RESEARCH   **Open Access**

A practical approach to cluster validation in the energy sector

Alexander Bogensperger[1,2]* and Yann Fabel[1]

*Correspondence:
abogensperger@ffe.de
[1]Forschungsstelle fuer
Energiewirtschaft e.V., Am
Bluetenanger 71, 80995 Munich,
Germany
[2]TUM Graduate School, Technical
University of Munich, Boltzmannstr.
17, 85748 Garching, Munich,
Germany

Abstract

With increasing digitization, new opportunities emerge concerning the availability and use of data in the energy sector. A comprehensive literature review shows an abundance in available unsupervised clustering algorithms as well as internal, relative and external cluster validation indices (cvi) to evaluate the results. Yet, the comparison of different clustering results on the same dataset, executed with different algorithms and a specific practical goal in mind still proves scientifically challenging. A large variety of cvi are described and consolidated in commonly used composite indices (e.g. Davies-Bouldin-Index, silhouette-Index, Dunn-Index). Previous works show the challenges surrounding these composite indices since they serve a generalized cluster quality evaluation. However, this does not suit individual clustering goals in many cases. The presented paper introduces the current state of science, existing cluster validation indices and proposes a practical method to combine them to an individual composite index, using Multi Criteria Decision Analysis (mcda). The methodology is applied on two energy economic use cases for clustering load profiles of bidirectional electric vehicles and municipalities.

Keywords: Unsupervised machine learning, Clustering, Cluster validation, MCDA, Energy

Introduction

With increasing amounts of data in the energy sector, the relevance of data analysis is increasing constantly. This is mainly caused by the rising numbers of smart meters and decentralized energy resources (DER) as well as sensors and actors in infrastructures and new assets (i.e., through sector coupling). This trend is causing a growing complexity in handling incoming data, purposefully utilizing it and managing the complexity of the system. This paper focuses on the utilization of data with a given goal in mind. In contrast to exploratory data analysis, the examination of unknown datasets is conducted with certain pre-conceived presumptions to identify new information, patterns and derive hypotheses concerning the individual research goals (Martinez et al. 2010; Tukey 1977). Especially now, in the early stages of the digitization of the energy industry, with newly available data

and tools, the importance of data analysis must not be overlooked. Unsupervised learning extends or simplifies this process and therefore gains an increasing practical importance within the industry. Especially with newly acquired data it bears many advantages such as

- the compression of information (reducing information complexity),
- simplification of complex and high-dimensional data,
- pattern recognition,
- the detection of outliers,
- knowledge expansion and an increased understanding of the data (Tanwar et al. 2015; Brickey et al. 2010).

Yet, while unsupervised learning becoming progressively more convenient with many available libraries, the process of data analysis with real world data remains a big challenge. The process of deriving the desired information out of specific datasets is highly individual and scientifically challenging. The extraction of valid clustering results, serving specific goals e.g., of a client or for a given real-world task is especially highly individual (Hennig 2020). The main research goals of this paper include the review and development of existing relative and internal cluster validation methodologies to compare different model results. Furthermore, an emphasis is put on the practical application of the methodology outlined in Hennig (2020) to build a bridge between experts in certain fields (here: energy economics) with machine-learning and data science experts. The resulting methodology is applied to energy-economic datasets in two different projects.

Literature review

The goal of this paper is to identify clusters for a given dataset without any prior knowledge about its structure but with certain goals in mind. The fact that countless clustering algorithms are available and easily accessible raises the challenge of identifying the individually best clustering result for a certain task and dataset. According to Rendón et al. (2011), there are three ways to evaluate the results of unsupervised clustering analysis to find the "best" clustering:

1 relative validation is used to tune the hyperparameters of an algorithm (i. e., number of clusters) to identify the best model. These relative validation methods may vary according to the machine learning algorithm used. One commonly used relative validation method is the elbow curve, used in conjunction with k-means (Syakur et al. 2018).
2 internal validation describes the identified clusters within a dataset by different algorithms and compares them.
3 external validation compares the clustering results to the ground truth and describes the error via selected indices.

The goal of this paper is to develop a practical methodology to identify the best clustering result out of a finite number of runs by applying different algorithms and varying hyperparameters on the same dataset. While options one and two are necessary to determine the optimal hyperparameters for a chosen algorithm (1) and to determine the "best" algorithm (2), option three is beyond the scope of this paper due to the lack of a ground truth. As stated in Hennig (2015); Hennig (2020); Hennig and Liao (2010); Metwalli (2020) and many more, there is neither a universally optimal clustering method nor a generally

applicable definition of a cluster. This is supported by the multitude of different algo-rithms described in literature, each having specific goals, strengths, and weaknesses in terms of clustering results, scaling and ease of use on different datasets. Selecting the indi-vidually best suited algorithm and comparing their results hence pose a challenge which is often overcome in a pragmatic approach, considering the size of the dataset, available computing power, ease of use of the algorithms or just personal preference. The first step to a scientifically viable clustering is to find a general or individual definition of a cluster, which is done in the following by a literature review.

Definition of clustering and clusters

Clustering can be described in a very general sense as a "method of creating groups of objects, or clusters in such a way that objects in one cluster are very similar and objects in different clusters are quite distinct" (Gan et al. 2007). More detailed definitions of clus-tering always use "metrics" to describe their goals, as shown in the definitions in Gan et al. (2007) by Bock (1989) and Carmichael et al. (1968). The authors describe objects in a cluster as closely related in terms of their properties with high mutual similarities (= low distances) and other objects out of the same cluster in close proximity. All clusters in a dataset should be clearly distinguishable, connected and dense areas in n-dimensional space and surrounded by areas of low density in n-dimensional space. These definitions show that, with a greater level of detail, the definitions of clusters vary strongly and might even be contradicting. It also shows that assumptions about the clusters have to be made in order to find a clustering result. Lorr (1983) proposed splitting clusters into two groups, as summarized in Lorr (1983):

- compact clusters have high similarity and can be represented by a single point or a center.
- "chained cluster is a set of datapoints in which every member is more like other members in the cluster than other datapoints not in the cluster" (Gan et al. 2007).

The challenge is either to find out the types of clusters that are present in a given dataset or find clusters that best match certain criteria (as seen in chapter "Application on energy economic use cases"). Yet with increasing usability and research in the field of data sci-ence and clustering algorithms, the number of easy-to-use algorithms is rising steeply. This is a challenge, as it makes it more difficult to choose the right algorithm, tune hyper-parameters, and choose the best result. The following chapters outline a methodology to overcome these challenges and use it with different real-world datasets.

Methodologies to identify the best clustering algorithm

Papers comparing different clustering algorithms (=relative validation) to identify a "best" solution usually do so to propose and validate new algorithms utilizing known datasets and a known ground truth (e.g., Hennig (2015); McInnes et al. (2017); Chen (2015); Kuwil et al. (2019); Das et al. (2008); Cai et al. (2020)) . Only very few of them utilize generalized metrics to compare the results and are completely unbiased (Hennig 2015). More general and axiomatic approaches characterizing clustering algorithms can be found in Acker-man and Ben-David (2009) , responding to Kleinberg (2002). Ackerman and Ben-David (2009) proposes a methodology to define cluster quality functions, individual goals for

these functions and then optimize towards it. A comprehensive connection between clustering goals, the structure of the datasets, clustering methods, and validation criteria can also be found in the works of Hennig et. al. (see Hennig (2020); Hennig (2015); Hennig and Liao (2010)). Hennig (2015) proposes a methodology to identify the optimal clustering algorithm for individual datasets. The paper focuses on pre-processing as well as the clustering itself. The choice of representation and measure of dissimilarity advocates for the attitude that correlating features should also be included in a dataset if they are essential for clustering and shows different ways to incorporate clustering in non-Euclidean space with different data types. The authors propose different (and optional) ways to transform features with nonlinear functions to influence the effect of distance measures and resulting gaps between datapoints within a feature. This helps to avoid unwanted effects of outliers in the dataset. Hennig (2015) Different methods of standardization, weighting and sphering of variables are further discussed. The authors highlight the impact of outliers on these methods and the effect of these methods on clustering results due to a (possibly even wanted) change of feature variance and refer to paper supporting these claims.

All in all, literature provides a wide range of internal and relative validation indices, suitable for clustering. Yet only a few sources focus on a more axiomatic approach to selecting the best clustering results purely based on a large range of validation indices. Hennig et al. 2020 provide a comprising methodology to standardize these indices to compare them (see chapter "Relative and internal cluster validation indices"). Kou et al. (2014) proposes a methodology for multiple criteria decision-making to select the best ensemble of validation criteria, interpretability, computation complexity and visualization for a specific challenge in financial risk analysis. Tomasini et al. (2016) propose a methodology using a regression model to determine "the most suitable cluster validation internal index.

Relative and internal cluster validation indices

To evaluate and compare different clustering results, a set of validation indices is required to benchmark the results of different algorithms (relative validation) or varying hyperparameters (internal validation). Thus, papers utilizing cluster validation indices (cvi) for relative or internal validation are introduced in the following. Puzicha et al. (2000) propose different separability measures based on clustering axioms. Cormos et al. (2020) focuses on internal validation criteria (sum of square error, scatter criteria, trace criteria, determinant criteria, invariant criteria) for large and semi-structured data as well as the performance of selected algorithms. Rendón et al. (2011) apply k-means and bisecting k-means with a variety of internal and external validation indices. All of them are composite indices, combining multiple validation indices into one generalized index. They include the commonly used Calinski-Harabasz-Index, Davies-Bouldin-Index, silhouette-Coefficient, Dunn-Index as well as a novel validity index (NIVA) (Rendón et al. 2008). This is also a common procedure in many energy related works. E.g. Yang et al. (2017) rely on the use of multiple composite indices (such as Calinski-Harabasz-Index, Davies-Bouldin-Index, silhouette-Coefficient, Dunn-Index) to detect building energy usage patterns using k-shape clustering. Proving their results with a known ground truth (external validation). Zhou et al. (2017) introduce a (fuzzy) cluster based model to identify patterns in monthly electricity consumption of households. They remark that no single cvi is always the best or performs best on any given dataset, datatypes or distance-measure. Hence, they apply

the COS index (composite index), they already used in previous works. It is comprised of a compactness, separation and overlapping indicator. Gheorghe et al. (2015) create representative zones to assess the renewable energy potential in Romania by using k-means. They validate their results internally with various indices related to the silhouette-index.

Akhanli and Hennig (2020) introduce two new composite indices to describe cluster homogeneity and cluster separation. Other internal validation indices can be found in Liu et al. (2010) and Vendramin et al. (2010). Kou et al. (2014) utilizes F-measure, normalized mutual information purity and entropy. Chou et al. (2002) introduce a point symmetry measure as a cluster validity measure. Wang et al. 2019 create a new composite index (Peak Weight Index) out of two composite indices (silhouette index and Calinski-Harabasz index). Many papers with practical relevance, including the field of energy and energy economics, utilize clustering techniques usually by applying only one clustering algorithm (e. g. Bittel et al. (2017); Siala and Mahfouz (2019)). If multiple algorithms are compared, generalized composite indices (e.g., Davies-Bouldin-Index, silhouette-index etc.) or a selected few indices such as sum of squared errors are used (Toussaint and Moodley; Schütz et al. 2018).

This overview shows the lack of scientific discussion of the comparison of different algorithms, especially in subject-specific scientific papers. Many scientific papers use one or multiple (composite) cvi, usually not providing much insights in the selection process or alternatives. A critical review or deeper analysis of the used index/indices is usually missing. This poses a risk since validating cluster results with different cvi on the same data set often produces very different results.

In Hennig (2020), Hennig et al. introduce different cluster validity indices (cvi) including their mathematical formulation and a suitable normalization. These cvi are normalized in such a way that 1 represents the best (possible) value and 0 the worst. An overview of these indices is given in Table 1.

Table 1 List of cluster validation indices used in this work

| Name	Abbreviation	Usage
Average within-cluster distance	I_{avg_wc}	Measure of similarity of objects/points in a cluster. The higher the index, the smaller the average within-cluster distance.
p-separation-index	I_{p-sep}	Measure of separation between clusters. Instead of minimum/maximum distance (prone to outliers) this can be calculated by the mean of a portion (p) between two clusters. The higher the index, the better the between-cluster separation.
Representation by centroids	$I_{centroid}$	Measure of how well a cluster is represented by its centroid. The higher the index, the better the representation.
Representation of dissimilarity structure by clustering	$I_{pearson}$	Measure of the dissimilarity structure denoted by the Pearson correlation between pairwise dissimilarities (e.g., Euclidean distances) and "clustering induced dissimilarity" (matching cluster). For increasing dissimilarity, objects/points should not be assigned to the same cluster. Hence for higher indices, pairwise dissimilarity correlates more strongly to clustering dissimilarity.
Within-cluster gaps	$I_{widestgap}$	Measure of the connectivity of a cluster. The higher the index, the smaller the within-cluster gaps.
Entropy	$I_{entropy}$	Measure for assessing the uniform size of clusters.
Parsimony	$I_{parsimony}$	Measure to express the preference for a lower number of clusters.
Density modes and valleys	$I_{densdec}$	Measure to quantify the density drop from cluster-mode to the edges of a cluster and the density-valleys between clusters.
Uniform within-cluster density	I_{cvdens}	Measure to quantify the within-cluster density levels. For higher indices, density is more uniform within the cluster.

Hennig (2015) shows the inherent clustering characteristics and tendencies of selected groups of algorithms (partially see chapter 4.3). It further proposes using different validation indices such as measurements of within-cluster homogeneity, cluster separation, homogeneity of different clusters, and measurements of fit, e.g., to a centroid. The author points out the importance of the stability of clustering (i. e. the influence of changes in the dataset on the clustering results). Generally, two types of indices can be distinguished. Simple validation indices (in analogy to cryptography one might call them primitive cluster validation indices) as shown above and composite indices. Composite indices (like the silhouette-coefficient) are not composed of a single cvi but combine multiple of them into one to create a measure of cluster quality. This measure might not suit every purpose well and rather aims for a more generalized approach. Hennig (2020) This paper will utilize the primitive indices over composite indices and create a task-specific composite index according to the clustering goal.

The literature review shows multiple challenges in the field of clustering. The number of available and easy-to-implement clustering algorithms increases steadily while mitigating certain weak points of the existing methods. This increases the difficulty of choosing the best algorithms for a given task. Evaluation metrics are manifold in different papers, a comprehensive overview and normalization to compare them is given in (Halkidi et al. 2016). The reviewed research also shows that existing composite indices (i.e., silhouette-Coefficient or Dunn-Index) that are a combination of primitive cvi might prove to be too generalized and not suitable for every specific task. Therefore, individual clustering goals and corresponding indices should be developed for every task. Hennig et al. introduce a methodology to normalize and calibrate cvi (Hennig 2020) and propose two general-purpose composite indices (Akhanli and Hennig 2020). They remark that, in particular, the weighting of indices poses a challenge to the creation of task-specific composite indices. While Hennig et al. lay the (mathematical) foundation to identify an individual "best" solution, they provide neither a methodology to identify the relevant indices nor a method for weighting them for a given task. Yet they provide the mathematical foundation to do so. The determination of individual cluster goals according to a specific task, selecting suitable algorithms, tuning and comparing them in order to select the "best" clustering results is outlined in the following paper. The focus of it is to include industry and clustering-specific expertise into the clustering process to create an individual composite index to compare clustering results. A methodology and a workflow to weight identified clustering goals is proposed in chapter "Weighting of clustering goals", improving the methodology of Hennig et al. by a multi-criteria decision analysis (mcda) and hence building the missing bridge from the mathematical foundation to a practical implementation. The method is applied on two energy-economic use cases in chapter "Application on energy economic use cases".

Methodology

The following paper builds on relative and internal cluster validation indices as well as their weighting and combination into a single composite index. The focus of this paper is to provide a practical workflow to conduct unsupervised cluster analysis for real-world tasks and apply it in the energy sector. It extends the methodology in Halkidi et al. (2016) by including a methodology for weighting the cluster goals using mcda. This requires a link between the mathematical formulation of cluster goals as provided in Hennig (2020)

Fig. 1 Methodology of cluster identification
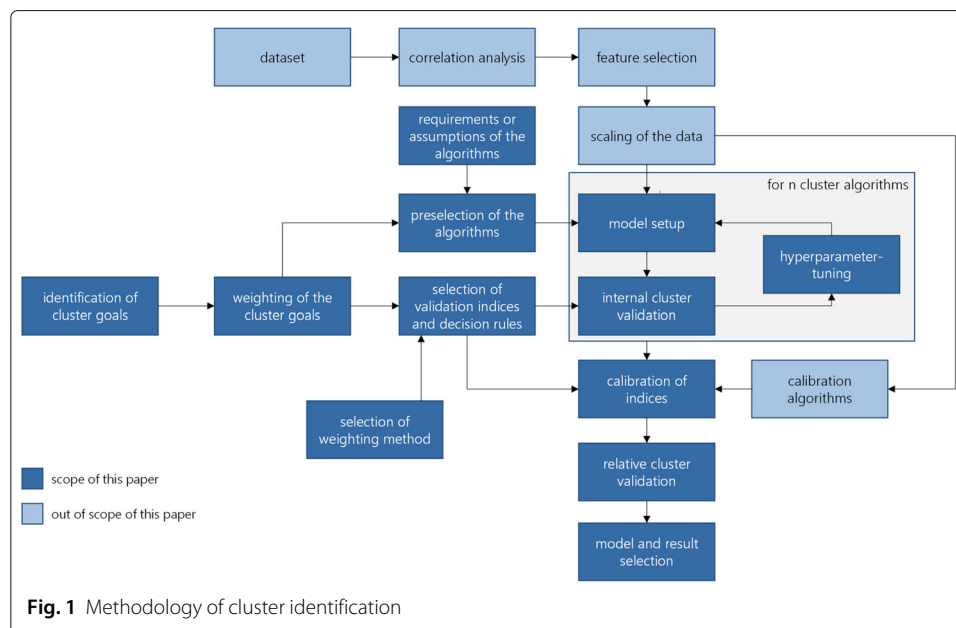
and the practical application according to possible clustering goals. The paper includes practical guidance as to which algorithms and validation indices to use in order to achieve an individual clustering goal. The methodology is outlined in Fig. 1, building upon the proposed validation indices in Hennig et al. (see Halkidi et al. (2016); Hennig (2020)).

The core methodology to identify clusters in an (already) pre-processed dataset builds on the following steps:

1 Identification of cluster goals: depending on the clustering task individual goals have to be chosen in order to choose the best result. In this step, goals are described in purely qualitative terms.

2 Weighting of clustering goals: by a multi-criteria decision analysis. The defined goals can be weighted by a single or by multiple decision makers (e.g., involved stakeholders)

3 Derivation of validation indices: the defined cluster goals (qualitative) must be transformed in mathematical statements utilizing existing validation criteria. Decision rules for these statements have to be formulated (min, max) and the validation criteria normalized [0, 1] to become comparable indices.

4 Preselection of suitable algorithms: by formulating cluster goals, validation indices and decision rules, some algorithms are no longer an option due to conflicting characteristics. The size of the dataset and available computing power are also included.

5 Model setup, internal validation and hyperparameter tuning: the pre-selected algorithms are set up and applied on the dataset. By internally validating the results with the selected cvi, hyperparameters can be tuned in order to iteratively improve the results.

6 Calibration of the clustering results: the resulting validation indices might differ in terms of variance. Hence calibration makes the indices comparable by identifying the normalization range via calibration algorithms.

7   Relative evaluation, model and result selection: the calibrated validation indices can be used to select the overall best model and determine the best clustering result.

The following chapters describe these steps in further detail.

Clustering goals and decision rules

The first logical step to conduct a cluster analysis is to derive task-specific clustering goals. These goals are individual and differ every time, as shown in chapter "Application on energy economic use cases". The clustering goals presented in Hennig (2020) are listed and explained in terms of common clustering goals in the following, whereas the similarity of two datapoints (in this study) is represented by their Euclidean distance. The lower the distance, the more similar two datapoints are, which corresponds to the general definition of clustering in chapter "Definition of clustering and clusters". Considering the nature of clustering, the clustering goals in Van Mechelen and Hampton (1993) can be split in three categories. While some goals describe the cluster definition "bottom-up" for the relation of datapoints and cluster to one other, they do not restrict the clustering result itself. Others a priori restrict the clustering results by their definition. The third category does not affect the clustering result directly but the process of clustering itself, by considering properties of algorithms, such as ease of use. In the following, potential clustering goals for the first two categories are introduced, explained if necessary, and linked to certain validation indices in chapter "Relative and internal cluster validation indices", if possible.

An overview of various clustering goals and corresponding indices described in Hennig (2020) is given in Table 2. However, an index for the representation of a cluster via a datapoint of the original dataset instead of an artificial datapoint (e.g., centroid) is missing. We therefore introduce the following index $I_{cp2cent}$ as described in Table 3.

Table 2 Description of clustering goals

Goal	Index
Within-cluster dissimilarities should be small: this implies that the points within a cluster are all relatively similar to one another.	I_{avg_wc}
Between-cluster dissimilarities should be large: clusters are clearly distinguishable and very different in their characteristics.	I_{p-sep}
Points of a cluster should be well represented by a centroid: a representative of the cluster (that is not an original datapoint) reflects the characteristics of the datapoints within a cluster in the best possible way.	$I_{centroid}$
Members of a cluster should be well represented by a specific datapoint within the dataset (=representative): a single point (that is an original datapoint) reflects the characteristics of the datapoints within a cluster in the best possible way	-
Clusters should correspond to connected areas in data space with high density: datapoints within a cluster always have very similar neighbors yet might not be very similar to every datapoint in the cluster (exception: spherical clusters).	$I_{widestgap}$
All clusters should have roughly the same size.	$I_{entropy}$
The density of clusters should be roughly the same.	I_{cvdens}
The number of clusters should be low (many indices increase with an increasing number (Hennig 2015))	$I_{parsimony}$
The number of clusters should be within a certain range of values.	$I_{targetrange}$[*]
It should be possible to characterize the clusters using a small number of variables: this is especially useful if the result is used for complexity reduction i.e., to create personas.	I_{pps}[*]

[*]Introduced in "Clustering of municipalities" section

Table 3 New index for good representation of data points

Goal	Index	Index Definition
Representation by data points	$I_{cp2cent}$	Measure of how well a cluster is represented by a single point out of its cluster (i.e., closest point x_{cp} to the centroid of the cluster c_i with $x_{cp} \in C_i$). The higher the index, the better the representation.

This index is viable if the features used for clustering are only a lower-dimensional representation of the actual datapoints (e.g., in spatial or time series clustering) and a centroid cannot be converted back in the original (higher) dimension.

Further, very specific restrictions and limitations as well as their mathematical formulation can be found in Hennig (2020). To perform clustering, the above goals must be specified according to the clustering task. Examples are shown in chapter "Application on energy economic use cases".
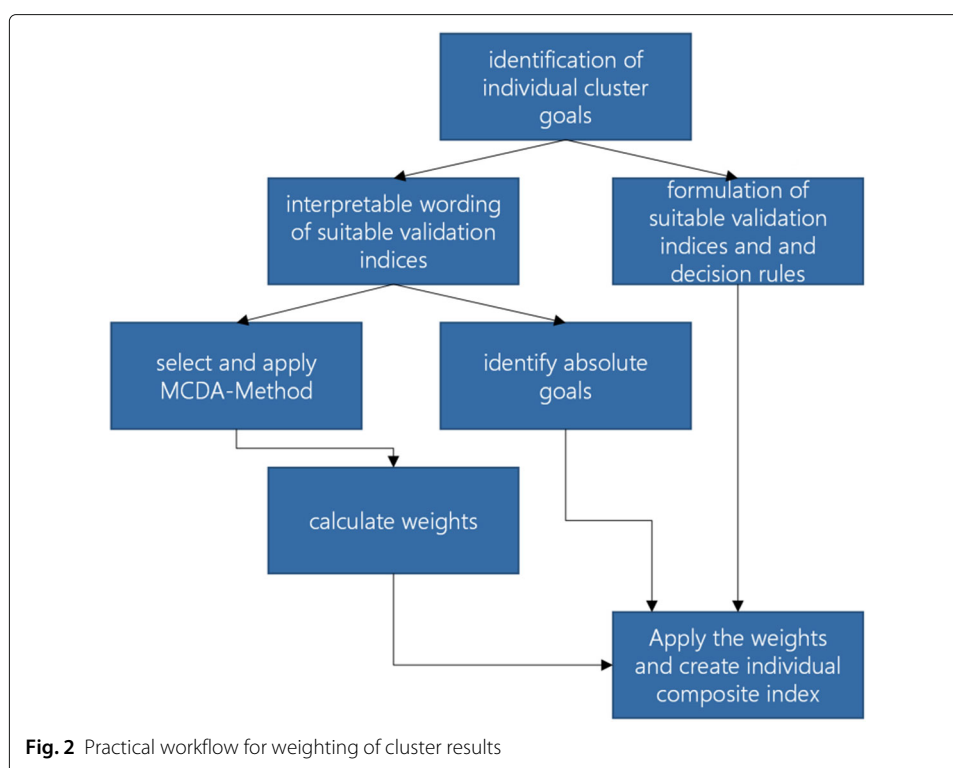
Weighting of clustering goals

Clustering is rarely a purpose in its own right. Especially in practical use cases there is always a specific goal in mind. For example, a customer segmentation analysis or a complexity reduction (see chapter "Application on energy economic use cases"). This paper focuses on energy economic use cases. Yet the methodology is applicable in any clustering task. In order to decide on a best solution among multiple algorithms and results and to simplify and objectify the clustering process, the normalized cvi can be aggregated into one composite index, as proposed in Hennig (2020). While Hennig et al. give a comprehensive methodology to apply validation indices on data and calibrate them, they do not specify how to find suitable individual weights for a distinct, individual goal. A methodology to weight individual clustering goals and therefore the validation indices is proposed in the following and summarized in Fig. 2:

The methodology consists of the following steps:

1. Identify general cluster goals, often set by the specific task and intended use of the results and/or the client
2. Decide on absolute goals: if a set threshold (e.g., minimum number of clusters) is not met, this result is discarded and is not be considered any further.
3. If not already necessary in step 1, find and mathematically formulate validation indices describing every remaining goal and find an understandable wording for them (depending on the decision makers). A list can be found in chapter 3.1.
4. Select and apply an mcda method to these remaining goals to weight them. The selection of the best mcda method depends on the setting and the involvement, knowledge and preference of the involved stakeholders.
5. Calculate the resulting weights of the applied mcda method(s)
6. Calculate an individual composite index by applying the weights to the underlying validation indices on which the understandable formulations are based.

With the second step being a "yes-or-no" decision or strict requirements, the fourth one represents a challenge, as stated in Hennig (2015). To rank certain interpretable goals (linked to mathematically formulated validation indices), we propose the application of "Multi-Attribute Decision Making Methods" (Xu 2015). The goal of these methods is to

Fig. 2 Practical workflow for weighting of cluster results

identify individual weighting factors for previously defined selection criteria (here: clustering goals). Weighting methods can be split in subjective methods (weights are based on the decision maker's judgment and require knowledge and experience in the field) and objective methods. These determine weights by mathematical algorithms or models (Zardari et al. 2015). In order to find a clustering result best suited to individual tasks or goals, subjective methods can be applied. Zardari et al. (2015) suggests among others the methods described in Table 4 to conduct a mcda.

In general, every method has its advantages and disadvantages (as summarized in Zardari et al. (2015)) and can be applied to quantify individual weights. Due to its properties enabling its use for silent negotiation, its easy application in a team, and its focus on unique collective results, we decided on the revised SIMOS method. This method has already been applied in the past in many practical and theoretical energy related projects (e.g. Samweber et al. (2017); Wang et al. (2009); Samweber (2017); Schmuck (2012)). This method builds on the collective and realm-specific knowledge of a team to identify a certain ranking among a set of decision variables (here: clustering goals) (Oberschmidt 2010). There are several variations and iterations of the methodology. The original procedure was introduced by Jean Simos in Simos (1990). It was revised in Figueira and Roy (2002); Pictet and Bollinger (2005) with the latter focusing on practical efficiency and the application with a single or multiple decision makers. Many stakeholders might be involved (e.g., multiple representatives of a client or members of a team) in real-world clustering tasks (as in chapter "Application on energy economic use cases"). The method thus aims at a collective elicitation of weights and thus a consensus among the participants. To apply the SIMOS method, the clustering goals must be understandable to all decision

Table 4 Restrictions and limitations for clusterings

Name	Explanation
Direct Rating	Every decision variable is assigned with an importance independent of the others (as in Likert scale questionnaires).
Ranking Method	Decision variables are ranked relative to one other. These ranks can be used to calculate weights using rank sum, rank reciprocal or rank exponent method.
Point Allocation	Decision makers allocate weights directly to decision variables. The result is normalized.
Pairwise Comparison Method	Decision variables are compared pairwise and the resulting pairwise weights are documented in a matrix. The resulting matrix is used to calculate the overall weights and a consistency ratio.
Swing Weighting Method	All decision variables are set to the worst score. Decision makers can change the score of individual variables by moving them to the best score. The rank of doing so determines the importance (Leijten et al. 2017).
Graphical Weighting Method	This graphical method utilizes a horizontal line to place decision variables relative to one other. Their distance determines their assigned weights.
(Revised) SIMOS Weighting Method	Decision variables are ranked relative to one other. Variables may share the same rank. The relative ranks can be increased by inserting empty ranks in between. In the last step, decision makers need to decide how many times more important the first variable is compared to the last. This rank is used to assign weights.
Fixed Point Scoring	Decision makers need to distribute a finite number of points to weigh decision variables.

makers. Therefore, instead of a mathematical formulation, the impact of a certain decision variable must be formulated in a clear (target group-specific) and interpretable way. Some suggestions can be found in chapter "Application on energy eco- nomic use cases". The SIMOS method then provides the necessary set of rules to rank these goals relative to one another. Based on the rank of the goals r and a selected weighting factor f, the exact weighting can finally be calculated by linear interpolation for any goal ϕ_i using the following formula from Wilkens (2012):

$$\phi_i = r_{min} + (f - 1)\frac{r_i - r_{min}}{r_{max} - r_{min}}$$

This methodology makes it possible to find relatively unbiased weightings ϕ_i (with $\sum_i \phi_i = 1$) for all defined goals. It also focuses purely on the task and is completely unbiased if applied prior to the clustering process. The generated ranking is applied to the underlying indices I_j to create a single composite index I_{agg} for a specific task according to Akhanli and Hennig (2020):

$$I_{agg} = \sum_{j=1}^{s} \phi_j I_j(C)$$

It must be stated that some evaluation criteria may correlate heavily. The inclusion of highly correlated evaluation criteria might by itself increase their weight (Akhanli and Hennig 2020). The set of decision rules generated in this way can be used to pre-select algorithms, optimize their respective hyperparameters and compare the results.

Algorithm pre-selection

In the first step after the determination of the clustering goals and decision rules, suitable algorithms have to be pre-selected. This step depends highly on many individual parameters:

1   length and feature space of the dataset
2   n-dimensional structure of the existing clusters in a dataset
3   characteristics of the algorithms
4   available computational power and time
5   ease of use
6   requirements for the clustering process (see chapter "Clustering goals and decision rules")

For reasons of scope, this topic will not be discussed further. Yet, some clustering algorithms are favored towards certain indices. After weighting them, the suitable algorithms should be selected. For example, k-means optimizes towards the best representation by a centroid ($I_{centroid}$) by minimizing the within-cluster sum of squares. Further, "axioms and theoretical characteristics of clustering methods" can be found in Hennig (2015) chapter 4.3.

Clustering

After the dataset has been prepared, the goals for the clustering have been set, and a range of suitable algorithms has been selected, clustering can be carried out.

Model setup, internal validation & hyperparameter -tuning

The models need to be setup and run to carry out clustering. The results must be evaluated with the selected indices in chapter "Weighting of clustering goals" and normalized (see Hennig (2020)) and the hyperparameters tuned in order to improve the models' results according to the defined goals.

Calibration

The different validation indices may have very small variance and are therefore sometimes hard to compare to those with high variance. Hennig introduces a calibration technique utilizing naïve, random clusterings and therefore a mean/standard deviation-based standardization (Hennig 2020). This is achieved by a "stupid k-centroids" and "stupid nearest neighbors" approach. Both have different assumptions about their results and thus help to increase the range of values of an index.

Scaling

In order to further simplify the decision process by calibrating the results, we further propose a simple scaling process. For any cvi, we set the best value to 1 and the worst value to 0. Since the value range of calibrated indices as proposed in Hennig (2020) is not limited between 0 and 1, a composite index based on weighted aggregation of selected indices could be dominated by single indices which would distort the original weighting. Hence, to compare selected clusterings, we scale their corresponding calibrated indices between 0 and 1. Assuming (for a specific index) that the mean of the "stupid" clusterings is always lowest, we scale the interval from 0 to the highest index to [0, 1]. Otherwise, the worst index of the selected clusterings is set as the lower limit. However, we do not scale $I_{parsimony}$, or $I_{targetrange}$ since they only depend on the number of clusters and are not calibrated, thus they are between 0 and 1 by definition.

Relative validation, model and result selection

After an individual, task-specific composite index (I_{agg}) is created and the clustering is carried out with different algorithms, the results are compared by utilizing the individual indices. The clustering result with the highest value is selected as the best overall result.

Application on energy economic use cases

In the following chapter, the introduced methodology will be applied to two use cases in the field of the energy economics from different research projects with varying goals. The datasets and tasks include the unsupervised clustering of municipalities and driving & load profiles of electric vehicles. The following chapters will give a brief overview of the tasks, data and results. The focus will be put on the methodology introduced in chapter "Methodology". Neither the dataset nor the performed pre-processing will be discussed in detail and will be found in their detailed respective publications.

Clustering of municipalities

Within the InDEED research project (03E16026A) an optimization and simulation framework for blockchain use cases within the field of labeling of renewable energies, p2p-trading and energy communities will be built. Due to computational limitations and the complexity of the optimization and simulation, the municipal level is to be considered. The goal of the clustering is to identify representative German municipalities that do exist and represent the other municipalities of the same cluster in the best way. In a later step the simulated economical potential of the use cases in representative municipalities will be used to calculate the potential in those municipalities that could not be simulated. In order to do so, a regression model will be applied to inter- and extrapolate the simulated potentials to non-simulated municipalities. The dataset consists of 11.994 municipalities, described with 27 selected features ranging from number of inhabitants and installed renewable capacities to peak load and geographical size.

Application of the method

The application of the SIMOS method worked smoothly with seven members of the project-team InDEED. The participants included experts with technical and economics background in energy economics, new business models and digitization, who functioned as product owners and were responsible for the evaluation of the simulation result. Additionally, one participant was responsible for the development of the simulation framework utilized on the clustering data. As described in chapter "Clustering goals and decision rules", clustering goals and decision rules were brainstormed in the team as qualified statements. During the brainstorming, the focus was set on understanding the statements and possible implications. The results were weighted according to chapter "Weighting of clustering goals". Qualified statements were then described mathematically building on chapter "Relative and internal cluster validation indices". The results can be seen in Table 5.

Clustering goals and decision rules

In addition to the ranks, the weighting factor f was determined as 13.2 resulting in the presented weights. Some requirements formulated by the participants, are not yet defined in "Relative and internal cluster validation indices" section. Hence, two

Table 5 Clustering goals and decision rules for municipalities

| Goal	Explanation	Mathematical formulation	Simos Rank	Weight in %
Members of a cluster should be well represented by a specific datapoint within the dataset.	This is necessary in order to a) simulate a real municipality and b) let it be as similar to other points in the cluster as possible. Input features are a lower dimensional representation of municipalities.	$\max(I_{cp2cent})$	13	13.20
The number of clusters should be as low as possible.	Since the resulting clusters are the basis for a subsequent optimization with high computation time, a lower number is favored.	$\max(I_{parsimony})$	9	9.13
Clusters should be clearly distinguishable.	Since one goal is to create "personas" with the clusters in order to improve explainability, clusters should be distinguishable.	$\max(I_{p-sep})$	9	9.13
Communities within a cluster should be structurally similar.	As similarity is defined by Euclidean distance, pairwise distances should correlate with cluster affiliation.	$\max(I_{pearson})$	9	9.13
The number of clusters should be between 5 and 30.	The experts in the simulation software estimate an upper limit of 30 possible simulations. In order to make the clustering viable, a minimum of 5 clusters was determined by the participants.	$\max(I_{targetrange})$	7	7.10
Within-cluster dissimilarities should be small.	This makes sure that not only the representative but also all datapoints in a cluster are comparable.	$\max(I_{avg\_wc})$	7	7.10
Clusters should be describable by a low number of features.	Next to having unique and distinguishable characteristics, in order to create understandable "personas", the number of characterizing features should be as low as possible.	$\max(I_{pps})$	5	5.07
Clusters should be relatively even in size.	A clustering with 90% of the datapoints in one cluster is not desirable. Hence the participants agreed on this parameter.	$\max(I_{entropy})$	1	1.00

qualitative statements with missing indices had to be formulated, see Table 6. This shows that an algorithmic or mathematical definition of new cvi is not only necessary, but a potential issue. Not any qualitative statements might be formulated as such.

Figure 3 shows the comparison of five clusterings with different algorithms and hyper-parameters (in A & B). With the chosen and weighted indices, the two clusterings with k-means best suit the needs of the use case. While both results (A & B) have high values in terms of $I_{cp2cent}$, the other algorithms perform relatively poorly in comparison. This

Table 6 New indices for municipality clustering

Name	Abbreviation	Usage
The number of clusters should be between 5 and 30.	$\max(I_{targetrange})$	Similarly to parsimony, the target range index assesses the number of resulting clusters k. If k is within this target range, the index is 1, if it is lower than the lower limit k_{min}, it increases linearly from 0 at $k = 0$ to 1 at k_{min}. For values larger than the upper limit k_{max}, the value decreases analogously, reaching zero at $k = k_{min} + k_{max}$.
Clusters should be describable by a low number of features.	$\max(I_{pps})$	This parameter builds on the predictive power score (PPS) (Sharma 2020). The PPS uses machine learning to find (pairwise) linear and non-linear relations between two feature vectors. The proposed index calculates the PPS between every feature vector and the clustering results. A threshold to imply a "good" correlation between features and results is set. The mean number of features describing the resulting cluster result well is used to derive a cvi according to the Parsimony (IP) with K_{max} as the dimensionality of the features.

is to be expected because the k-means optimize towards a minimum distance of cluster points to their respective cluster centroid. If the centroid has a neighboring point of the same cluster very close by, the results of $I_{cp2cent}$ are hence almost identical to $I_{centroid}$. The highly ranked I_{p-sep} performs the best in clustering A & B and very poorly in E. $I_{parsimony}$, a measure to express the preference for a lower number of clusters, is rather low overall due to the numbers of clusters ranging from 13 to 19. The newly introduced I_{pps} performs well in E, yet is still high in A & B. $I_{targetrange}$ is 1 for all clusterings since only results within that range were used for the comparison. Due to these clustering results, A is determined as the best overall result (out of the compared clusterings) for the needs of the project team with an I_{agg} (weighted average) of 0.514. This shows that not all clustering goals are met perfectly. Hence, further clusterings will be conducted in the future, to improve the results towards $I_{agg} = 1$. A specific publication introducing and validating the results is currently in progress.

Fig. 3 Cluster validation indices for municipalities
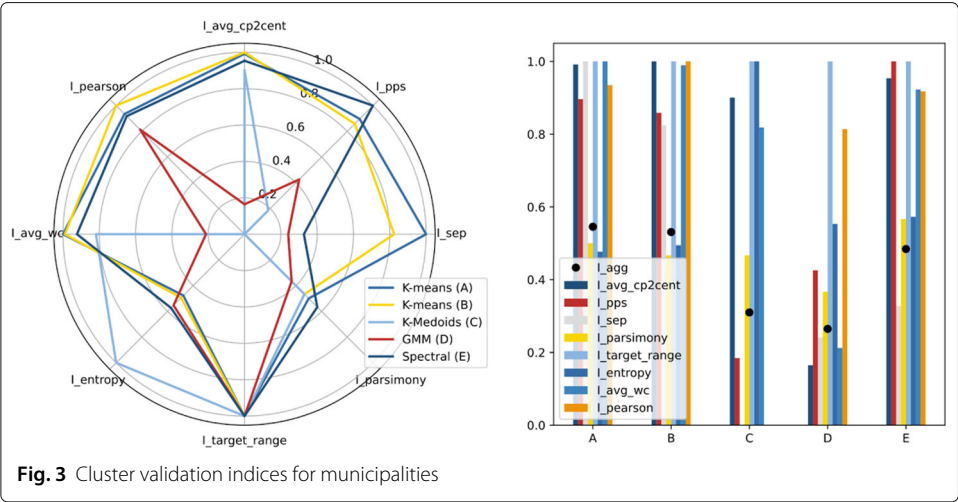
Clustering of driving & load profiles of electric vehicles

The BDL project focuses on the development of and research on bidirectional electric vehicles. One goal is to conduct a systemic evaluation of the impact of bidirectional electric vehicles in Germany. The optimization framework for this task is specified in Böing et al. (2018). In order to reduce complexity, the given driving & load profiles should be clustered in about 20-25 clusters. A preliminary analysis by the project team shows an anticipated optimum of model runtime and variance of load profiles in this range (i.e., the measured runtime of the model decreases by factor 3.2 if 25 instead of 1.000 load profiles are used). The dataset contains 9.997 load profiles represented in 337 features.

Application of the method

The method was applied by a team including six experts, four from the BDL project (01MV18004F) and two external clustering experts. The procedure was equivalent to chapter "Clustering of municipalities". The results can be seen in Table 7.

Clustering goals and decision rules

In addition to the ranks, the weighting factor f was determined as 5.25 resulting in the presented weights. The goal of this clustering was relatively comparable to chapter "Clustering of municipalities". With a different simulation framework in a far more

Table 7 Clustering goals and decision rules for driving & load profiles of electric vehicles

Goal	Explanation	Mathematical formulation	Simos Rank	Weight in %
The number of clusters should be as low as possible.	Since the resulting clusters are the basis for a subsequent optimization with high computation time, a lower number is favored.	$\max(I_{parsimony})$	6	33.0
Clusters should be relatively even in size.	The resulting representative driving & load profiles will be distributed according to their cluster size. If single clusters are overrepresented due to their size, the same driving & load profiles will be used and hence the desired variance will be low.	$\max(I_{entropy})$	6	33.0
Members of a cluster should be well represented by a specific datapoint within the dataset.	This is necessary in order to a) simulate driving & load profiles and b) have it be as similar to other points in the cluster as possible. Input features are a lower dimensional representation of driving & load profiles.	$\max(I_{cp2cent})$	5	27.7
Clusters should be describable by a low number of features.	Next to having unique and distinguishable characteristics, in order to create understandable "personas", the number of characterizing features should be as low as possible.	$\max(I_{pps})$	1	6.3

Fig. 4 Cluster validation indices for bidirectional load profiles of electric vehicles
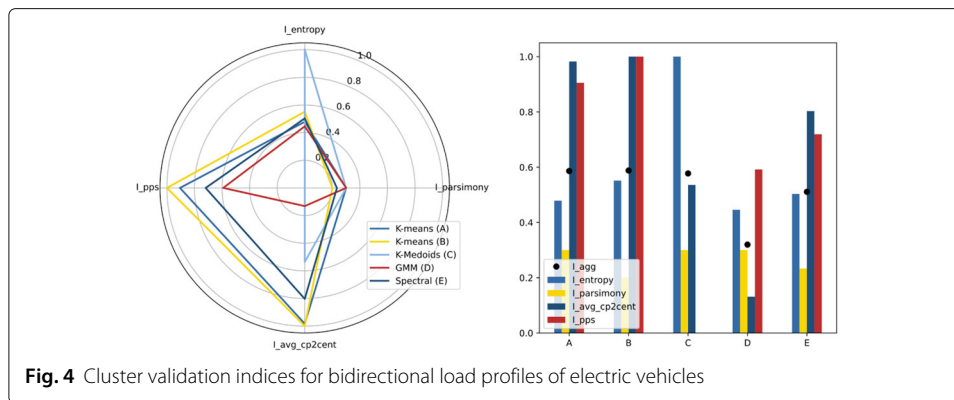
developed stage, the experts had very different goals, resulting in more clearly defined requirements.

The results as depicted in Fig. 4 show a big difference in terms of their cluster goals. While A and B show good results with $I_{cp2cent}$ (for an explanation, see chapter "Clustering of municipalities") and I_{pps}, their $I_{entropy}$ is relatively low compared to C. C has the overall lowest I_{pps} (0). A high $I_{parsimony}$ could not be reached in any of the clusterings, as it decreases with a higher number of clusters. All in all, this shows a tradeoff for all cluster results and the importance of the weighting process. For this use case, the k-means clustering A with 21 cluster reaches the highest I_{agg} of 0.81. Again, further clusterings will be carried out in order to improve the results.

Discussion

The proposed methodology is aimed at improving individual clustering results. Building on the previous works about cvi of Hennig et al. 2020, it adds a practical workflow as well as an mcda methodology to decide on individual weights and suggests new indices. This helps professionals in the field of data science and experts from different areas to identify the individually "best" clustering goals and benchmark different algorithms. The examples in chapter "Application on energy economic use cases" show promising results in the field of energy economics. The chosen cvi as well as their weights and resulting I_{agg} differ, even though the overall goal is relatively similar. This supports the need for the introduced methodology. However, the two examples also show that the set goals by the project teams could not be fully met by the clusterings. Even though, this method helps identifying the individually "best" result, it does not optimize towards it. The flaws of the methodology are outlined in the following:

- Result generation: the methodology is capable of comparing different clustering results with a single, individual composite index (I_{agg}). Generating the results still is challenging task and is of exploratory nature.
- Scalability: every exploratory approach comes with scaling issues. The bigger the dataset compared to the available computational power, the longer it takes to conduct the clustering itself and the calculation of the validation indices.
- Optimization towards indices: with defined indices, it should be possible to mathematically optimize towards a real "best" result. In the cases presented, the clustering was conducted manually. This process should be addressed in future works.

- Bias towards higher numbers of clusters: many indices improve with an increasing number of clusters. While the tendency of a clustering towards a lower number of clusters is expressed via the parameter "parsimony", it might still be weighted low or excluded by certain users.
- Correlation of indices: the resulting indices might correlate and hence be overrepresented even after the weighting. This should be addressed in future works.
- Missing indices: the two example showed that some indices had to be defined ($I_{cp2cent}$, I_{pps}, $I_{targetrange}$) after the mcda method. Depending on the complexity of the missing indices, their mathematical formulation might be time consuming and prone to error if defined incorrectly.
- Further validation: the methodology was conducted with two energy economic examples in different project teams. This showed that the application of mcda methods is possible and helps in tailoring an individual composite index. It also shows that comparing results can be simplified with an individual composite index I_{agg}. To prove the viability of the resulting composite indices, extended research in different fields of application has to be conducted. Further cases (e.g., deriving personas for marketing of utilities) will be applied in the future to show the universal usability. Further clusterings in the presented cases will be executed to improve the results.
- Detailed result analysis: due to scope and length restrictions, a detailed introduction, visualization and validation of the clustering results could not be provided in this paper. This will be addressed in further publications.

Summary and outlook

With ongoing digitization in many sectors, the importance of practical data-analysis, exploration and -usage is increasing significantly. A part of this process is the clustering of data for different practical reasons. These include the reduction and simplification of information complexity, pattern recognition, knowledge expansion, an increased understanding of the data or the detection of outliers. A growing field of use is the energy system analysis in order to reduce input complexity (see examples in chapters "Clustering of municipalities" and "Clustering of driving & load profiles of electric vehicles"). The literature review shows a wide variety of available clustering algorithms. However, it was also possible to identify a gap in their neutral comparison tailored to the individual requirements of practitioners. Most realm-specific papers provide little to no explanation on their cvi choice or choice in clustering algorithm(s). Existing literature presents generalized composite indices or a relatively mathematical formulation of individual cvi in the works of Hennig et al. 2020. While the former are relatively generalized and might not suit individual needs, the latter proposes a viable methodology but lacks a "bridge" to practical application. This paper focused on summarizing the necessary theoretical background as well as the status quo of the scientific discussion. A methodology was developed and proposed to help practitioners tailor an individual composite index to find the best clustering results according to their individual goals from a set of clustering results. This proposes an alternative to better define and achieve individual cluster objectives than with (often) randomly selected composite indices, as done in many cluster-related scientific studies. It creates a practical workflow for energy related projects, adds a mcda method to weight indices and adds further cvi to the method introduced by Hennig (2020). Two examples with different energy economical goals show that the method works

with practitioners. The practical application in mcda workshops showed that there were cvi missing. In this case, these indices need to be defined and mathematically formulated. The already existing composite indices, introduced in chapter "Literature review", may contain useful individual cvi, once decomposed into their components. $I_{cp2cent}$ was introduced in this paper due to practical needs and its viability shown in cases with high distances between centroids and datapoints from the original dataset. However, this also shows that the indices can correlate, which in turn can mean overrepresentation in individual composite index. I_{pps} was introduced in order to evaluate whether results are describable by a low number of features using non-linear-correlations (Sharma 2020). $I_{targetrange}$ was introduced to prefer not only lower number of clusters (as in $I_{parsimony}$) but numbers of clusters within a defined target range. The methodology proved viable to compare different clusterings of multiple algorithms towards individual goals. If the clustering goals can be reached with the provided datasets and specified I_{agg} can not be ensured with the methodology. Whether an optimization towards I_{agg} is possible, should be part of further research. The clusterings introduced in chapter "Application on energy economic use cases" will be used in further research and the respective papers concerning the results will be published. Further clusterings will be conducted to improve the results. Its application in other projects with different clients will prove its practicality in the future. All in all, the methodology can be helpful for data scientists and engineers to help find an optimal clustering result with clients or tasks with respective experts in this field with low or no prior knowledge on clustering.

Authors' contributions
AB developed the concept and methodology with a focus on applications within the energy sector. He wrote the paper and the clustering in the mentioned python framework. YF implemented the validation indices in python and prepared the visualizations. He provided critical feedback, validation, and helped with the finalization. The authors read and approved the final manuscript.

Availability of data and materials
Clustering input-features as well as results will be introduced in subsequent papers in detail. In this context, the data will be made publicly available on www.opendata.ffe.de.

Declarations

Competing interests
The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Published: 13 September 2021

References
Ackerman M, Ben-David S (2009) Clusterability: A theoretical study. In: Proceedings of the 12th International Conference on Artificial Intelligence and Statistics, PMLR 5. pp 1–8

Akhanli SE, Hennig C (2020) Comparing clusterings and numbers of clusters by aggregation of calibrated clustering validity indexes. Stat Comput 30(5):1523–1544

Bittel HM, Perera ATD, Mauree D, Scartezzini J-L (2017) Locating multi energy systems for a neighborhood in geneva using k-means clustering. Energy Procedia 122:169–174

Böing F, Murmann A, Pellinger C, Bruckmeier A, Kern T, Mongin T (2018) Assessment of grid optimisation measures for the German transmission grid using open source grid data. In: Journal of Physics: Conference Series. SciGRID International Conference on Power Grid Modelling, Oldenburg Vol. 977. pp 30–31

Brickey J, Walczak S, Burgess T (2010) A comparative analysis of persona clustering methods. AMCIS 2010 Proc:217

Cai J, Wei H, Yang H (2020) A novel clustering algorithm based on DPC and PSO. IEEE Access 8:88200–88214

Chen X (2015) A new clustering algorithm based on near neighbor influence. Expert Syst Appl 42(21):7746–7758

Chou CH, Su MC, Lai E (2002) Symmetry as a new measure for cluster validity. In: 2nd WSEAS Conference on Scientific Compuation and Soft Computing. pp 209–213

Cormos C-C, Petrescu L, Chisalita D-A (2020) Environmental evaluation of european ammonia production considering various hydrogen supply chains. Renew Sust Energ Rev 130:109964

Das S, Abraham A, Konar A (2008) Automatic kernel clustering with a multi-elitist particle swarm optimization algorithm. Pattern Recogn Lett 29(5):688–699

Figueira J, Roy B (2002) Determining the weights of criteria in the ELECTRE type methods with a revised Simos' procedure. Eur J Oper Res 139(2):317–326

Gan G, Ma C, Wu J (2007) Data Clustering: Theory, Algorithms, and Applications. York University, Toronto, Canada

Gheorghe G, Scarlatache F (2015) An assessment of the renewable energy potential using a clustering based data mining method. Case study in Romania. Energy 81:416–429. https://doi.org/10.1016/j.energy.2014.12.054

Halkidi M, Vazirgiannis M, Hennig C (2016) Method-Independent Indices for Cluster Validation and Estimating the Number of Clusters. Chapman and Hall/CRC, Boca Raton

Hennig C (2015) Clustering strategy and method selection. In: Rocci R, Murtagh F, Meila M, Hennig C (eds). Handbook of Cluster Analysis, 1st ed. Chapman and Hall/CRC. https://doi.org/10.1201/b19706

Hennig C (2020) Cluster validation by measurement of clustering characteristics relevant to the user. Data Anal Appl 1: Clustering Regression, Modeling-estimating, Forecast Data Mining 2:1–24

Hennig C, Liao TF (2010) Comparing latent class and dissimilarity based clustering for mixed type variables with application to social stratification (Research Report No. 308). Department of Statistical Science, University College London, London

Kleinberg J (2002) An impossibility theorem for clustering. In: Proceedings of the 15th International Conference on Neural Information Processing Systems (NIPS' 02). pp 463–470

Kou G, Peng Y, Wang G (2014) Evaluation of clustering algorithms for financial risk analysis using MCDM methods. Inf Sci 275:1–12

Kuwil FH, Shaar F, Topcu AE, Murtagh F, Applications ES (2019) A new data clustering algorithm based on critical distance methodology. Expert Syst Appl 129:296–310

Leijten F, Boland M, