Discussing Different Clustering Methods for the Aggregation of Demand Response and Distributed Generation

Cátia Silva, Pedro Faria, and Zita Vale
GECAD -Research Group on Intelligent Engineering and Computing for AdvancedInnovation and Development
Polytechnic of Porto
Porto, Portugal
cvcds@isep.ipp.pt;pnf@isep.ipp.pt;zav@isep.ipp.pt

Abstract— **With the introduction of the Smart Grid context in the current network, it will be necessary to improve business models to include the use of distributed generation and demand response programs regarding the remuneration of participants as a form of incentive. Throughout this article a methodology is presented which will aggregate generation units and consumers participating in DR programs. A comparison of clustering methods will be carried out in order to understand which one of them will be the most appropriate for the scenario studied. After grouping all the resources, the remuneration of the groups are made considering the maximum rate in each group. The hierarchical clustering proved to be the most appropriate because it grouped the resources so that the total cost for the aggregator was the minimum.**

Keywords—Aggregation, Demand Response, Distributed Generation, Clustering Methods

I. Introduction

Nowadays electric power systems are focusing on the Smart Grid concept. Currently, the electricity grid is comprised of a large number of decentralized renewable energy sources and intelligent infrastructures are being installed in conventional power systems to enable the supply of electricity in an intelligent and controlled manner [1],[2]. In addition, introducing the Smart Grid concept into the power grid revolutionizes the way consumers can interact thanks to Demand Response (DR) programs.

The idea behind DR can be defined as the incentive to modify the load diagram by promoting interaction and responsiveness by end-use consumers, [3]. That said, electric power systems should be transparent, flexible, reliable and carefully managed. Therefore, it is necessary to update and improve current business models, especially with respect to the remuneration of these resources in the context of Smart Grid. Based on the methodology presented in [4], in which a virtual energy player aggregates several small resources, including consumers participating in DR programs, this article will then determine the remuneration structure that best fits the goals of the aggregator .

There is a necessity for supporting the decision in remuneration of DR and Distributed Generation (DG) for their participation in DR programs. In this way, the methodology was drawn up in order to address the respective remuneration to the aggregate resources. Test results, for different aggregation groups and with different methods of aggregation, are provided so there is a possibility of more accurate a comparison between them. This could be very helpful on decision making to which number of groups is optimal and more advantageous for virtual power players to minimize the operation costs and give fair remuneration to all resources involved. Thus, several tariffs are created for each group, where each group has an energy price derived from the resources, which are defined according to the actual energy scheduled for each resource in several operation scenario.

The constant accumulation of data sets urges the need to analyze and organize the information so that it is properly handled. Clustering is one of the analytical methods used. This method causes the data to be organized into groups, clusters, and the objects inserted here have a similarity between them and a disparity with the remaining objects of the remaining groups. Aggregation of these resources through clusters aims at collecting common characteristics that best define the resources in a specific context, [5]. However, this method is not applied in reality, it is considered by the authors as a crucial factor for the success of the implementation of DR programs. The model proposes that the remuneration of the consumers that participated with this type of programs would be made considering the revenues of the energy sold in the energy market by this aggregator.

This paper is the further development of previous works, as previously mencioned. Here we attempt to compare some of the major clustering methods within which stand out Partitioning Methods such as k-means, Partitioning Around Medoids (PAM) and Clustering Large Aplications (CLARA); Hierarchical Clustering; Fuzzy Clustering, studying c-means; Density-based Clustering, through DBSCAN and, finally, Model Based Clustering. The aggregation will be done for different k clusters and the comparison of the remuneration in the same way.

Section I refers to the introduction of the theme and purpose of the paper. Section II reports all steps of the methodology proposed, with a more detailed explanation. Section III presents the methods of clustering that will be compared with a brief explanation and comparison between them. Only in section IV is presented the case study that was used, presenting the results for the choosen senario in section IV. Finally, section VI presents the main conclusions drawn from the work done.

The present work was done and funded in the scope of the following projects: CONTEST Project (P2020 - 23575), and UID/EEA/00760/2013 funded by FEDER Funds through COMPETE program and by National Funds through FCT.

.

II. Approach

The proposed methodology, explained throughout this section, will be divided into four essential phases, as shown in the Fig.1.

