



# Dynamic Fuzzy Clustering Method for Decision Support in Electricity Markets Negotiation

Ricardo Faia<sup>a</sup>, Tiago Pinto<sup>a</sup> and Zita Vale<sup>a</sup>

<sup>a</sup> GECAD – Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development – Polytechnic of Porto (ISEP/IPP), R. Dr. António Bernardino de Almeida, 431, P-4249-015 Porto, Portugal  
{rfmfa, tmcfp, zav}@isep.ipp.pt

## KEYWORD

*Artificial intelligence; clustering; electricity markets; fuzzy logic*

## ABSTRACT

*Artificial Intelligence (AI) methods contribute to the construction of systems where there is a need to automate the tasks. They are typically used for problems that have a large response time, or when a mathematical method cannot be used to solve the problem. However, the application of AI brings an added complexity to the development of such applications. AI has been frequently applied in the power systems field, namely in Electricity Markets (EM). In this area, AI applications are essentially used to forecast/estimate the prices of electricity or to search for the best opportunity to sell the product. This paper proposes a clustering methodology that is combined with fuzzy logic in order to perform the estimation of EM prices. The proposed method is based on the application of a clustering methodology that groups historic energy contracts according to their prices' similarity. The optimal number of groups is automatically calculated taking into account the preference for the balance between the estimation error and the number of groups. The centroids of each cluster are used to define a dynamic fuzzy variable that approximates the tendency of contracts' history. The resulting fuzzy variable allows estimating expected prices for contracts instantaneously and approximating missing values in the historic contracts.*

## 1. Introduction

Experiments with Artificial Intelligence (AI) began in the 50s, with Warren McCulloch and Walter Pitts, who created a model of artificial neurons (McCulloch and Pitts, 1943). The first steps of the AI were successful, but had several limitations (e.g. technology and remote programming languages). Computers were seen as machines that only performed mathematical operations, and the fact of having a machine that could think, caused admiration in people. AI tries to offer machines some of the aspects that are intrinsic to human nature. In fact, AI brings to many the image of a future where robots do all the things that a human being is capable of, with the ability to think and analyze, hence be self-sufficient. However,



in reality it became something more present in people's lives than this wishful thinking. This view, that AI is linked to the future is not true, it is present in various segments of society, whether at home, at work, on the street and most of the time goes unnoticed (Alvarado-Pérez et al., 2015; Chamoso and De La Prieta, 2015).

In addition to being applied in many branches of societal activities, AI is also applied in the world of electric energy markets, where it is often used to make forecasts of the values to support the market the user's decision (Pinto et al., 2015). AI tools are indispensable in the operation of electrical energy markets, as they offer support for the user. This need has been growing with the current changes in the electrical sector, such as the power sector restructuring. Competitiveness has been bringing with it new challenges with which the users have to deal (Meeus et al., 2005; Shahidepour et al., 2002).

The development of simulation platforms based in Multi-Agent Systems (MAS), which are integral parts of AI, is increasing as a good option to simulate real systems in which stakeholders have different and often conflicting objectives. These systems allow simulating scenarios and strategies, providing users with decision support according to their profile of activity (Li and Tesfatsion, 2009; Santos et al., 2016). MASCEM - Multi-Agent Simulator for Electricity Markets (Praça et al., 2003) is a simulation tool to study and explore restructured electricity markets. Its purpose is to simulate as many market models and player types as possible. The learning process of its agents is undertaken using MASCEM's connection with AiD-EM (Adaptive Decision Support for Electricity Markets Negotiations) (Pinto et al., 2014). AiD-EM provides decision support to electricity market players, allowing them to automatically adapt their strategic behavior to different contexts of negotiation. Additionally, AiD-EM is equipped with a portfolio optimization methodology, which enables players to decide the participation investment that should be made in each type of market, in order to optimize the potential profits from selling their power, or minimize the costs of buying the required amounts (Faia et al., 2015b; Pinto et al., 2015). The portfolio optimization considers the forecasted market prices that are expected in each alternative market that the supported player is allowed to participate in (e.g. day-ahead spot market, each session of the intraday market, bilateral contracts, forward markets).

Some of the advantages of using these AI tools are their responsiveness in execution time, and also to surpass the lack of information (Ferreira et al., 2015). Hence, the application of AI is essential for the development of this work. This paper proposes a methodology that estimates expected bilateral contract prices by analyzing the historical log of established contracts. Each bilateral contract is characterized by a specific amount of power and an associated price. However, predictions of expected prices for different amounts of power than those contained in the historic log are often required. Hence, an adequate estimative is essential. Moreover, so that a suitable portfolio optimization can be executed, the expected return prices for each possible amount of power are required. This, however, is impracticable due to the number of possible amounts (which tends to infinite when increasing the number of decimals of the power amount value). In order to overcome this difficulty fuzzy functions are used to estimate the price of each amount of energy given the historical information. Dynamic fuzzy membership functions are defined accordingly to the output of a clustering process, which groups the historical contracts depending on their similarity. The optimal number of clusters, which defines the number of fuzzy intervals, is calculated automatically, taking into account the preference for the balance between the estimation error and the number of fuzzy intervals. The proposed methodology allows estimating the large number of historic contract prices by means of a single fuzzy variable, hence reducing drastically the amount of variables to be considered by the portfolio optimization algorithm, and consequently contributing to the decrease of the optimization time. Additionally, the adaptation of the number of fuzzy intervals allows

considering more or less fuzzy functions in the problem to adapt the balance of execution time and quality of the estimation to each execution requirements.

After this introductory section, section 2 presents the proposed methodology, including the clustering process, the application of fuzzy logic, and the dynamic fuzzy clustering methodology. Section 3 presents some experimental findings that enable the assessment of the proposed methodology, and finally, section 4 presents the most relevant conclusions and contributions of this work.

## 2. Proposed methodology

This paper proposes methodology in which fuzzy logic (Zadeh, 1965) is used to estimate the expected electricity prices in certain electricity markets with specific characteristics. Relevant examples are bilateral contracts and forward markets, in which the price of electricity varies depending on the traded amount. Dynamic fuzzy membership functions are defined accordingly to the output of the clustering process (Jain, 2010), which groups the historical contracts depending on their similarity. The optimal number of clusters, which defines the number of fuzzy intervals, is calculated automatically, taking into account the Efficiency/Effectiveness (2E) preference, e.g. the balance between the estimation error and the number of fuzzy intervals (Faia et al., 2015a; Pinto et al., 2014). The higher preference for the optimization of the estimation error results in a fuzzy variable with an extended number of intervals, which causes an increase in execution time when using the fuzzy variable for the estimation for decision support purposes. If the user has the necessity for a quick response, it is proposed that it gives an indication of preference for a smaller number of clusters, although the error value will be higher. As shown by Figure 1, which describes the steps that are performed to obtain the final result.

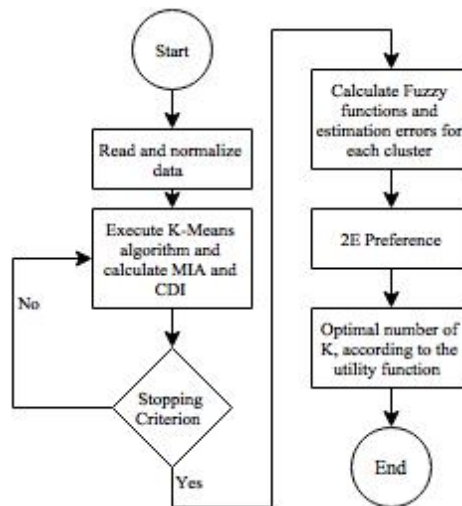


Figure 1. Flowchart of the proposed methodology

Initially, data is loaded and normalized so that the K-Means algorithm (Jain, 2010) can be used make the grouping of clusters. The K-Means algorithm requires that the exact number of clusters is defined, so

that they can be formed. In this case, this number will not be imposed; rather stopping criterion is defined, which obligates the algorithm to repeat its process until it verified. In each cycle, the clustering process is repeated, adding one more target group in each cycle. After the groups are defined, fuzzy functions are built to estimate the error for the group of numbers. After this phase, the user's preference for the balance between the execution time and the estimation quality is used. In each case, the user may give more importance to the number of clusters or incurred error in the estimation of electricity prices, by varying the weight between these two parameters. Once the weights are defined, a utility function is applied to select the optimal number of clusters. After the ideal number of clusters is found, the corresponding fuzzy functions are recovered.

## 2.1. Clustering methodology

The clustering technique consists in grouping objects with similar characteristics. This assembly causes the formation of groups in order to obtain aggregates with similar characteristics and maximize the similarity, i.e. trying to form a cluster with the highest possible homogeneity. On the contrary, one tries to minimize the similarity between the data of different groups. This technique is useful as it allows the user to draw knowledge from a data set. Many algorithms can be used to cluster analysis, however, the most popular is the K-means algorithm (Jain, 2010).

The K-means algorithm was implemented for the first time by James MacQueen in 1967 (MacQueen, 1967). This algorithm, as already mentioned, is used to partition the clustering, and it is necessary to provide information on how many groups a user will want to match. The algorithm ends when the combinations of objects and groups minimize specific criteria and after it converges to present a solution. Typically, the squared error (equation (1)) is used as stopping criterion, that is, accept the combination of objects and groups that minimize the sum of squared error between the center of the cluster and the belonging objects.

$$\min \sum_{i=1}^k \sum_{x \in C_i} ||x - \mu_i||^2 \quad (1)$$

This methodology considers a set of observations  $(x_1, x_2, \dots, x_n)$ , where each observation is a dimensional real vector, and  $n$  is the number of considered observations,  $C_i$  is a cluster and  $\mu_i$  is the *centroid* of  $C_i$ .

