negotiation is a recurrent theme in the literature of several fields, e.g., social psychology [19], economics and management science [20], international relations [21], and AI [22,23]. A relevant review on automated negotiation for computational agents with a particular focus on AI has been presented in [24]. According to this study, automated negotiation is generally composed of four phases: (i) preliminaries (the nature of negotiation); (ii) pre-negotiation (preparing and planning for negotiation); (iii) actual negotiation (moving toward agreement); and (iv) renegotiation (analyzing and improving the final agreement). However, several models consider the first two phases as a single initialization phase where all the necessary requirements, protocols and decisions that are essential before the actual negotiation process, are defined [24].

The initialization phase includes the selection of an appropriate initial strategy. The dual concern model for strategic choice in bilateral negotiation is proposed in [25]. This model stresses that negotiation strategies result from the combination of self-concern (own outcomes) and other-concern (other party's outcomes). A similar model of strategic choice is proposed in [26], and states that negotiation strategies result from the interplay of concern about own outcomes and concern about the relationship with the other party. According to [27,28] the most important pieces of data that must be considered in this phase are: (i) the intended limits and targets of the opponent(s); (ii) the negotiating history of the opponent(s); and (iii) the intended strategies of the opponent(s). Negotiators may speculate about the limits of the other parties and think stereotypically. However, they can also gather this data directly from them through the exchange of information prior or during the actual negotiation. Negotiators should also gather information about the past behavior of the other parties. Some works consider models that use information about the opposing negotiators (typically encoded into probabilistic distributions) to negotiate more effectively, such as [29,30]. Despite the efforts of these works, the authors are aware of no work on explicitly modeling the pre-negotiation step of gathering information either directly or indirectly about the opponent(s). In fact, AI researchers have traditionally neglected the pre-negotiation step of gathering information about opposing negotiators [24]. In summary, although some advances have been made regarding the pre-negotiation phase, several problems are yet far from being adequately addressed, such as the definition of models to choose the most appropriate parties to negotiate with, and how relevant information regarding competitors' history of previous negotiations

can be used to improve the decision making process, namely regarding the choice of the most suitable negotiation strategies and tactics [24].

Given the identified limitations in the field, this paper gives its contribution by proposing a novel methodology to support the decisions of bilateral contract negotiating players. The proposed model applies the game theory concept to enable the analysis of several distinct potential scenarios that the supported player is most likely to face when assuming negotiations. The alternative scenarios are created based on the historic analysis of the opponents' past actions. For this, forecasting methods are used, namely Artificial Neural Networks (ANN) [31] and Support Vector Machines (SVM) [32], among others. The forecasting results are then used by a fuzzy logic process to estimate the expected limit price values of the opponents when negotiating different amounts of power. A reputation model is also used [33], so that the decision takes into account, not only the expected negotiation prices, but also the benefit that establishing a contract with one or several players should represent to the supported player. Finally, several alternative decision methods are included to allow adaptation depending on the risk that the supported player is willing to face regarding the outcomes of the negotiation process. For this, a reinforcement learning algorithm is used, allowing the proposed model to learn which of the potential scenarios are most likely to represent a reliable approximation of the real negotiation environment that the player will face. The development of the proposed game theoretic based model aims at fulfilling an important gap in the field of bilateral contract negotiation, taking into account the advances that have been accomplished in parallel fields, such as AI and microeconomics.

The study presented in [34] reviews several models that game theory/microeconomics provide to study the problem of negotiation and bargaining. According to [34] there are two main approaches, which both model the preferences of the agents over the possible agreements by using the utility functions of von Neumann and Morgenstern [18]. The first approach is called cooperative, while the second is called non-cooperative or strategic. The bargaining problem was indeterminate by economics until the works by Nash [35,36] where the formal theory of bargaining, usually called axiomatic bargaining, was introduced. During the years several studies have extended and perfected this work; some relevant advances are the introduction of multilateral bargaining [37,38], and the modeling of the bargaining problem as an alternating—offers game [39], which approximates the mathematical models to the actual molds of most negotiations. Developments in game theoretic models have also potentiated their application to several research fields, such as energy (e.g., [40] presents an energy management model based on game theoretic assumptions, in [41] the optimization of the distribution system planning is performed using game theory, and [42] proposes a game theory based strategy for electricity market participation). Another field whose proposed game theory approaches are especially relevant is in production systems. The authors in [43] propose a novel cooperative game theoretical approach for distributed production planning of a reconfigurable enterprises. The study presented in [44] focuses on negotiation process in virtual enterprise formulation as a basic research to clarify its effective management. Each enterprise is defined as a software agent with multi-utilities and a framework of multi-agent programming with game theoretic approach is proposed as negotiation algorithm amongst the agents. A negotiation approach that enables an agent to efficiently model opponents in real-time through discrete wavelet transformation and non-linear regression with Gaussian processes is proposed in [45]. Based on the approximated model the decision-making component adaptively adjusts its utility expectations and negotiation moves. A generic framework for automated negotiation is described in [46],

which captures descriptively the social dynamics of the negotiation process. The proposed framework enables the agents to behave responsively to the changes in the environment. In the study presented in [47] an evolutionary game model is developed to observe the cooperation tendency of multi-stakeholders. The proposed game model studies the trade behavior which can be realized as strategies and payoff functions of the suppliers and manufacturers. These are relevant contributions, and their potential for application in the energy field is highly significant. However, the usual scope of application of game theoretic based models in production systems is in problems where all sides share common goals, e.g., to reach an optimal production or cooperation planning, even when considering the individual specific goals of each entity as well. In the context of application of this work the objective is different: to provide the best decision support to a single player, considering the expected behavior of the opponents. Thereby, the common goal or equilibrium between the involved entities is not relevant, just the individualist objective of the supported player: to maximize its own gain from negotiations. For this reason, it is necessary to propose new models, considering the already achieved advances in the area, but taking also into account the specific needs of the targeted problem.

After this introductory section, Section 2 presents the formulation of the addressed problem. Section 3 presents the proposed method, including the forecast based scenarios generation and the adaptive decision methods. The proposed methodology is tested and validated using realistic simulation scenarios, and the achieved results are depicted in Section 4. Finally, Section 5 provides the most relevant conclusions and contributions of this work.

2. Problem Formulation

The presented work concerns the decision support to electricity market negotiating players (both sellers and buyers when participating in bilateral negotiations. Bilateral negotiations can be undertaken with a single competitor, or by facing multiple competitor players. A seller player wishes to maximize its incomes by choosing to negotiate against the competitor or competitors that offer the best expectations of return. On the other hand, buyer agents desire to achieve as low contract prices as possible, thus the intention is to choose to negotiate against the competitor seller players that offer the lower expected contract prices. Besides the economic return of the established contracts, players should, when choosing the opponents, take also into account the expectation of opponents' ability to accomplish the contracted terms of the contract, *i.e.*, an opponent that allows achieving a very good contract price but that after the agreement is done, is not able to deliver the committed amount of power, may not be the most advantageous option. For this reason, the defined utility function considers not only and economic component, which assesses the potential economic gain form establishing a contract, but also a reputation component, which allows taking also into account with whom the contract is being settled. The balance between the two components is settled by the supported players' propensity to risk.

The formulation defined to address this problem considers the set of possible actions that the supported player is able to perform, *i.e.*, all the alternative ways to distribute the desired amount of negotiated power among the available competitor players. The proposed formulation also considers a range of alternative scenarios that represent different potential negotiation scenarios that the supported player may face when engaging the negotiation process.

The benefit for the supported player of adopting each action under each scenario is evaluated using a utility function. The utility function U_{as}, which is presented in Equation (1), allows assessing the outcome of action a in scenario s and ranges from 0 to 1.

$$U_{as} = rE'_{as} + (1 - r)R_a \qquad (1)$$

where r represents the supported player's propensity to risk, and ranges from 0 to 1. E'_{as} is the normalized economic gain (income when selling or cost when buying, assuming values from 0 to 1) of performing action a under scenario s, and R_a is the reputation component that results from negotiating the amounts defined in a with each of the corresponding competitor players (also ranging from 0 to 1).

U_{as} is defined so that a high propensity to risk considers almost exclusively the potential economic gain of the player, while neglecting the reputation of the players with whom contracts will be negotiated. On the other hand, a low value of propensity to risk means a larger weight to the reputation component, providing a larger influence to the reputation of competitor players and a lower importance to the economic gains, which results in safer and more reliable contract deals. U_{as} is thereby defined as a multi-objective function that allows maximizing the potential income of the player and minimizing the contract risk according to the preference (propensity to risk) of the supported player.

Both the reputation and the economic components range from 0 to 1, so that both components present a similar influence on U_{as}, depending only on the risk. The reputation component R_a is defined as in Equation (2).

$$R_a = \sum_{p=1}^{np} R_p \frac{A_{ap}}{TP} \qquad (2)$$

where p represents each competitor player from the total set of potential competitor players np. R_p is the reputation of player p, ranging from 0 to 1. A_{ap} is the amount of power that is allocated by action a to be negotiated with player p (ranging from 0 to TP), and TP is the total amount of power that the supported player needs to negotiate with all competitors. R_a thus represents the accumulated reputation of the players with whom negotiations will occur by undertaking action a. Each of the competitor players' reputations is relative to the percentage from the total power that is allocated to that player, *i.e.*, if the total amount of needed power is allocated to the negotiation with a single player, R_a is equal to R_p; contrarily, if the total amount is divided equally among two competitor players, the reputation of both will also contribute equally to the value of R_a.

The economic component E'_{as} defines the level of income or cost, in a scale from 0 to 1, by normalizing the actual income/cost values E_{as}, as presented in Equation (3).

$$E'_{as} = \begin{cases} \dfrac{E_{as} - E_{min}}{E_{max} - E_{min}}, & \text{when the supported player is selling} \\ \dfrac{E_{max} - E_{as}}{E_{mas} - E_{min}}, & \text{when the supproted player is buying} \end{cases} \qquad (3)$$

where E_{min} is the minimum value of E_{as} that results from all combinations action-scenario, and E_{max} is the maximum value of E_{as} from all combinations. By using Equation (3) when a player is selling, E'_{as} will assume the maximum value of 1 when the return of E_{as} is maximum, while when a player is buying, E'_{as} will assume the maximum value of 1 when the return of E_{as} is minimum (minimum buying cost).

E_{as} represents the absolute value of income/cost that results from transacting the amounts of power with the competitor players defined in action a under the expected prices from each player that result from scenario s, as defined in Equation (4).

$$E_{as} = \sum_{p=1}^{np} A_{ap} EP_{spA_{ap}} \qquad (4)$$

where $EP_{spA_{ap}}$ is the expected price of player p in scenario s, for the amount of power A_{ap}. $EP_{spA_{ap}}$ ranges from 0 to ∞, depending on the contract price forecasts.

E_{as} thus represents the total income/cost of the supported player when negotiating the amounts of power defined in action a with the competitor players defined in same action, and achieving the expected prices defined in scenario s.

The risk management capability provided by the utility function enables further adaptation to the decision making process, by not limiting the supported player to exclusively pursue the maximum possible economic gain, but also taking into account the potential benefit of the established contracts depending on the reputation of the competitor players with whom negotiations will take place.

The reputation is included in the model in order to endow the proposed decision support methodology with the capability of considering, not only the potential economic gain of the supported player in the undertaken negotiations, but also the benefit from a contract reliability standpoint. The reputation component represents the level of confidence that the supported player can have on the opponent's service, *i.e.*, in this case, the level of assurance that the opponent will fulfil the conditions established in the contract. Several works regarding the computational modeling of reputation and trust can be found in the literature, as discussed in [33], which provides an interesting review on the subject. The most recognized and globally accepted models are those resulting from the work of Sabater and Sierra [48]; the REGRET system, developed by these authors, accommodates several models for representing and assessing the reputation, trust, and credibility of different types of actors and players. The present work considers such proposed models to model the reputation of bilateral contract negotiating players.

The reputation R_p of the competitor player p is assessed from the perspective of the supported player sp. Two components are considered: the individual component $R_{sp,p}$, which represents the direct observations and experience of the supported player in regard to the subject competitor player; and the social component R_s, considering the perspective of the group in which each player is inserted, and also the prejudice regarding the player type. All reputation components range from 0 to 1. Group and player type, in the scope of this work, refer to the generation type of seller players (e.g., players that represent wind farms will tend to have a similar reputation, as they will have the same type of difficulty in fulfilling the agreed amount of power, as they are equally dependent on the wind speed), and consumer types, in case of buyer players (e.g., large industry, medium commerce, small players). The prejudice refers to the a-priori idea regarding the reliability of each player type. R_p is, therefore, defined in Equation (5).

$$R_p = w_i R_{sp,p} + w_s R_s \qquad (5)$$

where w_i and w_s are weights that are attributed to the individual and social component, respectively. The sum of both weights should be equal to 1, and these should reflect the confidence that the supported player has on its own experience and on the experience of others.

$R_{sp,p}$ is updated whenever a new observation is available. A positive or negative experience of the supported player regarding the subject competitor affects the new value of $R_{sp,p}$ as defined in Equation (6).

$$R_{sp,p} = \frac{NPE}{TNE} \tag{6}$$

where *NPE* represents the number of positive experiences and *TNE* the total number of experiences that the supported player has had with the subject competitor player.

The social component R_s allows, not only to include the opinion of others, but also to surpass the difficulties that arise from the usually very limited number of experiences that two players have directly with each other (it is unusual that players establish a large number of different contracts with the same player). Thus, the social component allows using information on similar players, and also to make use of the experience of other players regarding their personal experiences with the subject player. R_s is defined as in Equation (7).

$$R_s = w_{gp}R_{sp,Gp} + w_{gsp}R_{Gsp,p} + w_g R_{Gsp,Gp} + w_p P_s \tag{7}$$

where $R_{sp,Gp}$ represents the reputation of the subject competitor player's group form the perspective of the supported player, $R_{Gsp,p}$ represents the reputation that the subject competitor player has from the perspective of the supported player's group, $R_{Gsp,Gp}$ represents the reputation that the competitor player's group has from the eyes of the supported player's group, and P_s is the prejudice component.

$R_{sp,Gp}$ is defined by considering the individual reputation of all members that are part of the subject competitor player's group, as described in Equation (8).

$$R_{sp,Gp} = \sum_{p_i \in Gp} w_{sp,p_i} R_{sp,pi} \tag{8}$$

where $\sum_{p_i \in Gp} w_{sp,p_i} = 1$. $R_{sp,pi}$ is the reputation of member *i* of the subject competitor player's group from the standpoint of the supported player *sp*; and $w_{sp,pi}$ represents the weight that is given to each of these individual reputations of the group members. These weights can be defined according to the similarity of each group member with the subject player.

$R_{Gsp,p}$ is defined ty taking into account the opinion of each player that is part of the supported player's group in regard to the reputation of the subject competitor player, and is defined as in Equation (9).

$$R_{Gsp,p} = \sum_{gsp_i \in Gsp} w_{gsp_i,p} R_{gsp_i,p} \tag{9}$$

where $\sum_{gsp_i \in Gsp} w_{gsp_i,p} = 1$. $R_{gspi,p}$ is the reputation of the subject competitor player from the perspective of each member *i* of the supported player's group; and $w_{gspi,p}$ represents the weight that is given to each of these individual reputations of the group members. These weights can be defined according to the credibility of each group member from the perspective of the supported player.

$R_{Gsp,Gp}$ is defined by taking into account the opinion of each player that is part of the supported player's group in regard to the reputation of each member of the subject competitor player's group, and is defined as in Equation (10).

$$R_{Gsp,Gp} = \sum_{gsp_i \in Gsp} w_{gsp_i,Gp} R_{gsp_i,Gp} \tag{10}$$

where $\sum_{gsp_i \in Gsp} w_{gsp_i,Gp} = 1$. $R_{gspi,Gp}$ is the reputation of the subject competitor player's group from the perspective of each member i of the supported player's group; this reputation value is achieved by applying Equation (8) from the perspective of each player of the supported player's group. $w_{gspi,Gp}$ represents the weight that is given to each of these individual reputations of the group members. These weights can be defined according to the credibility of each group member from the supported player's perspective.

Considering the credibility of the opinions of other players to define the weights that will be attributed to their responses, requires analyzing the responses that are given from each player, and comparing them to the actual experience of the subject player; e.g., if a certain player attributes a large reputation value to the subject competitor player, and when the supported player establishes a contract with this opponent verifies that this player is not able to fulfill the contracted conditions, the supported player will not only update the reputation of the competitor player taking into account the bad experience, but will also update the credibility on the responses of the player that provided the misleading evaluation of the competitor player's reputation. The credibility update is performed by using Equation (6); however, in this case, the good or bad experience is not assessed by the player's ability in fulfilling the contracted terms, but by comparing its provided opinion regarding the opponent's reputation with the actual verified experience with the same opponent.

3. Proposed Methodology

The defined problem can be looked at as a multi-player game, where each player tries to maximize its own profit when selling and minimize its costs when buying. There is no optimal solution for this problem from the perspective of the complete set of players since the objectives of all players are contradictory, as the benefit of one player means the prejudice of the other. This is thus, exactly the nature of game theory [34]. From the global perspective of the whole set of players, the best solution would be to reach an equilibrium point between the actions and consequent outcomes of each player [36]. However, in this study the objective is to provide decision support to a single player, trying to maximize its own potential gain while completely disregarding the results of the other. For this reason, the proposed methodology adapts the general concepts of game theory to achieve the maximum potential gain for the supported player.