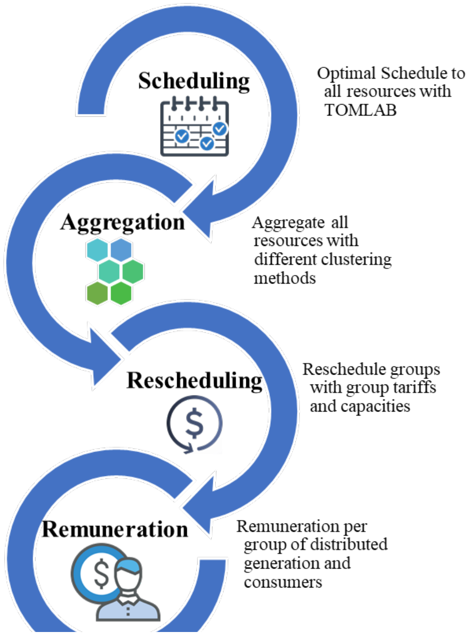

Scheduling
Optimal Schedule to all resources with TOMLAB
Aggregation
Aggregate all resources with different clustering methods
Rescheduling
Reschedule groups with group tariffs and capacities
Remuneration
Remuneration per group of distributed generation and consumers

Fig. 1. Proposed Methodology

This figure shows how an aggregator fits into the network infrastructure and how it handles the power market. In the first phase of this method, an optimization problem is formulated mathematically, minimizing operating costs taking into account all characteristics of the resources and the desired DR programs. The results of this optimization, which was the mixed-integer quadratic problem, were achieved through the MATLAB toolbox, TOMLAB.

In the second phase, the aggregation of resources all phase is done in order to provide the aggregator, with different groups, a considerable amount of energy for negotiation in the energy market. Here we chose the comparison of several clustering algorithms, then presented in section III, in order to understand the best clustering option for the studied data. This study was done through the software R, giving use to its potentialities for this type of analysis.

The third phase, rescheduling, is done create new tariffs by applying the max price of all resources in a group. In other words, there will be a group tariff for each cluster formed, where all resources in every single cluster, are remunerated at the same energy price. In this way, the aggregator could reduce its operating costs, benefiting from the potential by taking advantage of the full potential energy present in each of the formed groups.

The final phase, Remuneration of resources, is used as a motivation, an incentive, to the collaboration of all the resources associated with the aggregator in the operation of the network and as payment of the contribution of each of them in the final scheduling. The final remuneration to be paid is obtained through the tariffs of the previous phase.

III. Methods

In the perspective of machine learning, the technique of clustering, dividing data into different groups with similar objects, also known as clusters, is finding hidden patterns in the same information. This analysis is a method of unsupervised learning [6] and is used for the exploration of relationships within the existing set of patterns, and then organizing them into homogeneous groups. Unlike classification, a well-known method of supervised learning, no type of labeling is available in these standards for a priori differentiation. The measure to perceive the density of connection between objects within a cluster is called intra-connectivity. In such manner, the greater this intra-connectivity, the more certainty that the inserted objects have a very high degree of similarity between them. On the other hand, the concept of inter-connectivity measures the degree of connectivity between different clusters. It will be important that this value be low, meaning that each cluster is individually disparate from the rest, [7].

The first group to be analyzed, Partitioning Clustering, divides the data into non-overlapping subsets so that all data belongs to a cluster k. The number of k clusters to be generated must first be defined by the analyst. Partitioning Clustering includes k-means clustering or k-medoids clustering. The first one, can be represented by an algorithm with the homonymous name not happening the same with k-medoids clustering, which is commonly formed by PAM, [8].

K-means is the most common unsupervised machine learning algorithm for partitioning. One of the problems presented by this method is the sensitivity to noise and outliers. One of the variations of this method, defined by Hartigan-Wong in 1979, treats the total variation within a cluster as the sum of the squares of Euclidean distance between a point and the center of the cluster, assigning the point to the nearest k cluster.. After this step and throughout the algorithm, each cluster is represented by a new centroid, which corresponds to the average of the points assigned to the k cluster in question,[9].

PAM is an algorithm that is based on looking for objects, medoids, that can represent a cluster. Iteration after iteration, it is considered to exchange each medoid with each non-medoid and this exchange is validated only if there is an improvement over the criteria of the objective function - the minimization of the sum of the dissimilarities of all objects relative to the nearest medoid. PAM has a disadvantage relative to larger datasets. The problem of finding relatively small clusters in the presence of large clusters in the data set is a difficulty for this method. In situations where the dataset is greater than thousands of observations, typically PAMs are unsuccessful. Hence Clustering Large Aplications (CLARA) is an extension of this method to deal with this type of problems.

Clustering can be classified in Soft Clustering (Overlapping Clustering) or Hard Clustering (or Exclusive Clustering). In soft clustering, instead of putting each data point into a separate cluster, there is a probability of each point being allocated to a specific cluster. In hard clustering, each object is a member of that cluster completely or not. For example, K-means is considered Hard Clustering and the algorithms belonging to Fuzzy Clustering, like the most used C-means, are considered Soft Clustering, [10].

In Fuzzy Clustering it is considered a degree to which an element can belong to a given cluster, being this between 0 and 1. The points near the center of the cluster may have a degree greater than the points at the edges of the cluster. C-means is the most used within fuzzy clustering methods. The algorithm passes through the centroid of the cluster through the average of all the points, weighted by the degree of belonging to the cluster.

Unlike other algorithms, such as Partitional Clustering, in Hierarchical Clustering a number of k clusters are not required a priori, making it an asset. This method seeks to find a hierarchical structure according to a proximity matrix. They are usually presented in dendrograms or in binary trees. Therefore, if the analyst wants a more specific number of clusters, he will have to cut the dendogram at the desired level. Hierarchical Clustering (hclust) are subdivided into two groups: Agglomerative Clustering and Divisive Clustering. Agglomerative Clustering starts by treating all objects in the database as a singleton cluster, or a "leaf". Then, pairs of clusters are successively joined until all the information is in a single cluster, or "root". Divise Clustering is the opposite of agglomerative clustering since this algorithm works top-down, that is, it starts with root and divides until each of the objects belongs to an individual cluster. In this way, Agglomerative Clustering is a good option to identify smaller clusters. Already Divisive Clustering is the preferred choice when it comes to recognizing larger clusters, [11], [12].

Regarding density-based clustering, the dataset is grouped taking into account connectivity and density functions. One of the examples of this type of clustering is Density-Based Spatial Clustering and Application with Noise or DBSCAN. This algorithm was initially introduced in [13] and is characterized by identifying clusters of any shape in the data set containing noise and even outliers. The cluster definition in this algorithm is depicted through a region of points connected as a dense region collected from the dataset; and the other regions that are separated as sparse regions. The Euclidean distance is used to measure the similarity of points in the denser zone. DBSCAN requires two important parameters: epsilon ("eps"), which defines the radius of the neighborhood around any point x; and the minimum points ("MinPts") that define the minimum number of neighbors' points in radius eps. Through this feature, DBSCAN may not include all elements of the data set in the clustering leaving them in the k = 0 cluster belonging to the outliers and noise, [14].

There are more traditional methods, for example hierarchical clustering and k-means, which are heuristic methods and are not based on formal models. In addition, k-means normally initializes the algorithm in a random way, being able to obtain different results in different races, besides having to indicate the number of clusters a priori. An alternative to these methods will be Model Based Clustering (mclust) which considers the data to be formed by a mixture of underlying probability distributions, where each point can represent a different cluster or group. Unlike k-means, mclust uses a more flexible assignment, [14].

IV. Case Study

The case study where the proposed methodology was applied is a real distribution network composed of about 548 distributed producers and 20310 end-user consumers. The studied distribution network presents 30 kV with only one high voltage substation of 60 / 30kV, with the maximum capacity of 90 MVA. This distribution network consists of 937 buses where the aforementioned resources are scattered.

Demand side management can count with two main programs. One based on price, where consumers change their load by responding to changes in the price of electricity in real time, Real Time Pricing (RTP). And another based on incentives, in which consumers are paid at a fixed price per kW of reduced load.

Incentive based programs (Reduce, Cut) and RTP were applied in this study to different types of consumers: Domestic (DM), Small Commerce (SM), Medium Commerce (MC), Large Commerce (LC) and industrial (ID).

In terms of the type of distributed production, this study has Wind, Biomass, Small hydro, co-generation (CHP), Photovoltaic, Fuel cell and Waste-to-energy (WtE).

TABLE I presents the detailed information of these production units, showing the unit number by type, the unit operating price in m.u./kWh and the total available capacity. TABLE II shows the characterization for the types of consumers presented in this study and the possibilities of participation in the types of RD programs presented.

TABLE I. Distributed Generation Characterization

Designation	Nº of units	Capacity (kWh)	Price (m.u./kWh)
Wind	254	5 866.09	0.071
Co-generation	16	6 910.10	0.00106
Waste-to-energy	7	53.10	0.056
Photovoltaic	208	7 061.28	0.150
Biomass	25	2 826.58	0.086
Fuel cell	13	2 457.60	0.098
Small hydro	25	214.05	0.042
Total DG	548	25 388.79 kWh	

TABLE II. Demand Response Consumers Characterization

Designation	Reduce	Cut	RTP	Initial Price (m.u./kWh)
Domestic (DM)	●			0.12 (0.20)
Small commerce (SM)	●			0.18 (0.16)
Medium commerce (MC)		●		0.2 (0.20)
Large commerce (LC)		●		0.19 (0.20)
Industrial (ID)			●	0.15 (0.53)
Total Nº of DR	19 996	167	147	20 310
Total Capacity (kWh)	8 676	1 106	11 571	21 354.36

Since the main objective would be to generate clusters of existing resources, from the results of the optimization performed on the distribution network presented in the previous section, a scenario was constructed where only part of the resources belonging to this distribution network were grouped. Consequently, distributed resources such as Photovoltaic and Fuell-cell were withdrawn. In relation to DR, only the incentive-based program was studied, resulting in only consumers belonging to Small Commerce and Medium Commerce.

V. Results

Throughout this section are presented all the results obtained from this study as the analysis of them. Fig.2. presents the scheduling results for the scenario studied

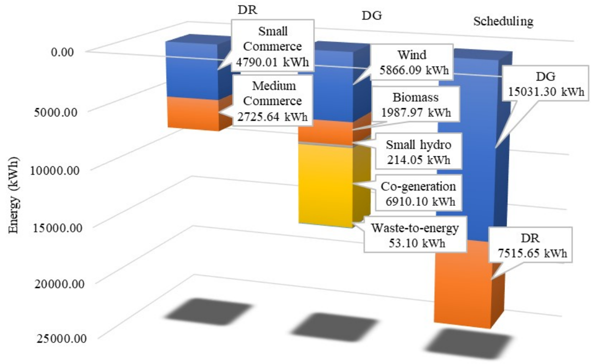


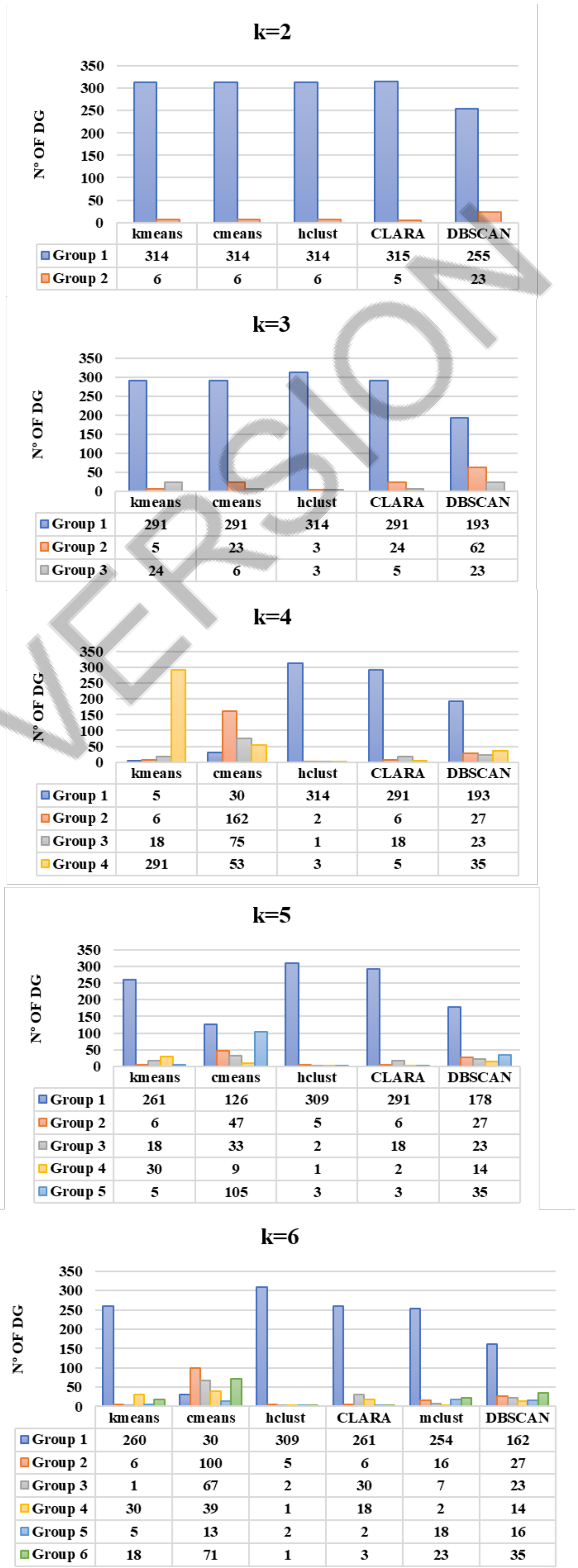
Fig. 2. Schedulling (phase 1) results for the scenario presented

The distributed producers represent the largest share in the global, presenting 66.67%. The rest is already guaranteed by the demand programs responds mostly to the Small Commerce type.