## 2.2. Fuzzy Logic methodology

Fuzzy logic works with sets of values to create response patterns. The fuzzy logic works with valued logics, which unlike the bivalent logic with only two values to represent a response that is true or false, the fuzzy logic has intermediate terms to give an answer. Intermediate terms do not say whether they are true or false, they only represent the truth as a value (Zadeh, 1965). Fuzzy logic has applications for decision support systems that examine various information, where knowledge of a given situation is passed to the system and so when an input value is passed to the application it will return some conclusions on that kind of input. In this case, fuzzy logic is used in order to support the estimation of prices for electricity. In this sense, so that the system can give an answer, it is necessary to incorporate

past situations that provide the historical logs. Figure 2, shows an example of the application of fuzzy logic to the addressed case.

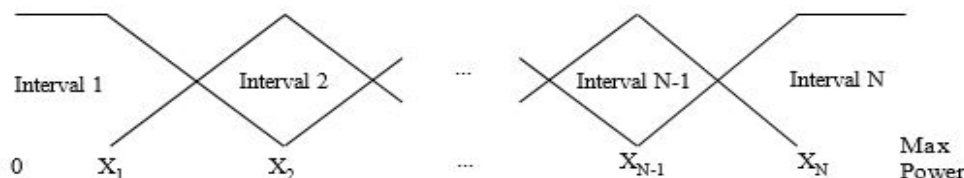


Figure 2. Fuzzy variables

From Figure 2 it is possible to observe the operation mode of fuzzy logic. It is necessary to establish ranges and limits for the fuzzy membership functions, and the maximum limit of the fuzzy set. The limits of the ranges are defined by the centroids obtained from the grouping of the data generated by the technical cluster. E.g.  $X_1$  to  $X_N$  of Figure 1 are the power amounts defined in each of the  $k$  clusters' centroids, hence  $N=k$ . The limits of each function assume the value of the preceding and following cluster centroids, which assume membership values of zero. All membership functions are triangular, except from the first and last.

The application of fuzzy logic becomes useful in electricity markets, because there are markets where the price of electricity varies with the amount to be negotiated. With a good grouping of data it is possible to obtain very similar values that estimate the real ones effectively.

In this sense fuzzy logic is used for smoothing the transition interval values. E.g. When negotiating 50 MW in the certain market (part of one power interval) the expected price is  $X$ ; When negotiating 51 MW in the same market, 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.

### 2.3. Dynamic Fuzzy Clustering

In this section an explanation is given on what has been created and used, namely the Mean Index Adequacy (MIA) and Cluster Dispersion Indicator (CDI) indices of clustering evaluation (Mahmoudi-Kohan et al., 2009), the utility function created for the problem, the used error description, and the used stopping criterion used (Dent et al., 2014).

MIA uses the Euclidian distance method to determine the value that reflects the quality of a cluster partition. MIA gives a value which relies on the amount by which each is compact, if the members in the cluster are close together the MIA value is low. This index is calculated as in (2).

$$MIA = \sqrt{\frac{1}{K} \times \sum_{k=1}^K d^2(x^{(k)}, \mu^{(k)})} \tag{2}$$

CDI equation (3) also depends on the distance between the members of the same cluster, but also incorporates the evaluation of the distances between the centroids for different clusters. Hence, this index assesses both the compactness of the clusters and the amount by which each cluster differs from the others.

$$CDI = \frac{\sqrt{\frac{1}{K} \sum_{k=1}^K \left[ \frac{1}{2 \cdot n^{(k)}} \sum_{m=1}^{n^k} \hat{d}^2(x^m, \mu^k) \right]}}{\sqrt{\frac{1}{2 \cdot K} \sum_{k=1}^K d^2(x^{(K)}, R)}} \quad (3)$$

After calculating the MIA and CDI rates, the stopping criterion is then applied. It is obtained in view of the values recorded by the indicators.

$$m1_i = \left| \frac{MIA_i - MIA_{i-1}}{K_i - K_{i-1}} \right| \quad (4)$$

$$m2_i = \left| \frac{CIA_i - CIA_{i-1}}{K_i - K_{i-1}} \right|$$

$$\begin{aligned} m1_{i-1} - m1_i &< 0 \\ \vee \\ m2_{i-1} - m2_i &< 0 \end{aligned} \quad (5)$$

Equations (4) and (5) were created incorporating the criteria of the stopping methodology. When one of the conditions of equation (5) is met, the algorithm returns a set of clusters that are close to the number of clusters where it stopped. E.g. if it stops at 7, the following set of values will be returned: [4, 5, 6, 7, 8, 9, 10], or three number following and prior the number at which the condition defined to stop. As can be seen by the equations, the stopping criterion is related to the calculation of the MIA and CDI.

