POLITECNICO DI TORINO
Repository ISTITUZIONALE

Assessment of Graph-Based Clustering Algorithms for Alternative Bidding Zone Configurations

Original
Assessment of Graph-Based Clustering Algorithms for Alternative Bidding Zone Configurations / Chicco, G., Colella, P., Mazza, A., Russo, A., Ilea, V., Quaglia, F., Limone, M., Bovo, C.. - ELETTRONICO. - (2024), pp. 1-6. (2024 AEIT International Annual Conference (AEIT) Trento, Italy 25-27 September 2024) [10.23919/AEIT63317.2024.10736712].

Availability:
This version is available at: 11583/2994296 since: 2024-11-11T17:17:05Z

Publisher:
IEEE

Published
DOI:10.23919/AEIT63317.2024.10736712

Terms of use:

This article is made available under terms and conditions as specified in the corresponding bibliographic description in the repository

Publisher copyright
IEEE postprint/Author's Accepted Manuscript

©2024 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collecting works, for resale or lists, or reuse of any copyrighted component of this work in other works.

(Article begins on next page)

21 November 2026

Assessment of Graph-Based Clustering Algorithms for Alternative Bidding Zone Configurations

Gianfranco Chicco
Pietro Colella
Andrea Mazza
Angela Russo
Politecnico di Torino
Dipartimento Energia "Galileo Ferraris"
Torino, Italy
gianfranco.chicco@polito.it
pietro.colella@polito.it
andrea.mazza@polito.it
angela.russo@polito.it

Valentin Ilea
Politecnico di Milano
Dipartimento di Energia
Milano, Italy
valentin.ilea@polimi.it

Cristian Bovo
Università degli Studi di Pavia
Dipartimento di Ingegneria Industriale e dell'Informazione
Pavia, Italy
cristian.bovo@unipv.it

Federico Quaglia
Mario Limone
Terna Rete Italia Spa
Dispacciamento e Conduzione
Roma, Italy
federico.quaglia@terna.it
mario.limone@terna.it

***Abstract*— This paper focuses on model-based algorithms to be applied for identifying alternative configurations in the context of a bidding zone review process for the European Day-Ahead Electricity Market operation that adopts a zonal model to manage the transmission grid congestions. In particular, the paper presents evolved graph-based clustering algorithms designed to define bidding zones formed only by interconnected nodes. Then, a method that evaluates the performance of the resulting alternative bidding zone configurations by calculating specific performance indicators is proposed. The clustering algorithms proposed are compared with other algorithms previously developed by the authors in their application to a large-scale model and extended data set related to the Italian Transmission Network, and relevant results are presented.**

Keywords*— *electricity markets, bidding zones, locational marginal prices, model-based, graph-based clustering, performance indicators.

I. INTRODUCTION

The presence of congestions in the transmission network perturbs the optimal economic solution of the electricity markets as it introduces a local difference in the energy value among the nodes, requiring dedicated approaches to handle them [1]. *Nodal pricing* schemes consider the transmission network modelled in its entirety [2],[3]. Without congestions, the nodal prices are the same in all the nodes, while the presence of congestions results in different nodal prices (Locational Marginal Prices, LMPs). On the contrary, *zonal pricing* schemes simplify the grid model by considering the aggregation of network nodes into bidding zones (BZ) with uniform electricity prices. Ideally, there are no (structural) congestions inside bidding zones, resulting in prices very similar with the nodal prices in the nodes of the same zone as if nodal pricing would be adopted. This could lead to definition of inefficient bidding zones configurations, e.g., zones made of a few nodes or zones made of portions not directly connected. Thus, only predictable and structural congestions should be considered in the definition of the BZs.

In the European Union, the internal electricity market has been progressively implemented since 1999, increasing competition at the wholesale level. The main goals of the process are: (i) maximising social welfare; (ii) optimal integration of renewable energy sources; (iii) maximising cross-border trade; and (iv) providing competitive energy prices [4][5]. To identify the optimal market clearing solution, complex optimisation algorithms are used, e.g., for the Electricity Day-Ahead Market (DAM) [6][7][8]. The European DAM is based on *zonal pricing* [9]. The *bidding zone* is defined as the largest geographical area within which market participants can exchange energy without generating congestions. The BZs should maximise economic efficiency and the opportunities for cross-zonal trading by maintaining security of supply. The BZ borders shall be based on long-term, structural congestions in the transmission network [9].

The identification of the BZs depends on the evolution of the transmission network and of the operating conditions in time; this is a crucial factor in the context of the energy transition toward clean energy sources that sees a major evolution of the generation fleet, in terms of both energy sources and their location. It is thus necessary that the BZs configurations are continuously checked to confirm their effectiveness and revised when needed. Thus, alternative BZ configurations are studied to be considered in a BZ review process. The approaches adopted could be *expert-based*, considering the expertise of the transmission system operator (TSO), or *model-based*, in which specific algorithms are executed based on selected information on the transmission network structure and operation. The expert-based approach is more suitable to apply incremental changes to the BZs with respect to the status quo considering a large set of criteria with complex interactions among them, while the model-based approach is suitable to conduct a "greenfield" analysis.

In the European Union, a BZ review was started by the Clean Energy Package [9] in 2019. The European Union Agency for the Cooperation of Energy Regulators (ACER) was asked to propose alternative configurations to be studied, based on discussions with the TSOs and data (including LMPs) provided by TSOs [10]. The approach identified in [11] applies model-based techniques in a "brown field" solution: existing BZs are considered as starting point in an iterative approach where BZs are progressively split: in a first step BZs are ranked according to two criteria (inner variance of LMPs and amount of loop and internal flows) and, then, the worst BZ is split using clustering algorithms applied on the LMP data. The procedure is applied iteratively until specific stop criteria are met.

Model-based approaches consist of the following steps: (i) calculate, using a nodal model of the transmission network,

electric grid-specific indicators related to optimal electricity market clearing and (ii) process these indicators using specially designed clustering algorithms to find the BZs as consistent areas of the network where the specified indicators present similar values.

The authors conducted a bibliography analysis on calculation of the grid-specific indicators in [12] and on clustering algorithms for alternative BZs identification in [13]. On the indicators calculation side, it has been noticed that the adopted algorithms use a DC power flow (PF) model while rarely considering explicitly the N-1 security criteria and unit commitment (UC) of the conventional units. In [14] the authors reduced the gap with a calculation model that explicitly represented the N-1 security criteria, while in [15] relevant UC constraints were introduced and the proposed model proved its efficiency when applied on a set of extended realistic data regarding the Italian Transmission Grid. On the clustering side, to properly consider the topological constraints of the solutions, ad hoc versions of the traditional algorithms have been developed, such as the penalty-based hierarchical and K-means clustering [16]. Moreover, a probabilistic approach and a procedure to properly manage multi-scenario analysis were proposed [17]. The method allows weighting the scenarios according to their probability of occurrence and provides feasible and promising solutions.

The clustering algorithms proposed are not able to identify solutions that well respect a very important electric grid constraint: the clusters must be formed only by interconnected nodes. Therefore, in this paper, some graph-based clustering algorithms are proposed in Section II.A. Then, Section II.B proposes a clustering procedure to properly manage multi-scenario analyses, while a postprocessing procedure to merge small clusters to avoid unrealistic BZ is presented in Section II.C. In Section III a method to evaluate the performance of the resulting BZ alternative configurations is proposed, while Section IV present relevant results on the Italian Transmission Network in a multi-scenario paradigm. Finally, Section 0 shows the main conclusions.

II. Methodology to Identify Possible Bidding Zones

A. Graph-based clustering algorithms

An electrical network can be powerfully represented by graphs, where nodes are the network buses, and edges are the network branches. The graph-based clustering algorithms combine the principles of traditional clustering algorithms with graph theory. These algorithms are promising for the BZ formation problem, since the topological constraints mentioned in Section II are intrinsically considered. In this paper, two graph-based clustering algorithms are implemented: graph-based k-means and graph-based hierarchical clustering. The implementation of this algorithm is described below.

1) Graph-based k-means

The implemented version of the Graph-based K-means (GKM) clustering algorithm requires three steps: the seed identification, the nodal initialization, and the changing phase. These phases are explained below.

First, the algorithm identifies the initial cluster seeds. Considering the desired number of clusters *K,* and the frequency distribution of the input feature (e.g., LMP), *K* equally spaced percentiles are computed. The cluster seeds are the nodes characterized by a feature closest to the computed *K* percentiles. For example, if $K = 5$ and the input nodal feature is the LMP, the nodes with an LMP closest to the 0^{th}, 25^{th}, 50^{th}, $75^{th,}$ and 100^{th} percentiles are identified as the seeds of the 5 clusters.

Secondly, all the nodes are associated with a cluster, considering their feature value and the graph topology, through an iterative process. The cluster centroids are computed as the mean value of the features of the nodes that belong to each cluster. The nodes with a direct link to clusters and not yet associated are identified. The pairwise Euclidean distance between their features and the cluster centroids is computed. The node with the lowest Euclidean distance is joined to the corresponding cluster. The cluster centroids are updated. The process is repeated until all the nodes are assigned to a cluster.