As previously mentioned, in this study the performance of six clustering methods was compared. Several scenarios were tested, varying the number of clusters for each of the methods. The number k of clusters from 2 to 6 was considered and tested. Since, as stated in section III, the PAM method does not handle very well the large dataset and, in order to be coherent, it was used the CLARA method for the two cases studied. The results obtained then group distributed producers and consumers into clusters, as shown in Fig.3. and Fig. 4, respectively.

The data set presented by Fig.3 can be considered small, with about 320 producers distributed. The performance of the methods of Partitioning Clustering, k-means and CLARA, remained similar across k clusters, differing only in some units and in group switching. In relation to hclust, aggregation is done in a more organized way, maintaining coherence throughout the tests. In relation to Fuzzy Clustering, c-means, due to its randomness, did not find a pattern, although k = 2 and k = 3 be similar to the other methods. As for density-based clustering, DBSCAN identified elements such as outliers and noise and thus did not group all the distributed producers, leaving about 42 elements, approximately 13% of the samples. For the mclust method, the characteristics of this method dictated that the optimal number of clusters for this data set would be k = 6 showing some similarity with the results obtained in the partitioning clustering methods.

Regarding the set of data grouped by consumers, the difference in size is highlighted in relation to the previously studied, with about 9910 consumers belonging to demand responder programs. This clearly affected the results obtained, noting already the difference between the methods of partitioning clustering. CLARA maintained a similar behavior throughout the 5 tests while k-means, in k = 6 completely change the values for each group. As said before, in the section III, this method normally initializes the algorithm in a random way which may affect the results. DBSCAN performed relatively better leaving only 0.12% of the data excluded from the aggregation. In relation to mclust, this time, the optimal number of clusters selected was k = 2 similar to its result for CLARA and c-means.



k=2

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	314	314	314	315	255
Group 2	6	6	6	5	23

k=3

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	291	291	314	291	193
Group 2	5	23	3	24	62
Group 3	24	6	3	5	23

k=4

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	5	30	314	291	193
Group 2	6	162	2	6	27
Group 3	18	75	1	18	23
Group 4	291	53	3	5	35