After the likely range of values for the optimal number of clusters is found, it is necessary to choose which of these is the optimal. For this, equation (6) is used.

$$K_{best} = \min(wc \times K_{norm} + we \times MADP_{norm}) \quad (6)$$

The  $K_{best}$  equation (6) is used to specify the utility, by considering the minimization of both variables, where  $K_{norm}$  and  $MADP_{norm}$  represent the normalized  $k$  and the error value, respectively; and  $wc$  and  $we$  are weights for the preference between  $k$  and the error value, according to the specification of the 2E balance. The value of both weights ranges from 0 to 1.

The corresponding Mean Absolute Deviation Percent (MADP) measurement error is calculated for each value of the clear cut range. For this it is necessary to calculate the fuzzy functions and compare them with the real recorded values. Equation (7) expresses the calculation of MADP.

$$MADP = \frac{\sum_{x=1}^N \|F_x - A_x\|}{\sum_{x=1}^N \|A_x\|} \quad (7)$$

where  $F_x$  represents the fuzzy estimated value, and  $A_x$  the real value.

As previously described, to apply the utility function it is necessary to normalize the values of both parameters, equation (8) is used for this purpose.

$$x_{i \text{ norm}} = \frac{x_i - \min(x)}{\max(x) - \min(x)} \quad (8)$$

The choice of the minimum number of  $k$  that best reduces the error is a multi-objective problem, where there are two goals that are divergent, e.g. when trying to improve one, the other worsens (the estimation error decreases when  $k$  increases, and vice-versa). By choosing the cluster number by using the proposed mechanism, not all of the solutions found in the search space that will be explored can be chosen. This is

because the result of the mechanism is one integer number (number of clusters), and in the set of integer solutions, not all are dominant in the two objectives (Hwang and Masud, 1979).

### 3. Results

This section exposes the results obtained by applying the proposed methodology. This case study counts on 50 observations between 1MW and 50MW, with different associated prices. The objective is to use the proposed methodology to estimate the price of different amounts of electric power.

Initially it is shown by Figure 3 the evolution of  $m1$  and  $m2$  (clustering error indices), for a number of clusters that varies from 1 to 10.

As it is possible to observe by Figure 3, the trend of the indicators is to decrease with an increasing number of clusters. However, as can be seen in Figure 3, these indicators present a tendency to decrease the error value at different times. The identification of this change is used to trigger the stop criterion, i.e. the use of the equation (5). In this case the stopping criterion was triggered when the number of clusters is 6, defining a set of possible values for the ideal number of clusters: [3,4,5,6,7,8,9]. Figure 4 presents the value of the MADP error for each fuzzy function created using each number of clusters (from 1 to 10).

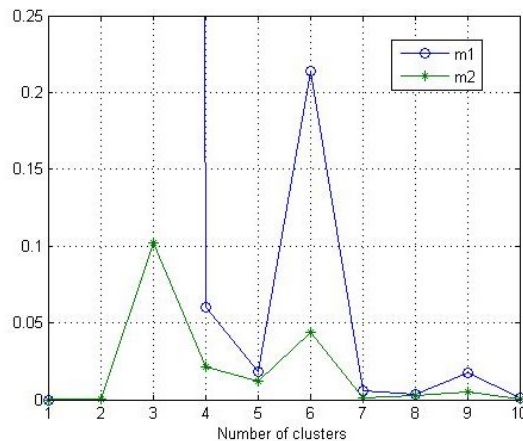


Figure 3. Evolution of the indicator  $m1$  and  $m2$

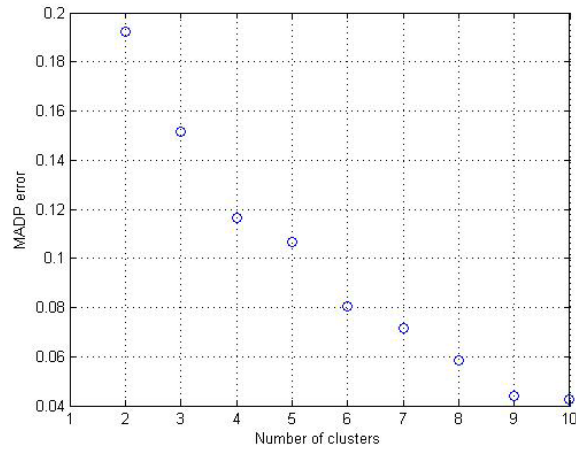


Figure 4. MADP Error evolution

The fuzzy function is defined for each different possible number of clusters of the set of possible values for the ideal number of clusters. The plot of Figure 4 represents the two objectives: the Y-axis represents the value of estimation error (MADP) and the X-axis defines the number of clusters. As mentioned before, this multi-objective problem has two objectives. This means that there is a set of solutions that are suitable to be accepted. The name of set solution is the Pareto front, which has one set of solutions that are discrete and finite (Hu et al., 2013). As has already been explained, with the objective of improving the solution from the standpoint of one objective, the other is deteriorated. This is visible from the graph of Figure 4, where a solution with a small error has a high number of clusters; on the other hand, with a small number of clusters, the error is higher. For this reason, in order to achieve a unique response, a weight value has to be attributed to each of the two objectives, indicating the importance that each has for the final solution.