Thirdly, the borders of the clusters are moved to have better results. The algorithm identifies the nodes on the border between two adjacent clusters and checks if there is any with a feature value closer to a centroid of a cluster different from the one at which it belongs. The algorithm verifies that the change does not create a number of clusters greater than the desired one and that all the buses that belong to the same cluster are interconnected.

2) Graph-based Hierarchical Clustering

Graph-based Hierarchical Clustering (GHC) is a method that merges the principles of hierarchical clustering with graph theory to group data points based on their inherent structures. The algorithm begins by treating each data point as its own individual cluster. The algorithm repeatedly merges the two interconnected most similar clusters. where there is at least one direct link. This similarity can be defined using various linkage criteria, such as single-linkage (minimum distance between points in different clusters), complete-linkage (maximum distance), or average-linkage (average distance). The merging process continues until all data points are grouped into a unique one and can be stopped at the desired number of clusters.

B. Clustering procedure in multi-scenario analysis

The operational states of an electrical system can be several, based for example on the different electrical demands, the generation power plants and electrical lines in service, and the amount of electrical energy from Renewable Energy Sources (RESs). Therefore, several congestion scenarios have to be simulated to obtain representative features (such as the LMPs) that can be used as inputs for the clustering process. For this reason, in previous papers, the authors have presented an ad-hoc procedure to effectively manage multiple scenarios, based on the consensus clustering algorithm [17].

In the first step of the process, a clustering algorithm is run to identify *K* clusters for each scenario considered singularly.

Then, in the second step, a similarity matrix $\mathbf{S} = \{s_{i,j}\} \in \mathbb{R}^{(M,M)}$ is formed. The similarity index $s_{i,j}$ for the nodes *i* and *j* is in the range (0,1) and represents the number of scenarios where the nodes are in the same cluster over the total number of scenarios analysed.

Finally, the matrix $\mathbf{S}$ is used as input for the spectral clustering algorithm that organized the *M* buses into the final *K* clusters. In this paper, the procedure has been improved, so that also the graph-based algorithm described in Section II-A can be executed.

C. Merging small clusters

After finding a solution with one of the specified algorithms, it is necessary to check that the clusters include a sufficient number of nodes. Indeed, it would not be appropriate to propose a BZ that is too small geographically, although the features considered in the input lead to this conclusion. For this reason, clusters with a number of nodes below a user-defined threshold are merged with the adjacent cluster characterized by the nearest centroid. In our case the threshold of 30 nodes is used, as a viable compromise between the two conflicting needs of having not-too-small zones and avoiding over-conditioning the clustering algorithm.

III. Clustering Solution Assessment

For each alternative BZ configurations, DAM sessions are run using the regulatory framework of the specific DAM and the market data employed to calculate the inputs of the clustering algorithms, i.e., the economic bids of suppliers and traders. The performance of the alternative BZ configurations can be then quantified by market efficiency indicators.

The case of Italy is considered where the DAM uses a *zonal price* mechanism with the BZs commercial exchanges managed in an ATC paradigm. This is suitable due to the structure of Italian Transmission Grid that allows, traditionally when expert-based approaches are used, the definition of radial BZ configurations. However, the alternative BZ configurations proposed by clustering algorithms can be meshed, resulting very difficult to manage in an ATC approach due to loop-flows. In these circumstances, for greater accuracy, the Flow-Based (FB) paradigm [18], largely used in the Integrated European DAM, is used and a methodology and mathematical models that follow closely the regulatory framework on capacity management [19] and clearing process [1] in European FB DAM are defined to evaluate the clustering solutions:

Step 1: *Definition of FB domain*. To identify the set of Critical Network Elements with a Contingency (CNECs) and their parameters (zonal PTDFs, Remaining Available Margin - RAM) it is necessary to have a good estimation of the DAM solution in terms of generation and demand dispatch, corresponding physical flows on the network branches and optimized Non-Costly Remedial Actions (NRA). The nodal market model presented in [15] and used to calculate the network indicators for clustering algorithms is adopted for this purpose with some differences: (i) the costly corrective actions (generation redispatch actions) are neglected; (ii) only the N-1 preventive security is thus considered; (iii) the Phase-Shifting Transformers' (PSTs) angles are the only NRAs considered and act preventively, and (iv) only the branches crossing two BZ or with a significant contribution to the power exchanges between BZs (present at least a zone-to-zone PTDF higher or equal to 5%) [15] and with flows above 60% of the maximum limit are considered as critical branches.