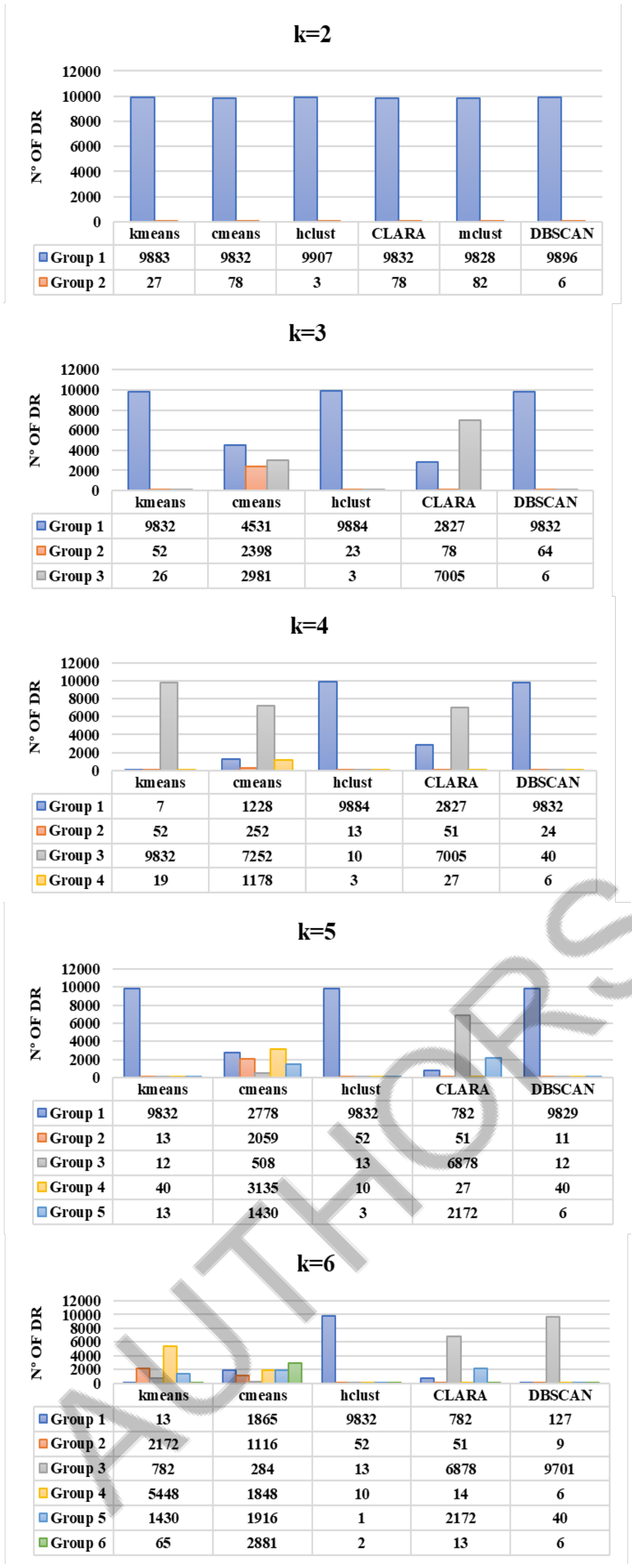
k=5

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	261	126	309	291	178
Group 2	6	47	5	6	27
Group 3	18	33	2	18	23
Group 4	30	9	1	2	14
Group 5	5	105	3	3	35

k=6

	kmeans	cmeans	hclust	CLARA	mclust	DBSCAN
Group 1	260	30	309	261	254	162
Group 2	6	100	5	6	16	27
Group 3	1	67	2	30	7	23
Group 4	30	39	1	18	2	14
Group 5	5	13	2	2	18	16
Group 6	18	71	1	3	23	35

Fig. 3. Comparison between methods for different k clusters (DG)



	kmeans	cmeans	hclust	CLARA	mclust	DBSCAN
Group 1	9883	9832	9907	9832	9828	9896
Group 2	27	78	3	78	82	6

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	9832	4531	9884	2827	9832
Group 2	52	2398	23	78	64
Group 3	26	2981	3	7005	6

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	7	1228	9884	2827	9832
Group 2	52	252	13	51	24
Group 3	9832	7252	10	7005	40
Group 4	19	1178	3	27	6

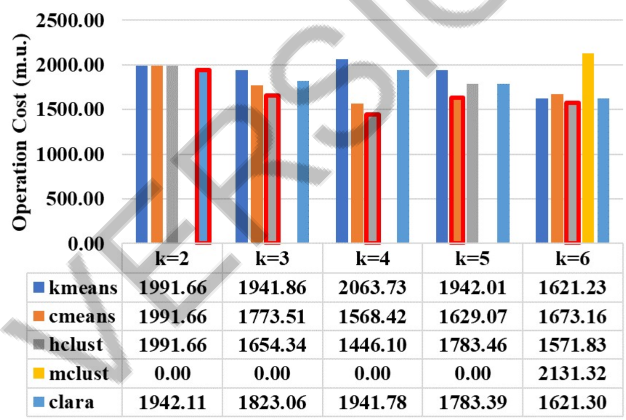
	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	9832	2778	9832	782	9829
Group 2	13	2059	52	51	11
Group 3	12	508	13	6878	12
Group 4	40	3135	10	27	40
Group 5	13	1430	3	2172	6

	kmeans	cmeans	hclust	CLARA	DBSCAN
Group 1	13	1865	9832	782	127
Group 2	2172	1116	52	51	9
Group 3	782	284	13	6878	9701
Group 4	5448	1848	10	14	6
Group 5	1430	1916	1	2172	40
Group 6	65	2881	2	13	6

Fig. 4. Comparison between methods for different k clusters (DR)

Looking more closely, Fig.5. shows the energy per group, according to the type of power source of the distributed producer. In this case, in order to compare all methods, k = 6 was chosen.

