

Assessment of Distributed Generation Units Remuneration Using Different Clustering Methods for Aggregation

Cátia Silva, Pedro Faria and Zita Vale

GECAD -Research Group on Intelligent Engineering and Computing for Advanced Innovation and Development, Polytechnic of Porto

Porto, Portugal

cvcds@isep.ipp.pt;pnf@isep.ipp.pt;zav@isep.ipp.pt

Abstract— The stakeholders that belong to the energy market will have to adapt to the changes that the implementation of the concept of Smart Grid imposes. This concept requires new business models that include the demand response programs, the use of distributed generation and especially the remuneration that will be made for their contribution. The exposed methodology can be presented as a solution for virtual power players in this new challenge. Throughout this article, this methodology was tested regarding the remuneration of aggregate groups of distributed generation. It will also be analyzed the meaning of this tariff for both sides - aggregator and producers.

Keywords— Remuneration, Smart Grids, Distributed Generation, Clustering

I. INTRODUCTION

Currently all the sectors are changing in a way to reduce the environmental problem that affects the world for many years. The consequences are irrefutable and scientific evidence, such as the depletion of the ozone layer, global warming and the extinction of thousands of species, awakens in humanity the duty to protect and preserve the planet.

The current situation implies change and it is believed that the solution, at least in the electric sector, can pass through the idea of endogenous renewable energies, making the concept of distributed production an indisputable reality. The present electricity grid infrastructure is not yet ready for the near future, although there are already efforts to transform the conventional power systems to allow the electricity supply in an intelligent and controlled manner.

The promotion of distributed generation (DG) closer to the places of consumption will be the key factor to reduce the need for transport over considerable distances - avoiding greater losses, and also to opt for a more environmentally friendly production nearby the places where it can be immediately consumed. As will be expected, the producers who will be available for this type of model will have to be remunerated according to their contribution, [1], [2].

The goal is then to promote electric energy systems that must be flexible, transparent, reliable and, mainly, very well managed. All players in these systems will then be affected, and markets will not be an exception. Therefore, it is essential to update the business models for these new changes, since the role of consumers is being revolutionized too. With the introduction of the Smart Grids concept, the possibility of two-way communication between utilities and energy end-users will be one of the key ideas of this innovation. This will allow the response of demand to changes in the network through appropriate Demand Response (DR) programs, [3], [4].

Therefore, it will be necessary to properly manage energy on the consumption side to optimize the

electrical system, one of the most complex and important systems today. Application of the Demand Side Management concept may involve applying a variety of conditions to improve energy efficiency, applying energy tariffs to encourage consumption at certain times of day to reduce peak, allowing and promoting a more responsive interaction of consumers, [5], [6].

In this way and based on [7], a methodology is proposed in which virtual energy players aggregates several small resources, including consumers participating in DR programs and DG. The methodology was conceived to address the respective remuneration to the aggregate resources. Different groups were tested, with different clustering methods, allowing the comparison and the possibility of choosing the most adjusted. This will be very beneficial to virtual power players (VPP) in decision making in the optimal number of groups that can minimize operating costs. This is not the only objective of minimization since it will also be necessary to allocate fair rates in the remuneration of all the resources that are involved. That being the case, for each group a tariff was created.

In this paper, the focus will be in the remuneration part. The main objective is to understand the effect of each tariff, for each group, in each type of resource. In the case study presented, only DG resources will be studied. In this way, considering some clustering methods such as k-means, cmeans, Clustering Large Applications (CLARA) and hierarchical clustering, it will be compared if the group to which a particular producer was inserted would be the most adjusted or, in other words, if this was translated into benefit to him regarding the lowest tariff in all groups.

The introduction to the topic of this article is presented in section I. For a more detailed explanation of the proposed methodology, it is presented in section II. The methods used for clustering in the case of study will be briefly discussed in section III. Sections IV and V present the case studied and the results obtained, respectively. Finally, section VI presents the main conclusions drawn from the study.

II. APROACH

Throughout this section, the methodology proposed and applied in this article will be dissected in detail. Fig.1. shows the steps followed.