From the solution of the problem, for every time window $t \in \boldsymbol{T}$, where $\boldsymbol{T}$ is the total number of time windows in the market session, the following results are retained: (i) the nodal power injections i_b, where $b \in \boldsymbol{B}$ is a bus of the network and $\boldsymbol{B}$ is the set of all buses; (ii) the PST angles $\varphi_{f,t}$, where $f \in \boldsymbol{PST}$ is a PST of the network and $\boldsymbol{PST}$ is the set of all PSTs and the set $\boldsymbol{CNEC}$ of the identified CNECs. The Generation Shift Key (GSK) of bus b in BZ a, GSK_b^a, is calculated as:

$$GSK_{b,t}^{a} = i_{b,t} / \sum_{m \in \boldsymbol{B_a}} i_{m,t} \quad (1)$$

where $\boldsymbol{B_a}$ is the set of buses of BZ a.

The PTDF of CNEC k with respect to variation in the net position of BZ a, $PTDFz_k^a$, is calculated as:

$$PTDFz_{k,t}^{a} = \sum_{b \in \boldsymbol{B_a}} PTDF_{k,t}^{b} \cdot GSK_{b,t}^{a} \quad (2)$$

where $PTDF_k^b$ is the PTDF of CNEC k with respect to the injection at bus b.

The RAM of CNEC k, $RAM_{k,t}$, is calculated as:

$$RAM_{k,t} = F_k^{\lim} - \sum_{f \in \boldsymbol{PST}} s_{k,t}^{f} \cdot \varphi_{f,t} \quad (3)$$

where $F_k^{\lim}$ is the flow limit of CNEC k and $s_{k,t}^f$ is the sensitivity of the flow in CNEC k at the change of the angle of PST f at time window t.

Step 2: *FB Market Model*. Considering the Italian market rules where simple step-wise bids are presented and hypothesising the demand inelastic, the adopted model is:

$$\min \sum_{t \in \boldsymbol{T}} \sum_{p \in \boldsymbol{UP}} \sum_{s \in \boldsymbol{S}} \pi_{p,t}^{s} \cdot P_{p,t}^{s} \quad (4)$$

such that

$$P_{p,t}^{UP} = \sum_{s \in \boldsymbol{S}} P_{p,t}^{s}, \forall p \in \boldsymbol{UP}, \forall t \in \boldsymbol{T} \quad (5)$$

$$y_{a,t} = \sum_{p \in \boldsymbol{UP_a}} P_{p,t}^{UP} - \sum_{b \in \boldsymbol{B_a}} P_{D,b,t}, \forall a \in \boldsymbol{BZ}, \forall t \in \boldsymbol{T} \quad (6)$$

$$\sum_{k \in \boldsymbol{UP}} P_{k,t}^{UP} - \sum_{b \in \boldsymbol{B}} P_{D,b,t} = 0, \forall t \in \boldsymbol{T} \quad (7)$$

$$F_{k,t} = \sum_{b \in \boldsymbol{B}} PTDFz_{k,t}^{a} \cdot y_{a,t}, \forall k \in \boldsymbol{CNEC}, \forall t \in \boldsymbol{T} \quad (8)$$

$$F_{k,t} \leq RAM_{k,t}, \forall m \in \boldsymbol{CNEC}, \forall t \in \boldsymbol{T} \quad (9)$$

$$F_{BS,t}^{min} \leq \sum_{m \in \boldsymbol{BS}} F_{m,t} \leq F_{BS,t}^{max}, \forall t \in \boldsymbol{T} \quad (10)$$

where $\boldsymbol{UP}$ is the set of production units (UP); $\boldsymbol{UP_a}$ is the sub-set of production units injecting power into BZ a; $\boldsymbol{S} = \{1,2,3\}$ is the set of supply bids steps of an UP; $\boldsymbol{BS}$ is the set of CNECs subject to a branch-set constraint on the total flow through an interface; $\pi_{p,t}^s$ is the selling price of supply bid step s by UP p at time window t; $P_{p,t}^s$ is the accepted quantity in production of supply bid step s by UP p at time window t; $P_{p,t}^{UP}$ is the power produced by UP p at time window t; $P_{D,b,t}$ is the power demanded by traders at bus b, at time window t; $y_{a,t}$ is the net position (export) of BZ a, at time window t; $F_{k,t}$ is the flow in CNEC k at time window t; $F_{BS,t}^{max/min}$ are the flow limits of the considered branch-set $\boldsymbol{BS}$.

Considering the demand inelastic, Eq. (4) minimizes the production costs instead of maximizing the social welfare. Eq. (5) defines the production of UP p; (6) defines the net position of BZ a; while (7) imposes the power balance at the level of the system. Eq. (8) defines the flow on CNEC k as a linear combination of BZ net positions weighted by the zonal PTDFs, while (9) constrain them to their thermal limits. Eq. (10) limits the total flow on a specified interface to its bounds; for the Italian case this is necessary to represent the voltage issues on the Central North – Central South interface [15].