Now it is possible to check the differences between methods that although the number of elements is similar, the type of power source is very different. K-means and CLARA assigned levels to all types of energy. With c-means we cannot find a logical sequence for assigning the values to these groups. Hclust grouped in group 1 the wind, biomass, small hydro, waste-to energy and co-generation elements with the lowest energy value, approximately between 3.27 kWh of Small Hydro and 114,85 kWh of Biomass. The remaining groups are composed only of co-generation with values higher than 233.37 kWh. Mclust separated energy source types by groups. DBSCAN considered values above 23.41 kWh as noise and outliers, leaving only wind and small hydro energy sources. Regarding the costs of operation by the aggregator after the application of the tariff, Fig.6. presents the result for each of the methods for all k clusters tested. The minimum values for each of the tests are indicated in red.



	k=2	k=3	k=4	k=5	k=6
kmeans	1991.66	1941.86	2063.73	1942.01	1621.23
cmeans	1991.66	1773.51	1568.42	1629.07	1673.16
hclust	1991.66	1654.34	1446.10	1783.46	1571.83
mclust	0.00	0.00	0.00	0.00	2131.32
clara	1942.11	1823.06	1941.78	1783.39	1621.30

Fig. 5. Comparison between methods for different k clusters (Total Cost)

In most cases, the hclust aggregation obtained the lowest value for operating costs, where k = 4 the lowest having obtained 1446.10 m.u. DBSCAN method was removed from this comparison since it would not include all the elements and would not of a fair comparison.

VI. Conclusion

The methodology proposed and presented in this paper aims to support energy resources aggregators. The method presents the resolution of an optimal schedule for the resources insert in the distribution network studied, that is then grouped through clustering methods. At this stage several methods are compared for aggregation, counting on k-means, c-means, hclust, mclust, clara and DBSCAN. It was tested the ideal number of clusters, between 2 and 6, both for the distributed producers' data set and for consumers of a demand response program. In general, the behavior of the methods was consistent with what was expected. It is concluded, however, that although it is an advantage in some situations, the fact that DBSCAN did not include all the elements, since it considered them outliers or noise, was detrimental in the evaluation. The next phase went through a rescheduling of the groups, calculating the tariffs for each of them with the intention of reducing the operating costs for the aggregator. It was then compared all methods, all tests and the remuneration were done group by group to define the fairer tariff.

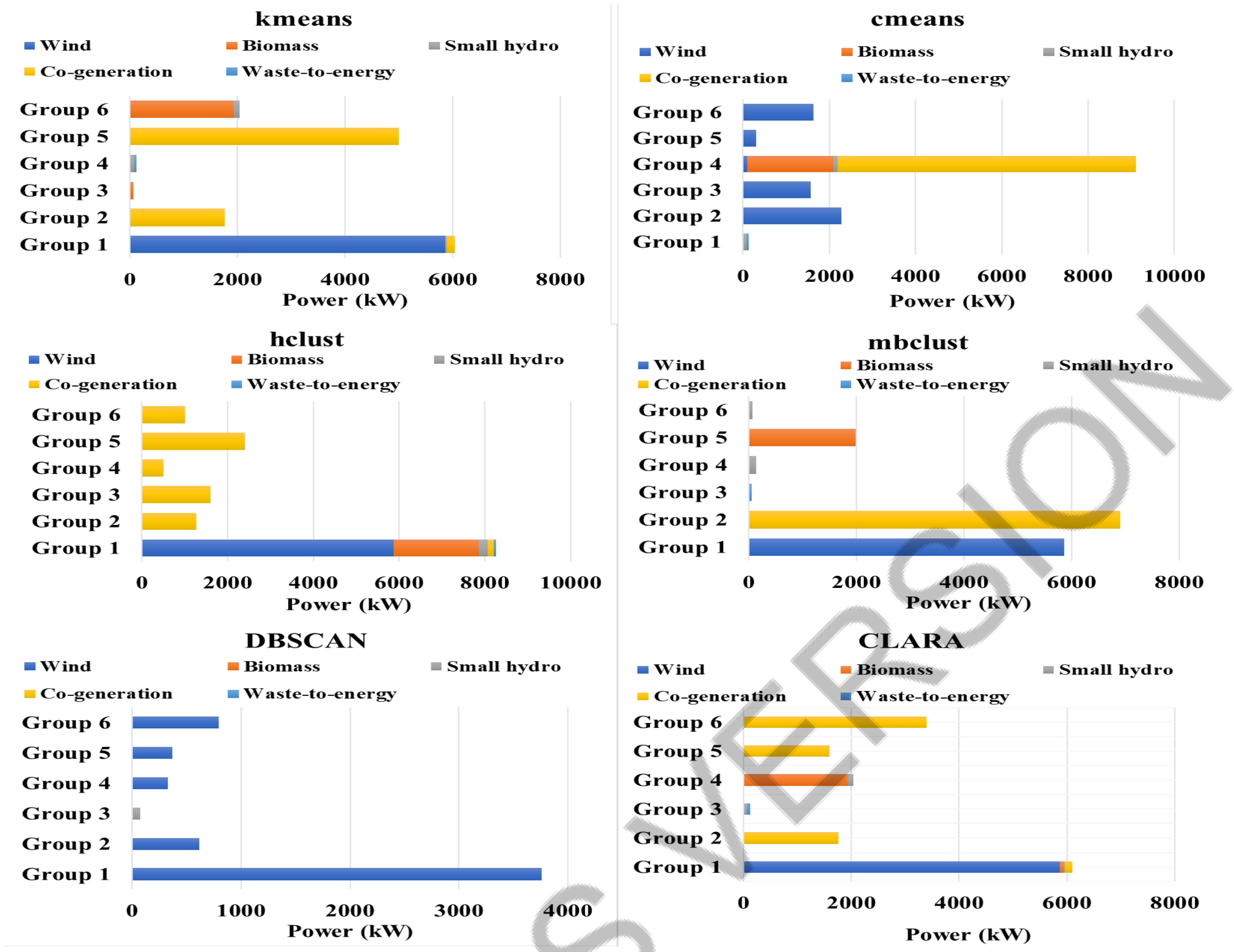


Fig. 6. Comparison between methods for k=6 (DG)

REFERENCES

[1] P. Wirasanti, E. Ortjohann, A. Schmelter, and D. Morton, “Clustering power systems strategy the future of distributed generation,” *Int. Symp. Power Electron. Power Electron. Electr. Drives, Autom. Motion*, pp. 679–683, 2012.

[2] P. Siano, “Demand response and smart grids - A survey,” *Renew. Sustain. Energy Rev.*, vol. 30, pp. 461–478, 2014.

[3] P. Faria, Z. Vale, and J. Baptista, “Demand Response Programs Design and Use Considering Intensive Penetration of Distributed Generation,” *Energies*, vol. 8, no. 6, pp. 6230–6246, Jun. 2015.

[4] P. Faria, J. Spínola, and Z. Vale, “Aggregation and Remuneration of Electricity Consumers and Producers for the Definition of Demand-Response Programs,” *IEEE Trans. Ind. Informatics*, vol. 12, no. 3, pp. 952–961, 2016.

[5] Z. Vale, H. Morais, S. Ramos, J. Soares, and P. Faria, “Using data mining techniques to support DR programs definition in smart grids,” in *IEEE Power and Energy Society General Meeting*, 2011, pp. 1–8.

[6] P. Rai and S. Singh, “A Survey of Clustering Techniques,” *Int. J. Comput. Appl.*, vol. 7, no. 12, pp. 1–5, 2010.

[7] R. Xu, “Survey of clustering algorithms for MANET,” *IEEE Trans. Neural Networks*, vol. 16, no. 3, pp. 645–678, 2005.

[8] M. Van Der Laan, K. Pollard, J. Bryan, M. J. Van Der Laan, and K. S. Pollard, “Journal of Statistical Computation and Simulation A new partitioning around medoids algorithm A NEW PARTITIONING AROUND MEDOIDS ALGORITHM,” *J. Stat. Comput. Simul.*, vol. 73, no. 8, pp. 575–584, 2003.

[9] M. Gi and M. Herbster, “K -means algorithm,” no. 13, pp. 354–360, 2014.

[10] D. Jyoti Bora and A. Kumar Gupta, “A Comparative study Between Fuzzy Clustering Algorithm and Hard Clustering Algorithm,” *Int. J. Comput. Trends Technol.*, vol. 10, no. 2, 2014.

[11] J. Pagel, M. Campion, A. S. Nair, and P. Ranganathan, “Clustering analytics for streaming smart grid datasets,” *Clemson Univ. Power Syst. Conf. PSC 2016*, 2016.

[12] A. Nasiakou, M. Alamaniotis, L. H. Tsoukalas, and G. Karagiannis, “A three-stage scheme for consumers’ partitioning using hierarchical clustering algorithm,” in *2017 8th International Conference on Information, Intelligence, Systems and Applications, IISA 2017*, 2018, vol. 2018–Janua.

[13] M. Daszykowski and B. Walczak, “Density-Based Clustering Methods,” in *Comprehensive Chemometrics*, vol. 2, 2010, pp. 635–654.

[14] M. P. Singh Pradeep, “Survey of Density Based Clustering Algorithms,” *Int. Conf. Inven. Comput. Informatics (ICICI 2017)*, no. Icici, pp. 313–317, 2017.