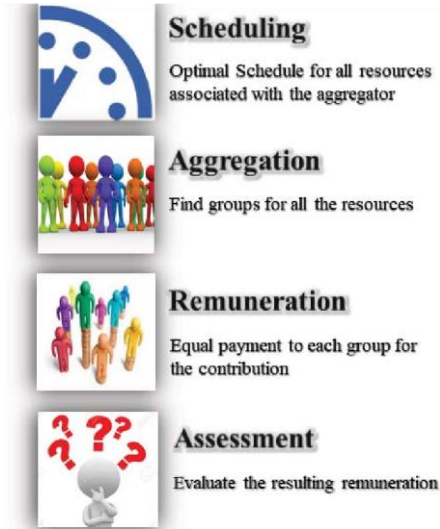


Fig. 1. Proposed methodology

The proposed methodology demonstrates the role of an aggregator in the grid infrastructure and how it can deal with the market in this sector. Firstly, an optimization will be performed to schedule all the energy resources aggregated being DG units, DR consumers and suppliers. The profiles used as parameter data include, for example, the elasticity of the demand price, the levels of comfort sought by each of them, the possibilities of direct control of load and the existence of production of either heat or electricity, ... The objective function is to minimize operating costs from the point of view of the VPP and remunerate fairly all participants. In this way, price and operation constraints were considered in this optimization, also operational constraints imposed by VPP to achieve its goals. The MATLAB software potentials were used through the toolbox, TOMLAB.

Then, the aggregation phase is performed to define suitable groups taking into account results from the first phase – schedule power for each resource. By grouping these small resources, VPP will have a considerable amount of energy for the negotiations in the market and allows the expansion of their participation in DR programs. Clustering methods were compared for several k – number of clusters. The methods will be presented in the following section more explicitly. This study was carried out using software R.

The remuneration stage is the motivation for continued collaboration of all resources hitherto associated with the aggregator in the network operation, compensation. Here, taking into account the tariffs that were designed specifically for each group, through the maximum of each one, all resource will be paid at the end of the schedule.

Finally, it will be tested if the group attributed to a set of producers belonging to different buildings, was the one that generated more benefit for the respective one. In other words, in what sense, different methods of clustering or even different numbers of k clusters will affect the final remuneration of producers.

III. METHOD

In this case study, after the optimal schedule for all resources, the above mention methods were applied. Only the results for the DG units are analyzed. These units are from consumers that can act as a prosumer – have a DG unit in their building. The main objective is to understand the benefits of using the proposed method for this type of producers.

The technique for dividing a set of data into groups of similar objects, finding hidden patterns, from a machine learning perspective, is called clustering. This method of unsupervised learning is used to explore relationships in a set of patterns and then succeed in organizing them into homogeneous groups.

There are several divisions for the type of clustering methods but in this paper only Partitioning Methods such as k -means, PAM and CLARA; Hierarchical Clustering; Fuzzy Clustering were studied and compared.

With these methods, it would be possible to form the best groups, finding patterns and similarities in the results of the optimization. Price and the schedule power for all the resources were used as parameters for the aggregation.

The most common method of unsupervised machine learning when it comes to partitioning is k -means. In 1979, Hartigan-Wong defined one of the possible variations of this method. The total variation within a cluster is then taken as the sum of the squares of Euclidean distance between a point and the center of the cluster, and then assigns the point to the nearest cluster. This method does, however, present a problem that can be considered serious for some cases since it cannot handle noise and outliers, [10].

In relation, k -medoids, Partitioning Around Medoids (PAM) is the most well-known algorithm and it is defined by the iterative search of objects, medoids, that can represent a cluster. The exchange between medoid and non-medoid is made and validated only when there is an improvement in the criterion of the objective function - the minimization of the sum of the dissimilarities of all objects relative to the nearest medoid. However, finding relatively small clusters in the large dataset is a difficulty for this method. In such situations, PAM may not achieve the best results. Hence CLARA is an extension of this method to deal with this type of problems.

Clustering methods can be organized into Hard and Soft Clustering. In the first, each data point either belongs to a cluster completely or not. In the case of soft clustering, rather than instead of putting each data point into a separate cluster, the probability or likelihood of that data point to be in those clusters is assigned. In 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. In fuzzy clustering, points close to the center of a

cluster, may be in the cluster to a higher degree than points in the edge of a cluster.