The proposed game theory based scenario analysis method has, therefore, the objective of supporting the decisions of bilateral contracts' negotiating players, namely concerning the pre-negotiation stage. The outputs of the proposed method are: (i) the selection of the most appropriate competitor players to negotiate with, aiming at optimizing the gain of the supported player in its transactions; (ii) the suggested amount of power that should be negotiated with each of the selected competitors in order to maximize the outcomes of the supported player; and (iii) the expected target price of each selected competitor player. These outputs are essential to enhance the results of the negotiation process, and are achieved through the application of a game theoretic based scenario analysis decision method, which evaluates the potential results of assuming different actions under distinct negotiation scenarios. The general process of the proposed methodology is illustrated by the diagram of Figure 1.
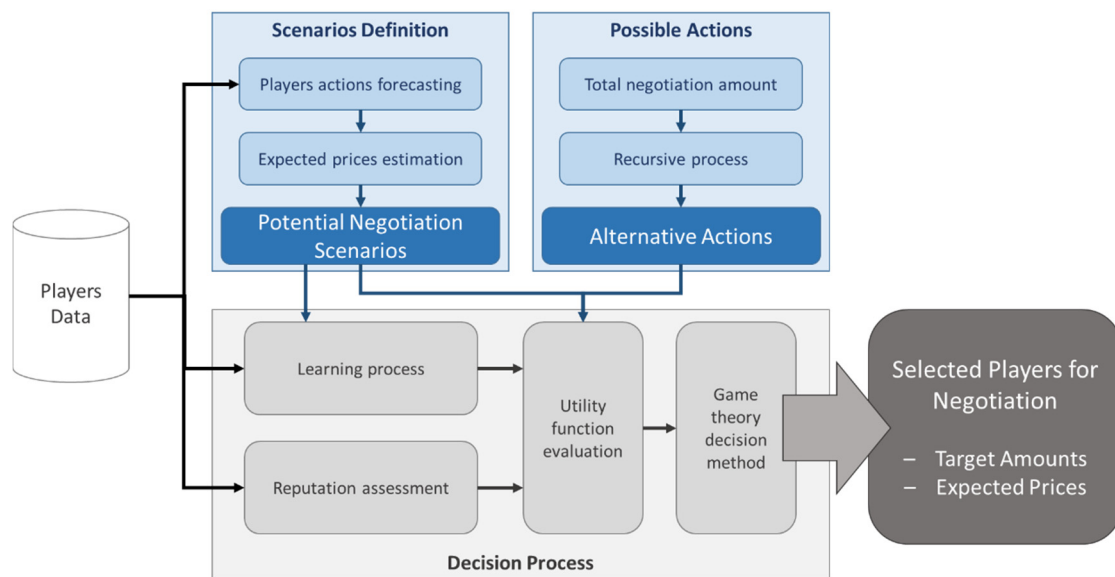
Figure 1. Decision process of the proposed methodology.

Figure 1 shows that the proposed methodology is composed by three main parts, as follows:

- Scenarios definition. As detailed in sub-Section 3.1 and considers the specification of different potential negotiation scenarios that the supported player may face when engaging the negotiation process. These alternative scenarios are created based on the analysis of the past results of the potential competitor players. Several forecasting methodologies are applied to predict the expected established contract price for each player, for different transacted amounts. Since the history log is often reduced, an estimation process is also required, to achieve the expected prices when negating amounts that are not possible to predict by the forecasting process;

- Possible actions definition. As presented in sub-Section 3.2, this process refers to the stipulation of the set of alternative actions that the supported player can undertake. The total amount of power that is intended to be negotiated is distributed among the potential competitor players by a recursive process, covering all the possible combinations;

- Decision process. The selected competitor players to negotiate with and the respective target amounts of negotiating power and expected prices result from the application of a game theoretic decision method, which uses the utility function presented in Section 2 to evaluate the potential outcome of each pair action-scenario. Hence, the result of assuming each alternative action under each scenario is calculated, using the reputation of the competitor players as the means to complement the assessment of the benefit for the supported player. Three distinct decision methods can be used: (i) a *Pessimistic* approach, (ii) an *Optimistic* approach, or (iii) the *Most probable* case. Reinforcement Learning Algorithms (RLA) [17] are used to provide the proposed method with learning capabilities, in order to perceive, throughout the time, which are the scenarios that present the higher probability of occurrence in each current context. The decision methods are presented in sub-Section 3.3.

A detailed description of the proposed pre-negotiation model is provided in the following sub-sections. As mentioned above, the pre-negotiation stage will be followed by an actual negotiation phase.

This phase can involve a simple bilateral negotiation between two players or alternatively a set of concurrent bilateral negotiations, *i.e.*, the supported player can negotiate simultaneously with several competitor players. Each negotiation will involve mainly an iterative exchange of proposals and counter-proposals regarding the prices for the energy.

3.1. Scenarios Definition

The alternative negotiation scenarios represent the alternative situations that the supported player can face when participating in bilateral contract negotiations. Scenarios are composed by the prices that are expected to be achieved from the negotiation with each of the potential competitor players, when negotiating different amounts of power. The amount is closely related to the expected price, since it is usual that a player agrees with different prices when negotiating distinct amounts of power. The expected prices from each player are calculated by several forecasting methodologies (as presented in sub-Section 3.1.1). However, predictions of expected prices for different amounts of power than those contained in the historic log are often required. Hence, an adequate estimative to is essential to reach the values that are not attainable via forecasting (as presented in sub-Section 3.1.2). Each estimative, based on the predicted prices resulting from each forecasting methodology, results in an alternative scenario.

3.1.1. Contract Price Forecasting

The prediction of competitor players' expected negotiation prices requires adequate forecasting techniques, able to provide adequate data analysis; namely of the historic of competitor players' past contracts, the amount of power that each price is associated to, and also the context to which the contract settlement refers to. The way each contract price is predicted can be approached in several ways, namely through the use of statistical methods, data mining techniques [49], neural networks (NN) [31], support vector machines (SVM) [32], or several other methods [17]. However, no method presents a better performance than all others in every situation, only in particular cases and contexts [17]. For this reason, and given that all forecasting methods are subject to some error degree, a set of different approaches is used, and the outcomes of each alternative are considered as basis to create a distinct scenario. In this way, the proposed methodology considers as many alternative scenarios as the number of different forecasting approaches.

The variables that are used by the forecasting algorithms are: the contract price, the amount of traded power, the target player and the context in which the contract settlement has occurred. The feature selection is performed by using the context analysis and definition methodology, which has been presented in [50]. This methodology is used to separate the historic data into different groups, which represent different negotiation contexts. This way, the forecasting processes consider only the data that refers to the same context as the one the decision support is intended to (e.g., negotiation of bilateral contracts for business days or weekends; directed to peak or off-peak hours of consumption, *etc.*). Thus, the contextualization of the forecasting process is enabled, resulting in forecasts that most reflect each current circumstances and context. Additionally, the forecast of a player's actions considers only the historic data of the same player. The used algorithms are listed as follows:

- A feed-forward ANN, trained with backpropagation using the historic contract prices of each subject player, for the amounts of power available in the history log.

- SVM using the exponential radial basis function (eRBF) as training kernel.

- Based on Statistical approaches. There are two strategies in this category:

  o Average of prices from the players' past actions database;
  o Regression on the contract prices historic data.

- Algorithms based on pattern analysis:

  o Sequences in the past matching the last few actions of the competitor player. In this approach are considered the sequences of at least 3 actions found along the historic of actions of this player. The sequences are treated depending on their size. The longer matches to the recent history are attributed a higher importance.
  o Most repeated sequence along the historic of actions of the target player.
  o Most recent sequence among all the found ones.

- Algorithm based on history matching. Regarding not only the player actions, but also the result they obtained. This algorithm finds the previous time that the last result happened, *i.e.*, what the player did, or how he reacted, the last time he performed the same action and got the same result.

- Algorithm returning the most repeated action of the target player. This is an efficient method for players that tend to perform recurrent actions.

The methods that are used in each decision support process depend on the requirements of the AiD-EM 2E balance management mechanism. When the requirement is the achievement of the best possible decision support results, all approaches are used, resulting in a large number of alternative scenarios to be analyzed by the decision method of the proposed methodology. On the other hand, when the execution time restraints are significant, only a few approaches are used (the faster ones to execute), so that the time demand of the decision support method is reduced.

The results of the forecasting process consider the expected contract prices for each competitor player, for the power amounts that are available in the history log. However, as explained before, the decision making process requires the expected prices of each player for each amount of negotiated power. For this reason, the estimation of the missing values is essential.

3.1.2. Contract Price Estimation

The decision making process requires the expected return prices for each possible amount of power, for each competitor player. This, however, is impracticable due to the number of possible amounts (which tends to infinite when increasing the number of decimal places of the power amount value). For this reason, a dynamic *fuzzy* variable that approximates the values of contract prices for different negotiated power amounts has been proposed in [51] and is used in the scope of this work. This methodology allows estimating the large number of historic contract prices by means of a single *fuzzy* variable, hence reducing drastically the execution time of the proposed methodology.

Historic contract information is limited, *i.e.*, the information concerns only prices for certain values of contracted power amounts. When it is necessary to achieve expected prices for contracts based on amounts of power that have never been negotiated before, these value has to be estimated. Using *fuzzy* logic, the estimative is done by defining power intervals, for which the expected price is similar. The *fuzzy* process allows smoothing the interval transition values, e.g., when negotiating 50 MW with a certain player (part of one power interval) the expected price is *X*; when negotiating 51 MW with the same player, amount of a different power interval, the expected price is *Y*. However, the difference from 50 to 51 MW is minimal, and not enough to represent a large difference in the expected price. The *fuzzy* process allows these transition values between different intervals to be smoother, avoiding abrupt price changes. Figure 2 shows the *fuzzy* variable that represents the different intervals.
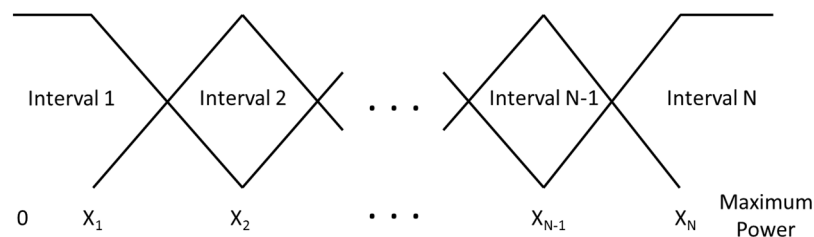

Figure 2. Dynamic *fuzzy* variable, adapted from [51].

The lower limit of the function variable is zero and the upper limit is the maximum power in the input data. Intervals are constructed according to the forecasted prices resulting from the algorithms presented in sub-Section 3.1.1. Each forecasted price defines the maximum membership value of each *fuzzy* function, *i.e.*, X_1 to X_N of Figure 2 are the power amounts for which a forecast has been performed. The limits of each function assume the value of the preceding and following price forecast, which assume membership values of zero. All membership functions are triangular, except from the first and last. The *fuzzy* variable is, therefore, dynamic, since its definition in done at runtime, depending on the number of performed forecasts, since these vary with the available historic data.

A distinct *fuzzy* variable is created to estimate the missing values that result from the forecasts of each alternative forecasting approach, for each competitor player. Each negotiation scenario is composed by the estimates of the expected contract prices of all competitor players, resulting from the values returned by each forecasting methodology.

3.2. Possible Actions Definition

Once the alternative negotiation scenarios are defined, it is necessary to identify the set of possible actions that the supported player is able to perform. The alternative actions consist in the amount of power that will be allocated to the negotiation with each competitor player. The most advantageous from these alternatives will be chosen by the decision method (presented in sub-Section 3.3) as the action that presents the most potential for the success of the supported player, by optimizing its benefit.

The definition of the alternative actions is done by distributing the total amount of desired negotiation power among the potential competitor players. Each alternative action comprehends a different distribution of the total amount among the competitor players, in a way that all combinations are represented by the different possible actions. This is achieved through the use of a recursive process that

guarantees that the total amount is always fully distributed among the players in each alternative action, and that all combinations are considered. The combinations range from allocating the total amount of power to the negotiation with a single competitor player, to the equal distribution of the desired negotiation amount among all competitor players.

The enlarged range of possible actions enables the decision method to consider the evaluation of a great number of alternative action-scenario combinations, thus facilitating the achievement of the most advantageous action for the supported player to perform, with the aim of increasing the quality of the outcomes of the negotiation process.

3.3. Decision Method

The decision method has the role of assessing the combinations action-scenario and choosing the action that presents the greatest potential benefit for the supported player. Figure 3 presents an illustration of the decision making process.
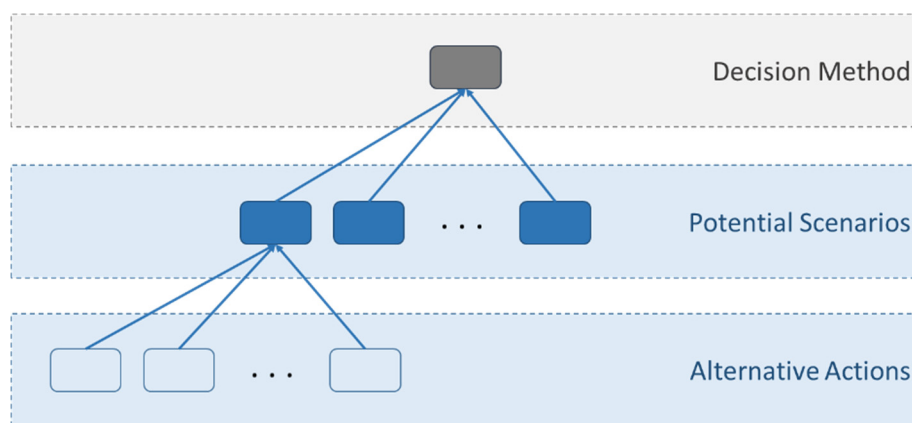

Figure 3. Decision making based on the evaluation of each action-scenario combination.

The evaluation of the action-scenario combinations is performed using the risk-based utility function presented in Section 2, which adapts the evaluation of the potential benefit of the supported player to the player's propensity to risk. The risk management is considered by including the reputation of each competitor player in the utility function evaluation; this way the supported player may choose to undertake negotiations with players that present a slightly lower potential profit, but compensate the gain by ensuring safer deals with players that present better reputations, which provide a different level of security, especially regarding the prospect of complying with the terms of the established contract. Finally, using the evaluation results of the performance of each action in each scenario, a decision method based on game theoretic concepts is used to make the final decision of what should be the best action for the player to perform. The decision method is also dependent on the player's propensity to risk, and can assume one of three approaches:

- *Pessimistic* approach. This decision method considers the usual mini-max game theoretic approach [34,42]. This method evaluates the global utility of each scenario individually, and chooses the action that presents the maximum utility (max) for the scenario with the minimum global utility (min). The global utility *GU* of scenario *j* is calculated as in Equation (11).

$$GU_{s_j} = \sum_{a \in A} U_{a_k s_j} \tag{11}$$

where a is each action from the set of all possible actions A. Hence, the global utility of scenario j is the sum of the utilities of applying each possible action under scenario j. The scenario with the lowest GU is chosen, and the action that presents the higher utility under this scenario is selected as the final action to be used by the supported player. This decision method allows the supported player to prepare for the worst case scenario it can find, and perform the safer action, which provides the best outcomes for the worst possible scenario.

- *Optimistic* approach. This approach uses the utility function evaluation of Equation (1) to find the action-scenario combination that presents the best gain among all combinations. The action that presents the higher possible gain is the one chosen as the final suggestion for the supported player to perform. This optimistic approach enables the supported player to risk, and perform the action that is able to provide the best possible gain under all scenarios.

- *Most probable* scenario. The third decision method uses a learning process to assess the probability of occurrence of each alternative scenario. The final chosen action is the one that presents the higher expected utility value for the most probable scenario. This approach allows the supported player to be prepared for the scenario that is the most likely to occur, and perform the action that should provide the best outcomes under this scenario. An adaptation of the Q-Learning algorithm [52] is proposed to undertake the learning process. Q-Learning is a very popular reinforcement learning method. It is an algorithm that allows the autonomous establishment of an interactive action policy. It is demonstrated that the Q-Learning algorithm converges to the optimal proceeding when the learning state-action pairs Q is represented in a table containing the full information of each pair value [53]. The proposed approach includes the contextualization of the learning process in order to avoid the over-generalization of the learning process, hence adapting the learning to each context. The basic concept behind the proposed Q-Learning adaption is that the learning algorithm is able to learn a function of optimal evaluation over the whole space of context-scenario pairs $c \; x \; s$. This evaluation thus defines the confidence value Q that each scenario is able to represent the actual encountered negotiation scenario s in context c. The Q function performs the mapping as in Equation (12).

$$Q : c \; x \; s \;\rightarrow U \tag{12}$$

where U is the expected utility value when selecting scenario s in context c. As long as the context and scenario states do not omit relevant information, nor introduce new information, once the optimal function Q is learned, the decision method will know precisely which scenario results on the higher future reward under each context. The reward r is attributed to each pair scenario-context in each iteration, representing the quality of this pair (how well does the scenario represent the real negotiation scenario under context c), and allows the confidence value Q to be updated after each observation. r is defined as in Equation (13).

$$r_{s,c,t} = \; 1 - norm \left| RP_{c,t,a,p} - EP_{s,c,t,a,p} \right| \tag{13}$$

where $RP_{c,t,a,p}$ represents the real price that has been established in a contract with an opponent *p*, in context *c*, in time *t*, referring to an amount of power *a*; and $EP_{s,c,t,a,p}$ is the estimation price of scenario that corresponds to the same player, amount of power and context in time *t*. All *r* values are normalized in a scale from 0 to 1, in order to allow the *Q(c, s)* function to remain under these values, so that the confidence values *Q* can be easily assumed as probabilities of scenario occurrence under a context. *Q(c, s)* is learned through by try an error, being updated every time a new observation (new contract establishment) becomes available, following Equation (14).

$$Q_{t+1}(c_t, s_t) = Q_t(c_t, s_t) + \alpha\left[r_{s,c,t} + \gamma U_t(c_{t+1}) - Q_t(c_t, s_t)\right] \tag{14}$$

where α is the learning rate; γ is the discount factor; and U_t (15) is the utility resulting from scenario *s* under context *c*, obtained using the *Q* function learned so far.

$$U_t(c_{t+1}) = \max_s Q(c_{t+1}, s) \tag{15}$$

The Q Learning algorithm is executes as follows:

- For each *c* and *s*, initialize *Q(c, s)* = 0;

- Observe new event;

- Repeat until the stopping criterion is satisfied:

   o Select the scenario that presents the higher *Q* for the current context;
   o Receive reward $r_{s,c,t}$;
   o Update *Q(c, s)*;
   o Observe new context *c'*;
   o $c \leftarrow c'$.

As the visiting of all scenario-context pairs tends to infinite, the method guarantees a generation of an estimative of Q_t which converges to the value of *Q*. In fact, the actions policy converges to the optimal policy in a finite time, however slowly. In order to accelerate the convergence process, not only the *Q* value of the chosen scenario is updated, but also that of all scenarios, since the *r* regarding all alternative scenarios can be computed by comparing the estimated prices by each scenario and the actual values that have been verified in a new contract agreement. After each updating process, all *Q* values are normalized, as in Equation (16), so that they are always kept in a scale from 0 to 1, thus facilitating the interpretation as the probability of each scenario in correctly representing the negotiation reality.

$$Q'(c,s) = \frac{Q(c,s)}{\max[Q(c,s)]} \tag{16}$$

In all three decision methods ties may occur when choosing the action with the higher utility value. In order to surpass this problem, the following tie-breaking conditions have been defined:

- When the supported player's propensity to risk is ≤0.5, the selected action is the one with the higher reputation component from all that present the same utility value;

- Otherwise, if the propensity to risk is >0.5, the selected action is the one with the higher economic gain component;

- If after the tie-break some action remain tied, the inverse condition is applied, *i.e.*, from the actions with the higher reputation component when the propensity to risk is ≤0.5, the one with the higher economic gain component is selected; the opposite is applied to the actions tied in economic gain when the propensity to risk is >0.5.

The decision method to be applied in any case can be selected directly by the supported player as input, or it can be defined dynamically depending on the risk aversion value. In this case, the *Pessimistic* decision method is used when the risk is ≤0.3; the *Optimistic* method is applied when the risk is ≥0.7, and the *Most probable* scenario is used otherwise.

4. Experimental Findings

This section presents a case study that demonstrates the advantages of using the proposed methodology. For this purpose, real data from the Iberian electricity market—MIBEL [54] has been used to assemble a historic database concerning the past log of established contracts of 37 electricity market participating players. This database is used to apply the proposed methodology and assess its performance, namely by comparing the achieved results (assignment of the negotiation amount among the set of potential competitor players) to the outcomes of allocating the total negotiation amount to a single player, which is the common approach in the pre-negotiation stage of bilateral contract negotiations.

All the simulations presented in this case study have been executed in a machine with one Intel® Xeon® E5-2620v2-2.10 GHz processor, with 12 cores, 16GB of Random-Access-Memory (RAM) and Windows 8.1 Professional.

The first part of this case study considers a set of 5 electricity market players as potential opponents, in order to facilitate the demonstration of the proposed methodology's decision support process and to allow a detailed description of results. The total negotiation amount to be allocated to be sold to the competitor players is of 10 MW. Figure 4 presents the price estimation error, using the Mean Absolute Percentage Error (MAPE) that results from the application of each of the 9 forecasting algorithms that have been presented in sub-Section 3.1.1.
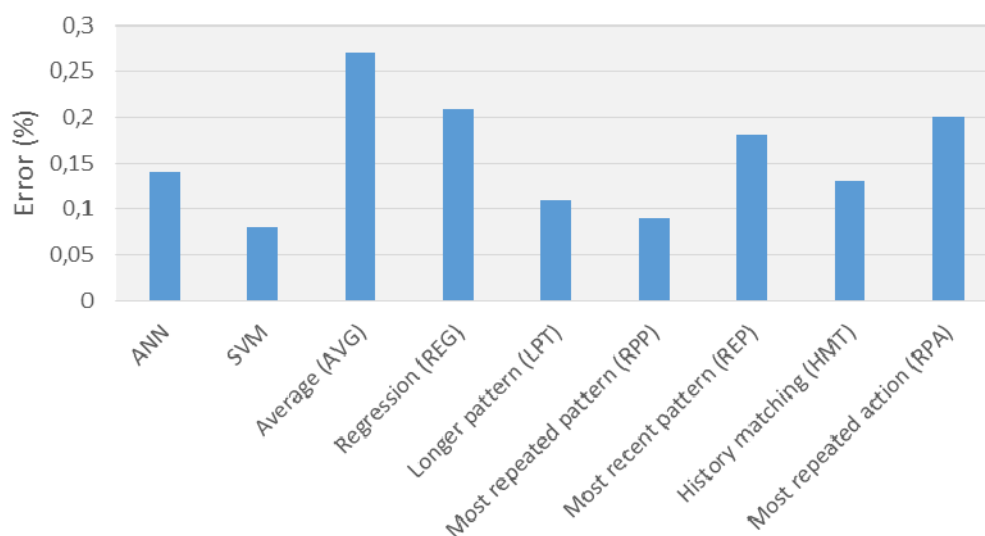

Figure 4. Mean Absolute Percentage Error (MAPE) estimation error using each of the considered forecasting algorithms.

From Figure 4 it is visible that algorithm that is able to achieve the best results for this case is the SVM, closely followed by the Most repeated pattern search and by the Longer pattern search. The worst estimations come from the application of the average and regression of the historic contracted prices.

In order to determine the *Most probable* scenario the Q-Learning based algorithm is applied, using $\alpha = 0.8$ and $\gamma = 0.2$, in order to provide the learning algorithm with a quick learning rate, with the aim at facilitating the fast adaptation to the most recent perceived events. The confidence value Q in each of the scenarios created using the forecasts resulting from each of the 9 algorithms throughout 25 observations (newly established contracts) is presented in Table 1.

Table 1. *Q* values of each scenario throughout 25 iterations.

Iteration	ANN	SVM	AVG	REG	LPT	RPP	REP	HMT	RPA
5	0.23	0.74	1	0.86	0.31	0.3	0.68	0.42	0.93
10	0.37	0.84	0.97	0.91	0.38	0.52	0.73	0.65	1
15	0.53	1	0.84	0.85	0.47	0.78	0.88	0.71	0.97
20	0.78	1	0.79	0.81	0.76	0.85	0.93	0.89	0.91
25	0.91	1	0.72	0.78	0.97	0.98	0.89	0.93	0.83

Table 1 shows that after 25 iterations the scenario that presents the higher Q value is the one based on the estimation from the forecasting results of the SVM, which corresponds to the strategy that also presents best MAPE results. During the first iterations, with very few data to train the most advanced algorithms, the approaches that have achieved the best results are the simpler ones, namely the average, regression and most repeated player's action. With the increase of the number of iterations, one can see the significant increase of the confidence value in the most complex algorithms, and a relative decrease of the Q value of the simpler strategies. Figure 5 presents the price estimation and forecasting results that compose the *Most probable* scenario (based on SVM).
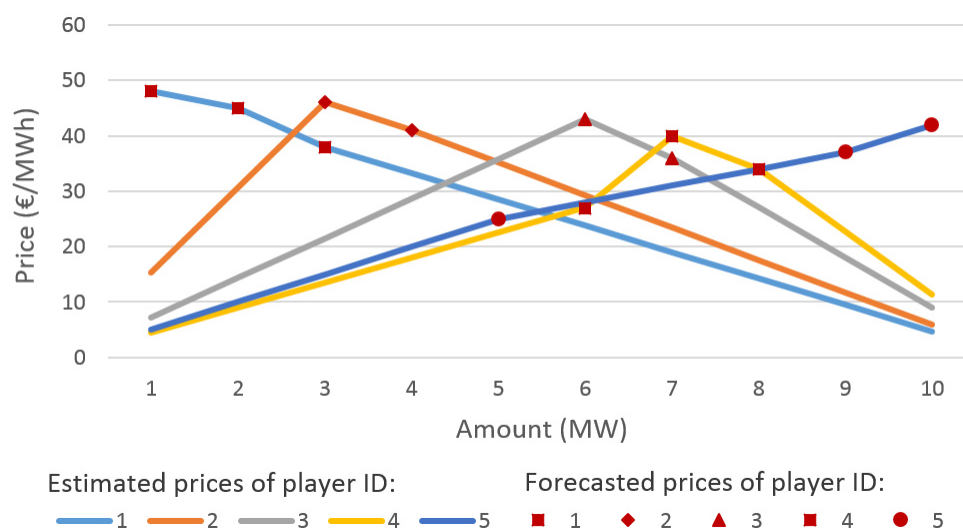

Figure 5. *Most probable* scenario estimation values.

Figure 5 shows that the forecasted (marked up) expected prices for each competitor player are very few. The reduced contract history of each player only enables the forecasting of a very strict amount of prices. This requires that the expected prices for the remaining amounts of power are estimated using

the fuzzy process introduced in sub-Section 3.1.2. It is important to notice that the estimation of prices for amounts that are not available in the historic log have a decreasing tendency towards the value of 0. Since no information is available regarding past contract settlements of a player for such amounts of power, it cannot be assumed that the player would be willing to negotiate at higher price values, thus the expectation is that these prices present a decreasing tendency, e.g., regarding opponent player ID 1, only three prices could be forecasting, concerning the amounts of 1 MW, 2 MW and 3 MW. The prices that refer to these amounts present a decreasing tendency, therefore this tendency is maintained throughout the remaining power amounts.

Executing the utility function that is necessary for the decision making process requires the use of reputation values referent to each competitor player. Since no real information regarding this aspect can be found, and using realistic values would require a sociologic study, which is out of the scope of this work, default reputation values have been attributed to each competitor player, in order to allow the test of the complete model. The reputation values that have been assigned to each competitor player are presented in Table 2.

Table 2. Reputation of the competitor players.

Player ID	1	2	3	4	5
Reputation	0.9	0.7	0.5	0.3	0.1

Table 2 shows that the reputation values assigned to the competitor players are all distinct and ranging from 0.1 to 0.9. This enables an easier verification of the influence of the reputation component on the proposed decision support model. Table 3 shows the utility values that result from the application of the proposed methodology to the present case, for different levels of risk propensity, considering the three decision methods, and the allocation of the total 10 MW to each of the 5 players.

Table 3. Utility function values for different risk propensity values.

Decision Method	Risk Propensity				
	1	0.8	0.5	0.2	0
Pessimistic	0.86	0.74	0.70	0.68	0.90
Optimistic	1.00	0.91	0.83	0.78	0.90
Most Probable	0.91	0.82	0.77	0.72	0.90
All Player 1	0.00	0.18	0.45	0.72	0.90
All Player 2	0.08	0.20	0.39	0.58	0.70
All Player 3	0.13	0.20	0.32	0.43	0.50
All Player 4	0.18	0.20	0.24	0.28	0.30
All Player 5	0.72	0.59	0.41	0.22	0.10

From Table 3 it is visible that, for a risk value of 1, which means that only the economic component is considered by the utility function, the *Optimistic* decision method is able to achieve the maximum utility value of 1. This occurs because the *Optimistic* method selects the action that presents the best possible outcome from all scenarios, thus the selected action-scenario combination is the one that achieves the maximum possible income. From Equation (3), the maximum income corresponds to a value of 1 in the economic component of the utility function evaluation. On the other hand, negotiating

the total amount of 10 MW with Player 1 results in an utility value of 0, since this is the worst possible action that can be performed; as can be seen by Figure 5, the expected negotiation price for Player 1 for the amount of 10 MW corresponds to the lowest possible price among all that are estimated. In fact, negotiating the total 10 MW with a single player, considering only the economic component, always leads to very bad incomes, with the exception of Player 5. This player (as presented in Figure 5) presents a relatively good expected price for the amount of 10 MW, hence resulting a fairly good utility value for a risk value of 1.

Looking at the other extreme, considering a risk value of 0 means that only the reputation component is considered by the utility function. When negotiating the total amount with a single player, by Equation (2), the reputation component is equal to the reputation of that player, therefore the utility value when negotiating the total amount with each player, for a risk value of 0, is equal to the reputation value presented in Table 2. All three decision methods have reached a utility value equal to the negotiation of the total amount with player 1. This occurs due to the exclusion of the economic component when risk is equal to 0, which means that the chosen scenario is irrelevant in this case, and only the choice of the opponent player(s), considering their reputation, is important. For this reason, regardless of the chosen scenario, the best action is always to negotiate the full amount with the player that presents the higher reputation, in this case Player 1.

For all intermediate risk values, the same tendency is always observed: the higher utility value is achieved by the *Optimistic* method, followed by the *Most probable* method, and with the *Pessimistic* method in third place. The negotiation of the whole amount with each player individually results in lower utility function values, for all players and for all intermediate risk propensity values.

Figure 6 presents the allocation results (amount to be negotiated with each competitor player) of the *Most probable* decision method, for the 5 considered risk propensity values.
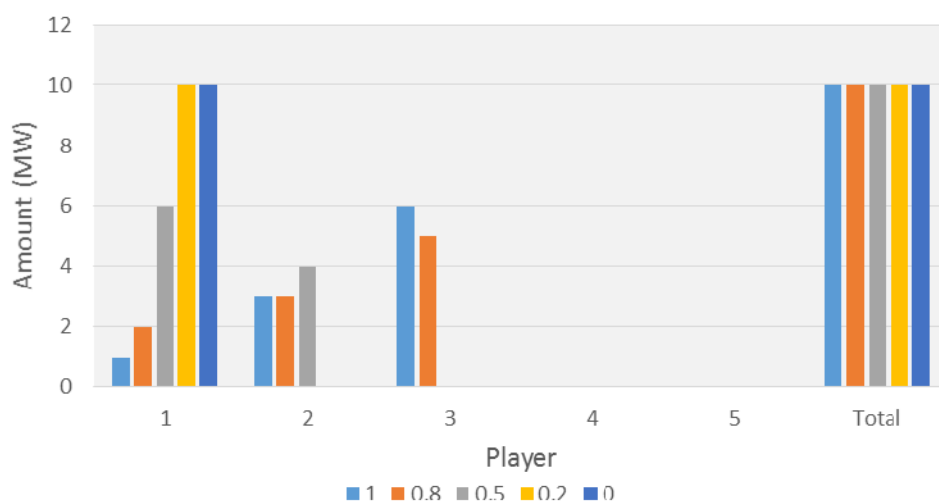

Figure 6. Actions resulting from the *Most probable* decision method for different risks.

By matching the results presented in Figure 6 to the estimation results of the *Most probable* scenario, depicted in Figure 5, it is possible to interpret the reasons why these actions have been chosen by the *Most probable* decision method. Considering a risk value of 1, where only the economic component is assessed, from the 10 MW, 6 MW are allocated to the negotiation with Player 3, 2 MW to Player 2, and 1 MW to Player 1. From Figure 5 one can see that these amounts correspond to the peak expected prices

from these 3 players, thus representing the higher possible economic gain considering the expected prices in the scope of this scenario (the most probable scenario). With a risk value of 0.8, the reputation component is already taken into account, even if just slightly. This is, however, enough to make the amount allocated to Player 3 decrease by 1 MW, which is now allocated to Player 1, which presents a much higher reputation value, and whose expected price for the negotiation of 2 MW is just slightly lower than the expected price for the negotiation of 1 MW. The gain in the reputation component compensates the slight decrease in the economic component. A risk value of 0.5 represents an equal consideration of both the economic and the reputation component. In this case, the negotiation with Player 3 no longer pays off due to the relatively poor reputation value of this player, and the 10 MW are negotiated between Player 1 and Player 2, which are the players with the higher reputation, and whose expected prices for the negotiation of respectively, 6 MW and 4 MW are still very reasonable. The cases where the risk value is set to 0.2 and to 0, which represent a prominence of the reputation component, result in the allocation of the total 10 MW to the negotiation with Player 1—the player with the higher reputation.

Figure 7 presents the execution time of the proposed methodology for different amounts of negotiation power, and for different numbers of potential competitor players.
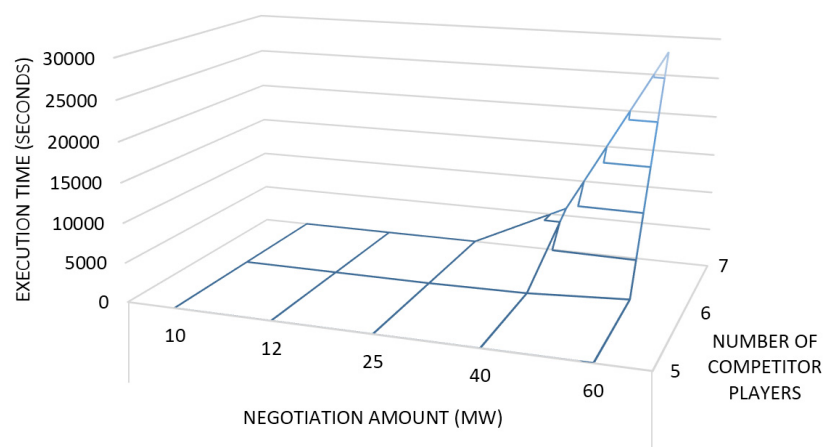

Figure 7. Execution time of the proposed methodology.

From Figure 7 it is visible that the execution of the proposed methodology for a relatively small amount of power (up to 25 MW) and for a reduced number of players takes only a few seconds to run. However, when both the negotiation amount and the number of alternative players increase, the required execution time also increases significantly. This increase is verified to the enlarged number of alternative actions that result from the negotiation of a large amount of power among a large number of players, which require an extensive number of alternative action-scenario evaluations. Moreover, the increase of the number of players also has a large implication of the time required to perform the forecast and estimation process for each player, for each distinct amount.

Table 4 presents the utility function results for the application of the proposed methodology to the negotiation of 10 MW among the complete set of 37 considered potential competitor players, for different values of risk propensity, and comparing the different decision methods with the negotiation of the total amount with a single player (the best one for each risk value). The reputation values for all players have been assigned randomly, and all parameterizations are kept equal to those of the previous case.

Table 4. Utility function values for different risk propensity values, considering 37 competitor players.

Decision Method	Risk Propensity				
	1	0.8	0.5	0.2	0
Pessimistic	0.89	0.83	0.80	0.84	0.98
Optimistic	1.00	0.98	0.92	0.93	0.98
Most Probable	0.96	0.92	0.87	0.89	0.98
Best Player	0.82	0.73	0.78	0.84	0.98

Table 4 shows that, as before, for a risk value of 0, all decision method reach the same utility value as the best player, since the total amount of negotiation is allocated to the player with the higher reputation (in this case the player with the higher reputation has a value of 0.98). For a risk value of 1, the *Optimistic* method is once again able to reach a utility function value of 1, by allocating the negotiation of the required 10 MW among the players that present the best possible economic gain in the scenario that presents the most advantageous price expectations. The negotiation of the total amount with a single player (the best player for each value of risk) always results in smaller utility function values than all three decision methods of the proposed methodology. From the three decision methods, the *Optimistic* is always the one that achieves the best expected outcomes, the *Pessimistic* is the method that presents the worst expected outcomes from the three (although assuming the safer option is still above the negotiation with a single player), and the *Most probable* decision method reaches intermediary utility function results, while representing the best action to take in the *Most probable* case scenario.

5. Conclusions

This paper proposed a methodology to provide decision support to electricity market negotiating players when participating in bilateral contract negotiations. The proposed methodology allows the supported player to decide the amounts of power that should be negotiated with its competitors in order to optimize its expected outcomes. For this, a game theoretic decision method is used and refined with a reinforcement learning approach that enables identifying the *Most probable* scenarios to occur in each context. The decision methods make use of a utility function that considers the reputation of the competitors in addition to the expected economic gain, to evaluate the outcomes of each combination action-scenario. Alternative scenarios are built from the expected price forecasts, using the competitor players' historic contract settlements, complemented by the estimative of missing expected price values.

Results show that the proposed methodology is able to achieve higher utility function values, using any of the three decision methods, than the utility that results from negotiating the total amount of desired price with any single player. As previously identified in the introductory section, the pre-negotiation stage of negotiations as been largely neglected, as well as the proper use of historic data to model opponents' behaviors and act accordingly, which is recognized by the authors themselves [24]. For this reason, there are no significant pieces of work to which the proposed methodology can be compared; therefore the only way to assess its advantage is by comparing the results of the proposed methodology with the expected outcome of negotiating the total amount with a single (the best) competitor player rather than distributing the negotiation amount among the possible opponents taking into account their expected performance. The outcomes of all three decision methods are always superior to those of negotiating the total amount with a single player, even for the *Pessimistic* decision method, which considers only the

best actions to perform under the worst possible negotiation scenario that may occur. Thus, in this case, even if everything goes as bad as possible, the proposed method is always able to provide a suggested action that is most advantageous for the supported player then the negotiation with a single player.

Regarding the execution time of the proposed method, the conclusion to be taken is that the proposed methodology is perfectly suitable to deal with a moderate negotiation amount of power and number of competitor players; but its application to larger problems is only possible as planning decision support (which is its main objective) to be taken some hours before the actual negotiations, or even during the previous days. The execution time can also be adapted by changing the 2E balance requirements of the AiD-EM decision support system, resulting in the use of less forecasting methods (using only the faster ones), and thus considering a reduced number of alternative scenarios. The result is the possible decrease of the achieved utility values, as result of the limited alternatives, but an important increase in execution time performance.

As future work, the improvement of the contract price forecasting process is an important aspect, namely by considering different feature selection approaches in order to test and compare their influence on the forecasting results. Additionally, the experimentation of different alternatives for the utility function definition is also suggested, by considering the different utility functions for loss and gain domains proposed by the prospect theory. In specific, the problem of drawdowns and drawups in electricity prices should be addressed.

Acknowledgments

Author Contributions

Tiago Pinto, Zita Vale and Isabel Praça conceived and designed the computational models; Tiago Pinto and Isabel Praça conceived and designed the experiments; Tiago Pinto performed the experiments; E. J. Solteiro Pires and Fernando Lopes analyzed the data; Zita Vale contributed with expertise in AI methods for power systems; Isabel Praça contributed with expertise in MAS simulation; E. J. Solteiro Pires contributed with expertise in optimization problems; Fernando Lopes contributed with expertise in automated negotiation; Tiago Pinto wrote the paper; Zita Vale, Isabel Praça, E. J. Solteiro Pires and Fernando Lopes revised and improved the paper.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Sharma, K.; Bhakar, R.; Tiwari, H. Strategic bidding for wind power producers in electricity markets. *Energy Convers. Manag.* **2014**, *86*, 259–267.

2. Meeus, L.; Purchalaa, K.; Belmans, R. Development of the internal electricity market in Europe. *Electr. J.* **2005**, *18*, 25–35.

3. Xavier, G.A.; Filho, D.O.; Martins, J.H.; Monteiro, P.M.B.; Diniz, A.S.A.C. Simulation of distributed generation with photovoltaic microgrids—Case study in Brazil. *Energies* **2015**, *8*, 4003–4023.

4. European Market Coupling Company Homepage. Available online: http://www.marketcoupling. com/ (accessed on 16 May 2015).

5. EUPHEMIA: Description and Functioning. Available online: http://www.apxgroup.com/wp-content/ uploads/Euphemia-Public-Documentation-to-be-published.pdf (accessed on 16 May 2015).

6. California Independent System Operator Homepage. Available online: http://www.caiso.com (accessed on 16 May 2015).

7. MISO Energy Homepage. Available online: http://www.misoenergy.org (accessed on 1 April 2014).

8. Operador Nacional do Sistema Elétrico, Electrical System Nacional Operator Homepage. Available online: http://www.ons.org.br (accessed on 16 May 2015).

9. Algarvio, H.; Lopes, F.; Santana, J. Bilateral Contracting in Multi-agent Energy Markets: Forward Contracts and Risk Management. In *Highlights of Practical Applications of Agents, Multi-Agent Systems, and Sustainability-The PAAMS Collection*; Springer International Publishing: Basel, Switzerland, 2015; pp. 260–269.

10. Bai, H.; Miao, S.; Ran, X.; Ye, C. Optimal dispatch strategy of a virtual power plant containing battery switch stations in a unified electricity market. *Energies* **2015**, *8*, 2268–2289.

11. Cerjan, M.; Matijaš, M.; Delimar, M. Dynamic hybrid model for short-term electricity price forecasting. *Energies* **2014**, *7*, 3304–3318.

12. Yoo, T.H.; Park, H.; Lyu, J.-K.; Park, J.-K. Determining the interruptible load with strategic behavior in a competitive electricity market. *Energies* **2015**, *8*, 257–277.

13. Praça, I.; Ramos, C.; Vale, Z.; Cordeiro, M. MASCEM: A multi-agent system that simulates competitive electricity markets. *IEEE Intell. Syst.* **2003**, *18*, 54–60.

14. Li, H.; Tesfatsion, L. Development of open source software for power market research: The AMES test bed. *J. Energ. Markets* **2009**, *2*, 111–128.

15. Koritarov, V.S. Real-world market representation with agents. *IEEE Power Energ. Mag.* **2004**, *2*, 39–46.

16. Cincotti, S.; Gallo, G. Genoa artificial power-exchange. *Agents Artif. Intell.* **2013**, *6*, 348–363.

17. Pinto, T.; Vale, Z.; Sousa, T.M.; Praça, I.; Santos, G.; Morais, H. Adaptive learning in agents behaviour: A framework for electricity markets simulation. *Integr. Comput. Aided Eng.* **2014**, *21*, 399–415.

18. Von Neumann, J.; Morgenstern, O. *Theory of Games and Economic Behavior*; Princeton University Press: Princeton, NJ, USA, 1947.

19. Pruitt, D.; Kim, S. *Social Conflict: Escalation, Stalemate, and Settlement*, 2nd ed.; McGraw Hill: New York, NY, USA, 2004.

20. Thompson, L. *The Mind and Heart of the Negotiator*, Prentice Hall: Englewood Cliffs, NJ, USA, 2005.

21. Snyder, G.; Diesing, P. *Conflict among Nations*; Princeton University Press: Princeton, NJ, USA, 1977.

22. Jennings, N.; Faratin, P.; Lomuscio, A.; Parsons, S.; Wooldridge, M.; Sierra, C. Automated negotiation: Prospects, methods and challenges. *Group Decis. Negotiat.* **2001**, *10*, 199–215.

23. Rahwan, I.; Ramchurn, S.; Jennings, N.; McBurney, P.; Parsons, S.; Sonenberg, L. Argumentation-based negotiation. *Knowl. Eng. Rev.* **2004**, *18*, 343–375.

24. Lopes, F.; Wooldridge, M.; Novais, A.Q. Negotiation among autonomous computational agents: Principles, analysis and challenges. *Artif. Intell. Rev.* **2008**, *29*, 1–44.

25. Pruitt, D.; Carnevale, P. *Negotiation in Social Conflict*; Open University Press: Philadelphia, PA, USA, 1993.

26. Savage, G.; Blair, J.; Sorenson, R. Consider both relationships and substance when negotiating strategically. *Acad. Manag. Persperct.* **1989**, *3*, 37–48.

27. Raiffa, H. *The Art and Science of Negotiation*; Harvard College Press: Cambridge, MA, USA, 1982.

28. Lewicki, R.; Barry, B.; Saunders, D.; Minton, J. *Negotiation*; McGraw Hill: New York, NY, USA, 2003.

29. Nguyen, T.; Jennings, N. Managing commitments in multiple concurrent negotiations. *Electron. Commerce Res. Appl.* **2005**, *4*, 362–376.

30. Li, C.; Giampapa, J.; Sycara, K. Bilateral negotiation decisions with uncertain dynamic outside options. *IEEE Trans. Syst. Man Cybern.* **2006**, *36*, 31–44.

31. Pinto, T.; Sousa, T.M.; Praça, I.; Vale, Z.; Morais, H. Support vector machines for decision support in electricity markets' strategic bidding. *Neurocomputing* **2015**, in press.

32. Pinto, T.; Sousa, T.M.; Vale, Z. Dynamic artificial neural network for electricity market prices forecast. In Proceedings of the IEEE 16th International Conference on Intelligent Engineering Systems (INES), Lisbon, Portugal, 13–15 June 2012; pp. 311–316.

33. Sabater, J.; Sierra, C. Review on computational trust and reputation models. *Artif. Intell. Rev.* **2005**, *24*, 33–60.

34. Gatti, N. Game Theoretic Models for Strategic Bargaining. In *Negotiation and Argumentation in Multi-Agent Systems: Fundamentals, Theories, Systems and Applications*; Lopes, F., Coelho, H., Eds.; Bentham Science Publishers: Sharjah, UAE, 2014; pp. 48–81.

35. Nash, J.F., Jr. The bargaining problem. *J. Econometric Soc.* **1950**, *18*, 155–162.

36. Nash, J.F., Jr. Two-person cooperative games. *J. Econometric Soc.* **1953**, *21*, 128–140.

37. Herrero, M. N-player bargaining and involuntary underemployment. Ph.D. Thesis, London School of Economics, London, UK, 1985.

38. Krishna, V.; Serrano, R. Multilateral bargaining. *Rev. Econ. Studies* **1996**, *63*, 61–80.

39. Roth, A.E. Bargaining experiments. In *Handbook of Experimental Economics*; Kagel, J., Roth, A.E., Eds.; Princeton University Press: Princeton, NJ, USA, 1995; pp. 253–348.

40. Gao, B.; Zhang, W.; Tang, Y.; Hu, M.; Zhu, M.; Zhan, H. Game-theoretic energy management for residential users with dischargeable plug-in electric vehicles. *Energies* **2014**, *7*, 7499–7518.

41. Li, R.; Ma, H.; Wang, F.; Wang, Y.; Liu, Y.; Li, Z.; Li, R.; Ma, H.; Wang, F.; Wang, Y.; *et al*. Game optimization theory and application in distribution system expansion planning, including distributed generation. *Energies* **2013**, *6*, 1101–1124.

42. Pinto, T.; Praça, I.; Vale, Z.; Morais, H.; Sousa, T.M. Strategic bidding in electricity markets: An agent-based simulator with game theory for scenario analysis. *Integr. Comput. Aided Eng.* **2013**, *20*, 335–346.

43. Argoneto, P.; Bruccoleri, M.; Lo Nigro, G.; Perrone, G.; Noto La Diega, S.; Renna, P.; Sudhoff, W. High level planning of reconfigurable enterprises: A game theoretic approach. *CIRP Ann. Manuf. Technol.* **2006**, *55*, 509–512.

44. Kaihara, T.; Fujii, S.; Iwata, K. Virtual enterprise coalition strategy with game theoretic multi-agent paradigm. *CIRP Ann. Manuf. Technol.* **2006**, *55*, 513–516.

45. Chen, S.; Weiss, G. An approach to complex agent-based negotiations via effectively modeling unknown opponents. *Expert Syst. Appl.* **2015**, *42*, 2287–2304.

46. Li, M.; Bao Vo, Q.; Kowalczyk, R.; Ossowski, S.; Kersten, G. Automated negotiation in open and distributed environments. *Expert Syst. Appl.* **2013**, *40*, 6195–6212.

47. Ji, P.; Ma, X.; Li, G. Developing green purchasing relationships for the manufacturing industry: An evolutionary game theory perspective. *Int. J. Prod. Econ.* **2015**, *166*, 155–162.

48. Sabater, J.; Sierra, C. REGRET: A reputation model for gregarious societies. Available online: http://www.emse.fr/~boissier/enseignement/sma05/exposes/sabater00regret.pdf (accessed on 10 June 2015).

49. Jain, A.K. Data clustering: 50 years beyond K-Means. *Pattern Recognit. Lett.* **2010**, *31*, 651–666.

50. Pinto, T.; Vale, Z.; Sousa, T.M.; Praça, I. Negotiation context analysis in electricity markets. *Energy* **2015**, *85*, 78–93.

51. Faia, R.; Pinto, T.; Vale, Z. Dynamic fuzzy estimation of contracts historic information using an automatic clustering methodology. In *Highlights of Practical Applications of Agents, Multi-Agent Systems, and Sustainability-The PAAMS Collection*; Bajo, J., Sáchez-Pi, N., Hallenborg, K., Méndez, N.D.D., Pawlewski, P., Lopes, F., Botti, V., Julian, V., Eds.; Springer International Publishing: Basel, Switzerland, 2015.

52. Rahimi-Kian, A.; Sadeghi, B.; Thomas, R.J. Q-learning based supplier-agents for electricity markets. *IEEE Power Eng. Soc. Gen. Meeting* **2005**, *1*, 420–427.

53. Juang, C.; Lu, C. Ant colony optimization incorporated with fuzzy Q-learning for reinforcement fuzzy control. *IEEE Trans. Syst. Man Cybern. Part A Syst. Humans* **2009**, *39*, 597–608.

54. MIBEL Data Files. Available online: http://www.omie.es/aplicaciones/datosftp/datosftp.jsp?path=/ (accessed on 16 May 2015).