Additionally, model (4)-(10) includes the constraints related to the operation of a basin hydroelectric UP, detailed in [15]. Even if the rules of the Italian DAM do not allow hydroelectric constraints within the model, they have been considered here to quantify the strategic behavior (connected to the reservoir capacity) of such plants in the market.

Step 3: *Performance indicators*. From the solution of the FB DAM, performance indicators of the market, hence of the evaluated BZ configurations, are calculated. One indicator is

the Social Welfare (SW) that quantifies globally the benefit of both suppliers and traders, and which maximization is the main goal of the market. Here the demand is inelastic, so the maximization of the SW becomes the minimization of the total production cost of the market; thus, the smaller this number, the more efficient the market. Another indicator is the total Congestion Rent (CR) of the market, i.e., the financial surplus of the market resulting from different zonal prices in presence of CNEC congestions. The CR for each CNEC is calculated as the product between the Lagrange multiplier of the flow limit constraint (10) and the value of the CNEC flow; the total CR is obtained as the CR sum over all CNECs. Ideally, the grid infrastructure should allow the best economic solution without CNEC congestions thus, the CR should be null; in reality, structural congestions are difficult to avoid, hence, the smaller the CR, the more efficient the market. Finally, the CNEC flows calculated by the DAM clearing model are not exact values due to the approximations introduced by the zonal model (e.g., a relevant error is introduced by (1), (2) if the predicted FB DAM solution differs, even slightly, by the actual solution) so residual congestions can appear when the FB DAM solution is deployed to the nodal model – solving these congestions in an Ancillary Services Market (ASM) stage can have high costs. Moreover, due to the filtering criteria of the CNEC (the 5% rule) some "local" congestions may not be considered by the FB DAM. Thus, a third performance indicator is the total quantity of residual congestions (intended as the sum of MWh above the considered limits of all branches in the nodal model) to be mitigated by ASM. This indicator was chosen instead of a financial one (e.g., cost of mitigating the residual congestions) due to the difficulty in estimating the ASM prices over long periods of time as they present high volatility.

IV. Application to the Italian Case Study

A. Input data

The network considered is the Italian High Voltage transmission grid described in detail in [15]. The year 2019 is considered and each day is reduced to six significant hours. This data set is extended under a Monte Carlo approach by replicating the year many times and considering unprogrammed faults (UF) and scheduled maintenances (SM) of the electric branches and UPs: the number of UF and SM in the considered period is randomly generated using the *Poisson distribution* and the *mean number of UF/SM per year* (μ_P) as parameter. The duration of the UF/SM event is randomly generated using the *Gamma distribution* and the *mean and standard deviation of the event duration* (μ_γ and σ_γ) as parameters. The moment of the event is chosen randomly and iteratively in the time slots still not occupied by previous events until the list of generated events is exhausted. Based on the Italian TSO experience, the parameters used for electric lines are $\mu_{P,UF} = 0.05$ UF/year $\mu_{\gamma,UF} = 1.7$ day/year $\sigma_{\gamma,UF} = 1.2$ day/year, $\mu_{P,SM} = 1$ UF/year, $\mu_{\gamma,SM} = 3.1$ day/year $\sigma_{\gamma,SM} = 3.4$ day/year while the parameters used for UPs are $\mu_{P,UF} = 1$ UF/year $\mu_{\gamma,UF} = 1$ day/year $\sigma_{\gamma,UF} = 6$ day/year, $\mu_{P,SM} = 1$ UF/year, $\mu_{\gamma,SM} = 14$ day/year $\sigma_{\gamma,SM} = 7$ day/year. Twenty years of data were generated and the nodal market algorithm of [15] was deployed to calculate the nodal indicators.

B. Clustering Solutions

The clustering solutions are presented in Fig. 1 for five clusters, chosen arbitrarily as an example. As input data of the graph-based K-means and hierarchical clustering algorithms, the column array $\mathbf{x} = \{x_i\} \in \mathbb{R}^{\{M,1\}}$ is provided, where x_i is the mean value of the LMP for node i. For the consensus clustering, each hourly scenario is instead considered in step 1, where the K-means or hierarchical clustering algorithm is executed; the desired number of clusters is obtained in step 2 by the spectral clustering. Considering the results in Fig. 1, the graph-based versions of the algorithms provide more reasonable solutions than the classical ones. With the classical K-means and hierarchical clustering, some clusters are formed by a few number of nodes (Fig. 1-a, Fig. 1-b) or by nodes not interconnected (Fig. 1-b).

Some similarities can be detected. As expected, Sicily is often separated from the Calabria borders. Two exceptions are the solution proposed by graph-based K-means (Fig. 1-a) and the consensus where the hierarchical algorithm was executed at step 1 (Fig. 1-f). In the first case, the nodes in Sicily are merged into a unique cluster with some nodes of the continent. In the second one, some Sicilian nodes are merged with the Calabria ones. Another recurring result is a cluster border in the centre of Italy. It is important to emphasize that in the graph-based hierarchical clustering (Fig. 1-c) and the consensus clustering algorithms (Fig. 1-e, Fig. 1-f), the North and central-northern Italy are merged into a unique cluster.

An uncommon solution is proposed by the graph-based K-means, where most of the Italian East coast nodes are grouped. The solutions proposed by the two consensus clustering algorithms (Fig. 1-e and Fig. 1-f) are similar, except for the South and central-southern Italy. In the first case, a cluster in the Campania area is identified, but Puglia and Calabria are merged. In the second one, a bigger cluster for the central-southern is proposed, while Puglia and Calabria are split into two distinct clusters.

C. Assessment of the clustering solutions

The procedure in Section III was run for all the BZ configurations shown in Section IV-B to calculate the performance indicators. The BZ configuration given by the hierarchical clustering did not converge due to the single-bus BZ (red in Fig. 1-d) that in many hours has a null net injection ($i_{b,t} = 0$) and its GSK and PTDF, cannot be calculated; thus, results for this configuration are not presented further.

Fig. 2 shows the total generation cost. The graph-based k-means presents the best value, while the graph-based hierarchical the worst. However, given the values' scale, all methods provide very similar results. Indeed, the corresponding unitary price (total cost / total volume traded) varies very little among configurations, i.e., 57.47 ÷ 58.19 €/MWh, so the indicator does not clearly indicate the best one.

Fig. 3 shows the congestion rent. Again, the graph-based k-means presents the best value, while k-means the worst, followed closely by the others. The results could suggest an optimal efficiency in managing the structural congestions by the graph-based k-means algorithm, but it could also mean the opposite: the structural congestions are seldomly detected.

Indeed, Fig. 4, which shows the total quantity of residual congestions, confirms the latter hypothesis: the graph-based k-means has the highest value and the k-means algorithm the second highest. In general, the numbers of Fig. 4 and the differences between the lowest and highest values are very high; considering also that these quantities need to be mitigated by re-dispatching actions in the Ancillary Services Markets where the prices are very high, this indicator may give the best idea of clustering algorithms performance.

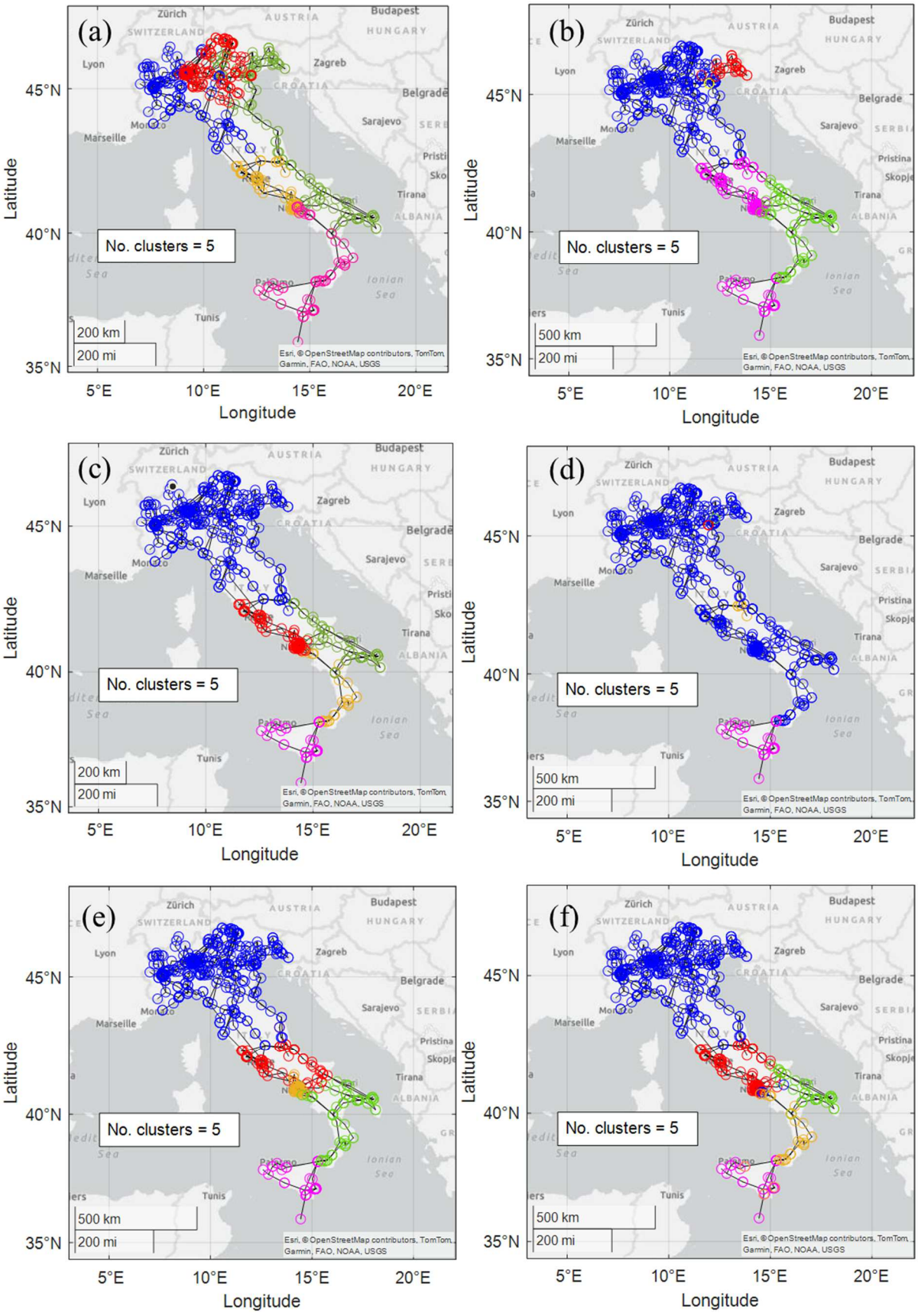


Fig. 1. Clustering solution obtained by: a) graph-based k-means; b) k-means; c) graph-based hierarchical; d) hierarchical; e) consensus (step 1: k-means, step 2: spectral); f) consensus (step 1: hierarchical, step 2: spectral)

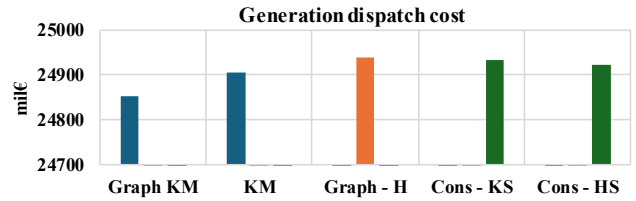


Fig. 2. Generation dispatch cost indicator

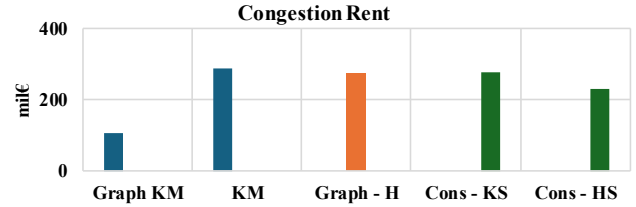


Fig. 3. Congestion rent indicator

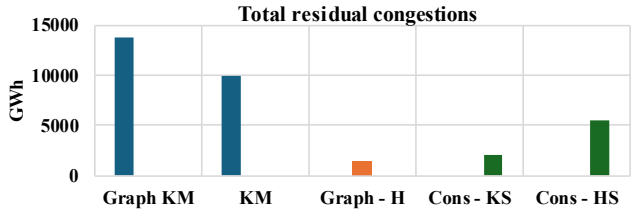


Fig. 4. Total quantity of residual congestions indicator

A CNEC well represented in FB DAM implies accurate zonal PTDFs, which occurs if BZs are homogeneous and, ideally, contain all the nodes with similar nodal PTDFs (thus, nodal injections change the flow similarly). This is not the case for graph-based k-means: (i) the BZs in the North part (see Fig. 1-a) have no clearly defined borders but are mixed and (ii) the CNECs in Sicily area are on the natural border between Sicily and the continent, manifest in N-1 conditions when connect the island radially; this means, with the slack node near the border with France, that the nodal PTDFs of the island are unitary while the other null. Thus, including buses with high nodal injections from the continent (high-capacity UPs or high demand) will give very inaccurate representation of CNEC in the zonal model due to a zonal PTDF of Sicily area very far from unitary value (the theoretical value). The k-means algorithm does not present the first problem, but Sicily BZ is again very badly represented (see Fig. 1-b). The graph-based hierarchical (Fig. 1-c) and the first consensus algorithm (Fig. 1-e) don't show these issues: the values of the indicator are similar and very low; the best is the graph-based hierarchical which provides, visually, the most balanced BZ alternative configuration. For the second consensus algorithm (Fig. 1-f) the value of the indicator is not good due to merging some Calabria BZ nodes (yellow) within the Sicilian BZ.

V. Conclusion

To assist the bidding zone review process, this paper proposed specially tailored graph-based clustering algorithms to be applied to economic nodal indicators, e.g., LMP, and over a large data set to obtain alternative bidding zones configurations. The paper also proposed a methodology to calculate performance indicators to evaluate the results. Tests on the Italian High Voltage Transmission grid have compared the various graph-based clustering algorithms with others, more standard, and showed the better efficiency of the later, with the best being the graph-based hierarchical algorithm. The developed methodology is also suitable for highly meshed and larger power systems. Finally, it is worth mentioning that the proposed configurations could imply a change in BZ borders and even an AC meshed zonal configuration, resulting in the necessity of updating cross-zonal limits or even capacity calculation methodologies.

References

[1] S. Stoft, *Power System Economics*: Designing Markets For Electricity, Wiley, NY, 2015.

[2] P. Cramton, "Electricity Market Design", Oxford Review of Economic Policy, Volume 33, Number 4, 2017, pp. 589–612.

[3] F.A. Wolak, "Wholesale electricity market design", Handbook on Electricity Markets, Chapter 4 (2021): 73-110.

[4] L.H. Lam, V. Ilea, and C. Bovo, "Impact of the price coupling of regions project on the day-ahead electricity market in Italy", Proceedings of the 2017 IEEE Manchester Power Tech, Manchester, UK, 18–22 June 2017, pp. 1–6.

[5] C. Graf, F. Quaglia, F.A. Wolak, "Simplified electricity market models with significant intermittent renewable capacity: Evidence from Italy", no. w27262. National Bureau of Economic Research, 2020.

[6] L.H. Lam, V. Ilea, and C. Bovo, "A thorough comparison among various mathematical approaches to compute PUN in Italy", Proc- of the 2018 15th International Conference on the European Energy Markets (EEM), IEEE, Lodz, Poland, 27–29 June 2018; pp. 1–5.

[7] L.H. Lam, V. Ilea, and C. Bovo, "European day-ahead electricity market coupling: Discussion, modeling, and case study", Electr. Power Syst. Res., vol. 155, pp. 80–92, Supplement C., 2018.

[8] L.H. Lam, V. Ilea, and C. Bovo, "New clearing model to mitigate the non-convexity in european day-ahead electricity market", Energies, Volume 13, Issue 18, Article number 4716, September, 2020.

[9] Regulation (EU) 2019/943 of the European Parliament and of the Council of 5 June 2019 on the internal market for electricity (Electricity Market Regulation), Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32019R0943&from=EN.

[10] ENTSO-E, "Report on the Locational Marginal Pricing Study of the Bidding Zone Review Process", 30 June 2022, https://eepublicdownloads.entsoe.eu/clean-documents/Publications/Market%20Committee%20publications/ENTSO-E%20LMP%20Report_publication.pdf

[11] ACER, Decision no. 11/2022 of the European Union Agency for the Cooperation of Energy Regulators on the alternative bidding zone configurations to be considered in the bidding zone review process, Ljubljana, Slovenia, 8 August 2022.

[12] C. Bovo et al., "Review of the Mathematic Models to Calculate the Network Indicators to Define the Bidding Zones", Proceedings of the 2019 54th International Universities Power Engineering Conference, UPEC 2019, Bucharest, Romania, 3–6 September 2019.

[13] G. Chicco et al., "Overview of the clustering algorithms for the formation of the bidding zones", Proc. 54th International Universities Power Engineering Conference (UPEC 2019), Bucharest, Romania, 3-6 September 2019.

[14] C. Bovo et al., "Optimal Computation of Network Indicators for Electricity Market Bidding Zones Configuration Considering Explicit N-1 Security Constraints", Energies, vol. 14, art. 4267, 2021.

[15] F. Quaglia et al., "Advanced Model-based Approaches to be Applied in the Context of a Bidding Zone Review", Proc. 2023 115th AEIT International Annual Conference (AEIT 2023), Rome, 5-7 October 2023.

[16] P. Colella et al., "Model-based identification of alternative Bidding Zone Configurations from Clustering Algorithms applied on Locational Marginal Prices," in 2020 55th International Universities Power Engineering Conference (UPEC), 2020, pp. 1–6.

[17] P. Colella et al., "Model-Based Identification of Alternative Bidding Zones: Applications of Clustering Algorithms with Topology Constraints," Energies, vol. 14, no. 10, p. 2763, 2021.

[18] K. Van den Bergh, J. Boury, E. Delarue, "The Flow-Based Market Coupling in Central Western Europe: Concepts and definitions", The Electricity Journal, vol. 29, pp. 24-29, 2016.

[19] ACER, "Core CCR TSOs' regional design of the day-ahead common capacity calculation methodology in accordance with Article 20ff. of Commission Regulation (EU) 2015/1222 of 24 July 2015", June, 2018.

[20] Epex Spot, "PCR Market Coupling Algorithm", 2016. Available online: https://www.epexspot.com/en/market-coupling/pc