One of the advantages to Hierarchical Clustering (hclust) in relation to the previous methods resides in the fact that, previously, it is not necessary to know the number of clusters. As the name implies, this method hierarchically structures the data set according to the proximity matrix. The algorithm can start, for example, in a singleton cluster (leaf) and iteratively joins these into pairs of clusters until it finds a cluster that agglomerates the entire data set, root. The opposite algorithm is also considered, in other words, to start in root and subdivide to the leaf. [11]

IV. CASE STUDY

In the application of the methodology proposed in the previous section, a case study was created through a part of a real Portuguese distribution network, composed of about 548 distributed producers and 9910 end-user consumers.

The DMS applied in this case study has two major consumer programs: price-based and incentive-based. In the first case, consumers

change their consumption patterns with responses to real-time electricity price changes (RTP). On the other hand, through incentive-based programs, consumers are paid at a fixed price kW of reduced load.

Price-based (RTP) and Incentive based programs (Reduce, Cut) were applied in this study to different types of consumers: Domestic (DM), Small Commerce (SM), Medium Commerce (MC), Large Commerce (LC) and industrial (ID). TABLE I presents the characterization for the types of DR consumers and the possibilities of participation in the types of DR programs presented.

TABLE I. 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 N° of DR	19 996	167	147	20 310
Total Capacity (kWh)	8 676	1 106	11 571	21 354.36

Talking about the distributed resources, several types were considered in this study, namely wind, biomass, small hydro, co-generation, photovoltaic, fuel-cell and waste-to-energy. TABLE II presents the detailed information for each DG, showing the unit number by type, the unit operating price in m.u./kWh and the total available capacity.

TABLE II. DISTRIBUTED GENERATION CHARACTERIZATION

Designation	N° 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	

After the optimization phase, the aggregation of resources through several methods was performed. As mentioned in section II, four clustering methods were tested for several k clusters. With this result it was possible to carry out the third phase of the methodology, rescheduling and then the remuneration of the respective groups according to the new tariff.

In this article, it will then be ascertained in which situations and in which methods the DG producers were or were not benefited by being assigned to a particular group in relation to the others. Thus, 2592 scenarios were chosen to study this problem. From that point on, some producers of four types of DG were selected: wind, waste-to-energy, biomass and co-generation. The results of this study will then be presented in the next section.

V. RESULTS

It is in this section that the results from all DG units of the case study will be presented and analyzed. Some producers were chosen to be studied after being grouped with the clustering methods presented earlier. The objective is to understand if the resulting groups bring more benefit for the producer side.

Since were studied k clusters between k = 4 and k = 6 for all proposed methods, six group tariffs were created to remunerate every DG unit that collaborate with the aggregator. This groups were formed through the maximum tariff of all the objects in it. In this way, everybody would benefit.

TABLE III. TARIFFS FOR EACH GROUP

Group	k=4 (m.u./kWh)	k=5 (m.u./kWh)	k=6 (m.u./kWh)
1	0.00120	0.00124	0.00124
2	0.00124	0.00123	0.10034
3	0.00120	0.00120	0.00124
4	0.10034	0.00120	0.00123
5	-	0.10034	0.00120
6	-	-	0.00120

Through TABLE III it is possible to verify the value of each tariff and which one has the lowest price. From this, it is deduced that the lowest one has a value of about 0.00120 m.u./kWh. Thus, for k = 4 the lowest tariff is in group 1 and 3. For k = 5, two groups equate the lowest tariff as well, groups 3 and 4, as in k = 6, these being groups 5 and 6.

In the figures that will be analyzed after, there are two distinct areas, the areas where the tariff is lower, as previously mentioned, marked in green. The remaining areas represent the remaining tariffs. The figures are divided into groups of clusters, between 4 and 6, which are represented as letters a, b and c, respectively. Each one of them shows the results for all methods in order to understand to which group a particular producer is assigned and if it's the one with the lowest tariff. The methods are represented in the horizontal axis and the different groups in the vertical axis and all these assumptions are applied to all the results in this section.

A. Wind

In the first place, it will be presented, through Fig. 2 results for Wind.