In Table 1 are expressed the values of the utility function for different combinations of weights for the minimization of the two objectives: number of clusters (*wc*) and estimation error (*we*).

| we  | K (number of clusters) |           |           |           |           |           |     | wc  |
|-----|------------------------|-----------|-----------|-----------|-----------|-----------|-----|-----|
|     | 3                      | 4         | 5         | 6         | 7         | 8         | 9   |     |
| 0,1 | 0,1                    | 0,2171106 | 0,3582992 | 0,4839011 | 0,6254085 | 0,7632176 | 0,9 | 0,9 |
| 0,2 | 0,2                    | 0,2675545 | 0,383265  | 0,4678023 | 0,5841503 | 0,6931018 | 0,8 | 0,8 |
| 0,3 | 0,3                    | 0,3179984 | 0,4082308 | 0,4517034 | 0,5428921 | 0,6229861 | 0,7 | 0,7 |
| 0,4 | 0,4                    | 0,3684423 | 0,4331967 | 0,4356045 | 0,5016339 | 0,5528703 | 0,6 | 0,6 |
| 0,5 | 0,5                    | 0,4188863 | 0,4581625 | 0,4195056 | 0,4603757 | 0,4827545 | 0,5 | 0,5 |
| 0,6 | 0,6                    | 0,4693302 | 0,4831283 | 0,4034068 | 0,4191176 | 0,4126388 | 0,4 | 0,4 |
| 0,7 | 0,7                    | 0,5197741 | 0,5080942 | 0,3873079 | 0,3778594 | 0,342523  | 0,3 | 0,3 |
| 0,8 | 0,8                    | 0,570218  | 0,53306   | 0,371209  | 0,3366012 | 0,2724073 | 0,2 | 0,2 |
| 0,9 | 0,9                    | 0,6206619 | 0,5580258 | 0,3551101 | 0,295343  | 0,2022915 | 0,1 | 0,1 |

Table 1. Utility function values for different weights for number of cluster /estimation error





From Table 1 it is visible that attributing a small value of weight for the error and high for  $k$ , results in achieving a solution with a small  $k$  (where the minimization of the two objectives reaches its smaller value). On the other hand, a high weight for the error minimization objective results in the choice of larger  $k$ . Since the proposed method can only choose an integer value of  $k$  (as it represents an integer number of clusters), the set of solutions is discrete and finite  $k \in [3, \dots, 9]$ .

Table 1 shows that only  $K$  equal to 3, 4 and 9 were chosen as the optimal  $k$  (shaded values). As can be seen in Table 1 nine combinations of weights are considered, but many more could have been selected, depending on the number of decimal places.  $K$  equal to 5, 6, 7 and 8 were not selected as ideal in any of the considered cases. However, these may be selected by assigning weighing intervals of smaller amplitude.

Figure 5 shows three applications of the proposed methodology, varying the number of clusters. These illustrations show the results of Table 1 graphically. In all three figures it is possible to observe the actual data are represented in blue, the estimated data in green and the centers of the clusters in red. Figure 5 a) is a representation with three clusters resulting in four different areas of fuzzy functions. For the three clusters was made the following combination of weights [(we, wc); (0.1,0.9); (0.2,0.8) and (0.3,0.7)]. In Figure 5 b) we have a representation of four clusters that result in five different estimations areas, and can be obtained from the following combination of weights [(0.4,0.6) and (0.5,0.5)]. Figure 5 c) shows a representation with nine clusters and ten different estimation areas prices for the nine clusters. The following combinations of weights has been considered: [(0.6,0.4); (0.7,0.3); (0.8,0.2) and (0.9,0.1)]. As expected, a larger number of clusters results in a smaller estimation error, but on the other hand longer delay the process.

The case with three clusters results in an average MADP of 0.1517. Using four clusters results in an average MADP of 0.1163 and with nine clusters the average MADP is 0.0440, which is about 1/4 of the error when using three clusters.

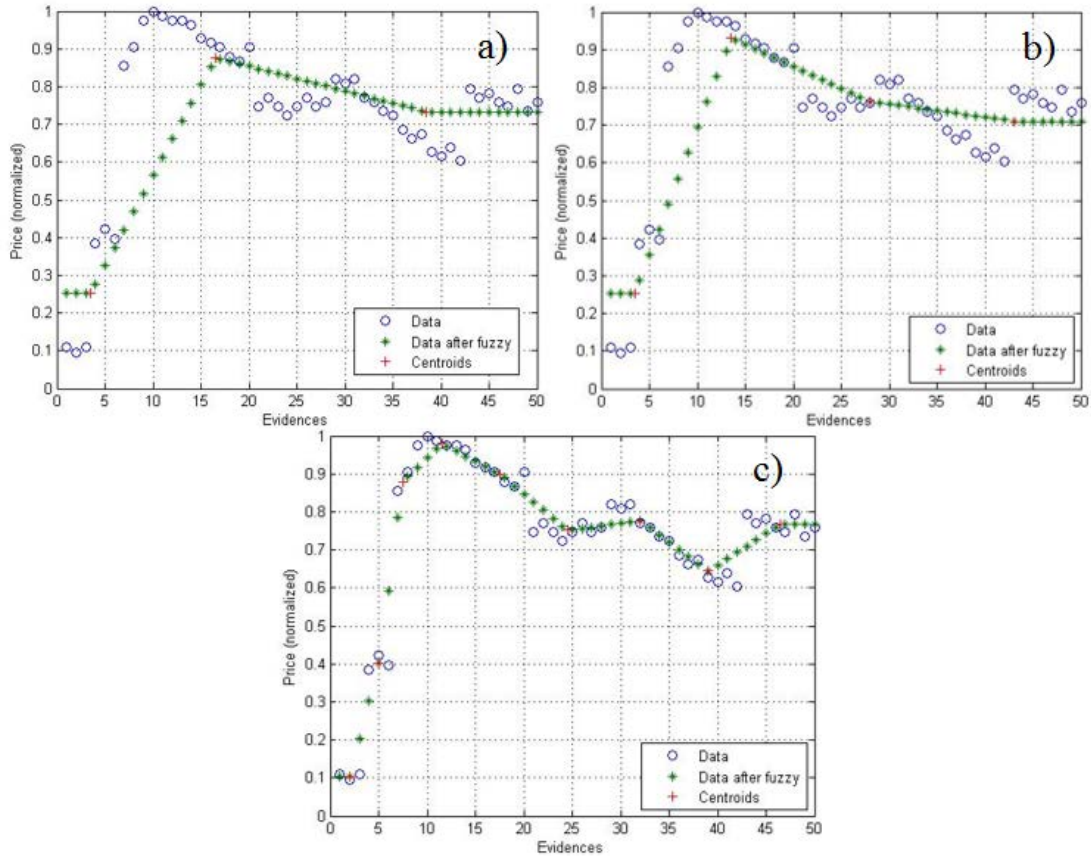


Figure 5. Dynamic fuzzy estimation values for different optimal number of clusters, namely: a)  $K=3$ ; b)  $K=4$ , and c)  $K=9$

An explanatory example of how fuzzy logic is applied is presented as follows. The k-means algorithm not only allows to cluster data into different groups, but also returns the coordinates of each cluster centroid – representative values of each cluster. The explanatory example is generated with the same problem, by building the fuzzy functions related to the case of optimal  $K=9$  – as seen from Figure 5 c. Table 2 presents the coordinates of the clusters centroids, from  $K=1$  to 9, where  $x$  stands for the amount of power (in MW), and  $y$  represents the price (€MWh).

| K        | C1    | C2     | C3  | C4     | C5     | C6     | C7     | C8     | C9    |
|----------|-------|--------|-----|--------|--------|--------|--------|--------|-------|
| x (MW)   | 2     | 5      | 7,5 | 11,5   | 17,5   | 24,5   | 32     | 39     | 46,5  |
| Y (€MWh) | 8,666 | 33,333 | 73  | 81,333 | 74,667 | 62,375 | 64,429 | 53,429 | 63,75 |

Table 2. Coordinates of the clusters centroids



Figure 6 represents the application of fuzzy logic with the data of Table 2, which is achieved by applying the K-means algorithm. In Figure 6, the yellow dots represent the cluster centroids. In this example the value for 6 MW and 11 MW are estimated.

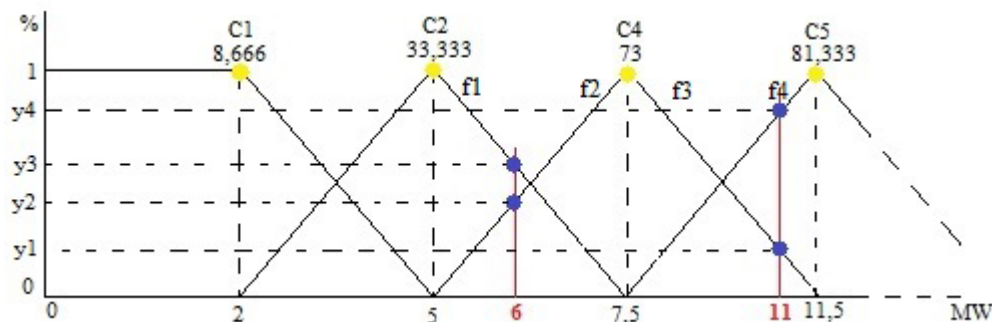


Figure 6. Illustration of fuzzy logic application

As one can see from Figure 6 the values displayed in red are those that are intended to be estimated. Each value intersects two lines, and thus it is necessary to calculate the ordinates (yy) of the intersection points. In the figure, each interval, except from the first, has two straight downwards lines (negative slope) and a rising line (positive slope), so it is necessary to calculate the slopes of each of these lines, as well as the interception. With this, the well-known expression (9) is used for each of the lines. Table 3 presents the values of the slopes for each of the intervals of power considered in Figure 6.

$$f(y) = mx + b \tag{9}$$

| Interval no. | 1 |   | 2        |          | 3    |     | 4     |        | 5        |          |
|--------------|---|---|----------|----------|------|-----|-------|--------|----------|----------|
| Limits       | 0 | 2 | 2        | 5        | 5    | 7,5 | 7,5   | 11,5   | 11,5     | 17,5     |
| <b>m</b>     | 1 |   | -0,33333 | 0,333333 | -0,4 | 0,4 | -0,25 | 0,25   | -0,16667 | 0,166667 |
| <b>b</b>     | 1 |   | 1,666667 | -0,66667 | 3    | -2  | 2,875 | -1,875 | 2,916667 | -1,91667 |

Table 3. Slopes and y axis value at source

The data presented in Table 3 has been obtained from the data presented in Figure 6. With this information expressions (10), (11), (12) and (13) can be written, as follows:

$$f1(y) = -0,4x + 3 \tag{10}$$

$$f2(y) = 0,4x - 2 \tag{11}$$

$$f3(y) = -0,25x + 2,785 \tag{12}$$

$$f4(y) = 0,25x - 1,875 \tag{13}$$

By substituting the value of y related to the power amount for each it is intended to estimate the expected price, and multiplying by the respective centroid, the following values are obtained. By summing the different values, the price estimation presented in Table 4 is achieved:



- $y_1 = (-0,25 \times 11 + 2,785) \times C_4 = 0,125 \times 73 = 9,125$
- $y_2 = (0,4 \times 6 - 2) \times C_4 = 0,4 \times 73 = 29,2$
- $y_3 = (-0,4 \times 6 + 3) \times C_2 = 0,6 \times 33,333 = 20$
- $y_4 = (0,25 \times 11 - 1,875) \times C_5 = 0,875 \times 81,333 = 71,167$

| Amount (MW) | Real Price (€) | Estimated Price (€)  |
|-------------|----------------|----------------------|
| 6           | 33             | $y_2 + y_3 = 49,2$   |
| 11          | 82             | $y_1 + y_4 = 80,292$ |

Table 4. Price estimation

## 4. Conclusion

The problem of choosing the optimal  $k$  in clustering is very popular, and a consensual approach has not yet been discovered. This paper proposes a methodology to automatically achieve the optimal  $k$ , depending on the objective preference of the user. In this case, the optimal  $k$  is used to estimate bilateral contract prices through the application of a dynamic fuzzy variable, which is defined automatically depending on the clustering results, taking into account the clustering evaluation process, in order to reach the best estimate for the specific objectives of the user (faster execution time, i.e. least number of clusters, or better estimation error, i.e. larger  $k$ ).

The proposed methodology includes the use of the K-means clustering algorithm, whose results are evaluated automatically so that a dynamic fuzzy variable can be defined automatically using these clustering data. The price estimation of the fuzzy variable is also evaluated automatically, taking into account the preference weight for both objectives (2E), and the final result is achieved.

The results of the presented experimental findings show that the proposed methodology is able to adapt itself to the input data, being capable of providing adequate results for when the user requests a better estimation accuracy (less error), and also for when a faster execution time is required. The estimation performed by the dynamic fuzzy variable is able to follow the real data series' trend in all cases, being the quality of the estimation a result of the user's preference.

In addition to the results obtained in (Faia et al., 2015a) the methodology proved to be able to solve the problem by presenting a valid solution. The 2E mechanism is used to give the user the freedom to choose their preference, which allows you methodology adapted to any user. In the case where there is a small amount of data is normal that is not as visible to the applicability 2E mechanism but where the set of data and high baste is essential for a high number of clusters may lead to a high execution time.

As future work it is proposed to try to apply other aggregation algorithms, and the use of other indicators which may be used to create the stopping criterion.

## Acknowledgements

This work has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 641794 (project DREAM-GO) and from FEDER Funds through COMPETE program and from National Funds through FCT under the project UID/EEA/00760/2013.

## 5. References

- Alvarado-Pérez, J.C., Peluffo-Ordóñez, D.H., Therón, R., 2015. Bridging the gap between human knowledge and machine learning. *ADCAIJ Adv. Distrib. Comput. Artif. Intell. J.* 4, 54.
- Chamoso, P., De La Prieta, F., 2015. Swarm-Based Smart City Platform: A Traffic Application. *ADCAIJ Adv. Distrib. Comput. Artif. Intell. Journal*; Vol 4, No 2.
- Dent, I., Craig, T., Aickelin, U., Rodden, T., 2014. Variability of Behaviour in Electricity Load Profile Clustering; Who Does Things at the Same Time Each Day? In: Perner, P. (Ed.), *Advances in Data Mining. Applications and Theoretical Aspects: 14th Industrial Conference, ICDM 2014, St. Petersburg, Russia, July 16-20, 2014. Proceedings.* Springer International Publishing, Cham, pp. 70–84.
- Faia, R., Pinto, T., Vale, Z., 2015a. Dynamic Fuzzy Estimation of Contracts Historic Information Using an Automatic Clustering Methodology. In: Bajo, J., Hallenborg, K., Pawlewski, P., Botti, V., Sánchez-Pi, N., Duque Méndez, N.D., Lopes, F., Julian, V. (Eds.), *Highlights of Practical Applications of Agents, Multi-Agent Systems, and Sustainability - The PAAMS Collection SE - 23.* Springer International Publishing, pp. 270–282.
- Faia, R., Pinto, T., Vale, Z., Pires, E.J.S., 2015b. Portfolio Optimization for Electricity Market Participation with Particle Swarm. 2015 26th Int. Work. Database Expert Syst. Appl.
- Ferreira, A.S., Pozo, A., Gonçalves, R.A., 2015. An Ant Colony based Hyper-Heuristic Approach for the Set Covering Problem. *ADCAIJ Adv. Distrib. Comput. Artif. Intell. J.* 4, 1.
- Hu, X.B., Wang, M., Paolo, E. Di, 2013. Calculating Complete and Exact Pareto Front for Multiobjective Optimization: A New Deterministic Approach for Discrete Problems. *IEEE Trans. Cybern.*
- Hwang, C.-L., Masud, A.S.M., 1979. *Multiple Objective Decision Making — Methods and Applications, Lecture Notes in Economics and Mathematical Systems.* Springer Berlin Heidelberg, Berlin, Heidelberg.
- Jain, A.K., 2010. Data clustering: 50 years beyond K-means. *Pattern Recognit. Lett.* 31, 651–666.
- Li, H., Tesfatsion, L., 2009. Development of open source software for power market research: The AMES test bed [WWW Document]. *J. Energy Mark.*
- MacQueen, J., 1967. Some Methods for Classification and Analysis of MultiVariate Observations. In: *5th Symposium on Mathematical Statistics and Probability.* pp. 281–297.
- Mahmoudi-Kohan, N., Moghaddam, M.P., Bidaki, S.M., 2009. Evaluating performance of WFA K-means and Modified Follow the leader methods for clustering load curves. *Power Syst. Conf. Expo. 2009. PSCE '09. IEEE/PES.*
- Mcculloch, W., Pitts, W., 1943. A Logical Calculus of the Ideas Immanent in Nervous Activity, *Bulletin of Mathematical Biophysics.*
- Meeus, L., Purchala, K., Belmans, R., 2005. Development of the Internal Electricity Market in Europe. *Electr. J.* 18, 25–35.
- Pinto, T., Morais, H., Sousa, T.M., Sousa, T., Vale, Z., Praça, I., Faia, R., Pires, E.J.S., 2015. Adaptive Portfolio Optimization for Multiple Electricity Markets Participation. *Neural Networks Learn. Syst. IEEE Trans.*
- Pinto, T., Vale, Z., Sousa, T.M., Praça, I., Santos, G., Morais, H., 2014. Adaptive Learning in Agents Behaviour: A Framework for Electricity Markets Simulation. *Integr. Comput. Eng.* 21, 399–415.
- Praça, I., Ramos, C., Vale, Z., Cordeiro, M., 2003. MASCEM: a multiagent system that simulates competitive electricity markets. *IEEE Intell. Syst.*
- Santos, G., Pinto, T., Praça, I., Vale, Z., 2016. MASCEM: Optimizing the performance of a multi-agent system. *Energy* 111, 513–524.
- Shahidehpour, M., Yamin, H., Li, Z., 2002. Market Overview in Electric Power Systems. *Mark. Oper. Electr. Power Syst. Sched. Risk Manag.*
- Zadeh, L.A., 1965. Fuzzy sets. *Inf. Control* 8, 338–353.