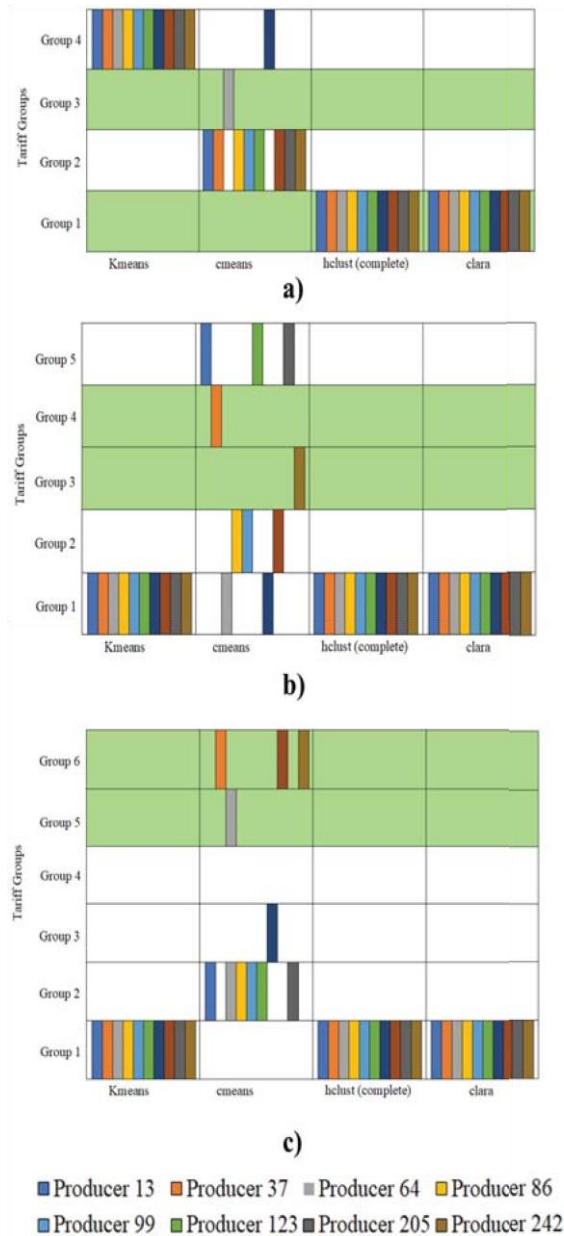


Fig. 2. Wind: Comparison between methods for a set of producers a) k=4 b) k=5 c) k=6

In analyzing Fig. 2, it can be seen that in none of the cases, the K-means method attributed the chosen producers to the lowest tariff. For c-means, the largest number of producers with the lowest tariff was $k = 6$. Since hclust and CLARA, the results obtained for the chosen producers are equal and only in $k = 4$ it assigns the lowest tariff, but always maintaining the same group for all k clusters.

B. Waste-to-Energy

Fig. 3 shows the comparison between the methods for a set of producers using Waste-to-energy.

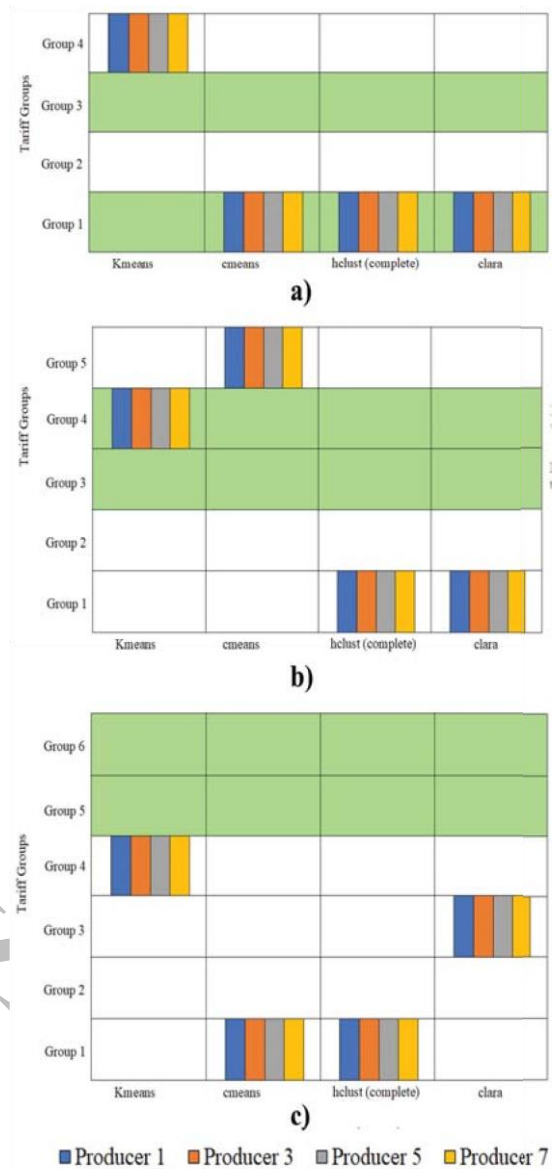


Fig. 3. Waste-to-Energy: Comparison between methods for a set of producers a) k=4 b) k=5 c) k=6

In this case, at $k = 4$, only K-means did not allocate the lowest tariff, placing all producers chosen in group 4, the highest tariff group with about 0.10034 m.u./kWh. On the other hand, in $k = 5$, k-means was the only one which attributed the group of producers to the lowest tariff. Finally, in $k = 6$, none of the methods assigned the lowest tariff to the producers, with k-means placing them in group 4 at 0.00123 m.u./kWh, CLARA in group 3 at 0.00124 m.u./kWh and the other two methods in the group 1 also at 0.00124 m.u./kWh.

C. Biomass

In relation to Fig. 4, it presents the comparison of the clustering methods but in this case with respect to Biomass.

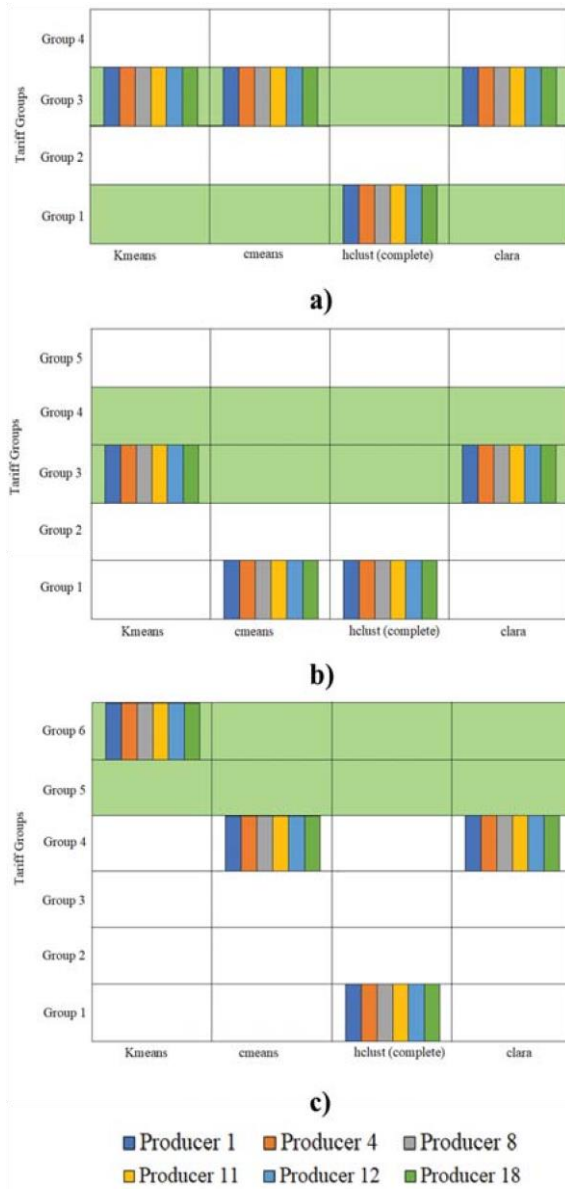


Fig. 4. Biomass: Comparison between methods for a set of producers a) k=4 b) k=5 c) k=6

In $k = 4$, all methods assigned the lowest tariff to all producers chosen. The same did not happen in $k = 5$

since k-means and CLARA attributed these 6 producers to group 3 receiving the lowest tariff, but c-means and hclust allocated them to group 1 receiving 0.00124 m.u./kWh. Finally, $k = 6$, where k-means kept the producers at the lowest rate in group 6. For c-means and CLARA, they assigned group 4, which could receive 0.00123 m.u./kWh. Hclust, dictated that producers receive 0.00124 m.u./kWh.

D. Co-generation

In Fig. 5, it shows the results obtained for the selected co-generation producers.

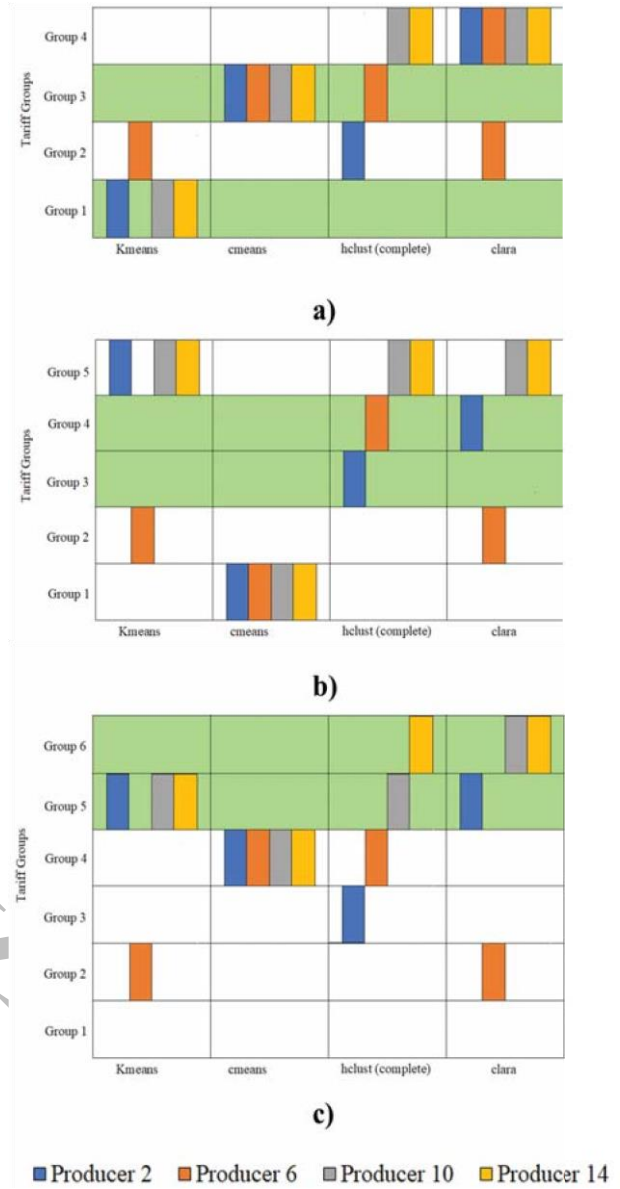


Fig. 5. Co-generation: Comparison between methods for a set of producers a) k=4 b) k=5 c) k=6

For the k-means method, in $k = 4$, only one of the producers did not receive the lowest tariff. The opposite happens in $k = 6$. In $k = 5$, none of them received the lowest rate. With c-means, only at $k = 4$, the producers receive the lowest tariff. Through hclust, producer 6 was the only one that received the lowest tariff at $k = 4$. At $k = 5$, producer 2 joined the last one, at the lowest tariff. The remaining two producers, only in $k = 6$, received the lowest tariff using this method. Finally, CLARA, in $k = 4$, all received tariffs other than the lowest. In $k = 5$, only producer 2 received the lowest tariff and at $k = 6$, only the producer 6 doesn't received the lowest one.

In order to gain a clearer picture of the impact on a given building, a producer of each type of DG from the previous set was selected. TABLE IV shows the percentage comparison between the tariff of the groups attributed to each one of the producers by the methods studied, in relation to the tariffs in other existing groups.

TABLE IV. TARIFF COMPARISON

Wind: Producer 37					
Group	Price (m.u./kWh)	k-means	c-means	hclust	clara
		1	4	1	1
1	0.00124	-	-3.28%	-	-
2	0.00123	0.81%	-2.47%	0.81%	0.81%
3	0.0012	3.28%	0.00%	3.28%	3.28%
4	0.0012	3.28%	-	3.28%	3.28%
5	0.10034	-195.12%	-195.27%	-195.12%	-195.12%
Waste to Energy: Producer 1					
Group	Price (m.u./kWh)	k-means	c-means	hclust	clara
		4	5	1	1
1	0.00124	-3.28%	195.12%	-	-
2	0.00123	-2.47%	195.16%	0.81%	0.81%
3	0.0012	0.00%	195.27%	3.28%	3.28%
4	0.0012	-	195.27%	3.28%	3.28%
5	0.10034	-195.27%	-	-195.12%	-195.12%
Small Hydro: Producer 15					
Group	Price (m.u./kWh)	k-means	c-means	hclust	clara
		4	5	1	1
1	0.00124	-3.28%	195.12%	-	-
2	0.00123	-2.47%	195.16%	0.81%	0.81%
3	0.0012	0.00%	195.27%	3.28%	3.28%
4	0.0012	-	195.27%	3.28%	3.28%
5	0.10034	-195.27%	-	-195.12%	-195.12%
Biomass: Producer 8					
Group	Price (m.u./kWh)	k-means	c-means	hclust	clara
		3	1	1	3
1	0.00124	-3.28%	-	-	-3.28%
2	0.00123	-2.47%	0.81%	0.81%	-2.47%
3	0.0012	-	3.28%	3.28%	-
4	0.0012	0.00%	3.28%	3.28%	0.00%
5	0.10034	-195.27%	-195.12%	-195.12%	-195.27%
Co-generation: Producer 14					
Group	Price (m.u./kWh)	k-means	c-means	hclust	clara
		5	1	5	5
1	0.00124	195.12%	-	195.12%	195.12%
2	0.00123	195.16%	0.81%	195.16%	195.16%
3	0.0012	195.27%	3.28%	195.27%	195.27%
4	0.0012	195.27%	3.28%	195.27%	195.27%
5	0.10034	-	-195.12%	-	-

Analyzing from the point of view of the producer and by groups, those who were assigned to group 1 benefited from a slightly higher remuneration than those who were on group 2, not reaching 3 %; for groups 3 and 4, since the tariffs were the same, the difference is slightly more significant reaching 3.28 % since it is the group 5 that presents a greater disparity with the rest since the difference is around 195 %. Clearly, for all producers, it would be more beneficial to be assigned to group 5, and the c-means method attributed this same group to more producers.

VI. CONCLUSION

The methodology presented in this article had as main objective to provide the aggregator with all the necessary tools to be successful in its activities.

Four phases were created: optimization, aggregation, remuneration and finally, an assessment, in this case, for the DG producers. In this article, we focused on the last stage. The lower rates of the groups compared to the DG resource allocated to a given group were compared. With this analysis it was possible to verify which group and method was the more benefit for the producer side. This analysis was performed according to the group that was assigned to each resource and by each method of clustering.

ACKNOWLEDGMENT

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.

REFERENCES

- [1] P. Palensky and D. Dietrich, "Demand side management: Demand response, intelligent energy systems, and smart loads," *IEEE Trans. Ind. Informatics*, vol. 7, no. 3, pp. 381–388, 2011.
- [2] G. J. Tsekouras, P. B. Kotoulas, C. D. Tsirekis, E. N. Dialynas, and N. D. Hatziaargyriou, "A pattern recognition methodology for evaluation of load profiles and typical days of large electricity customers," *Electr. Power Syst. Res.*, vol. 78, no. 9, pp. 1494–1510, 2008.
- [3] S. Ghosh, X. A. Sun, and X. Zhang, "Consumer profiling for demand response programs in smart grids," *2012 IEEE Innov. Smart Grid Technol. - Asia, ISGT Asia 2012*, pp. 1–6, 2012.
- [4] R. Deng, Z. Yang, M.-Y. Chow, and J. Chen, "A Survey on Demand Response in Smart Grids: Mathematical Models and Approaches," *IEEE Trans. Ind. Informatics*, vol. 11, no. 3, pp. 1–11, 2015.
- [5] 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.
- [6] Z. Vale, H. Morais, S. Ramos, J. Soares, and P. Faria, "Using data mining techniques to support DR programs definition in smart grids," in *2011 IEEE Power and Energy Society General Meeting*, 2011, pp. 1–8.
- [7] 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.
- [8] P. Rai and S. Singh, "A Survey of Clustering Techniques," *Int. J. Comput. Appl.*, vol. 7, no. 12, pp. 1–5, 2010.
- [9] A. Kassambara, *Practical Guide to Cluster Analysis in R*, 1st ed. .
- [10] L. Morissette and S. Chartier, "The k-means clustering technique: General considerations and implementation in Mathematica," *Tutor. Quant. Methods Psychol.*, vol. 9, no. 1, pp. 15–24, 2013.
- [11] C. N. Silla and A. A. Freitas, "A survey of hierarchical classification across different application domains," *Data Min. Knowl. Discov.*, vol. 22, no. 1–2, pp. 31–72, 2011.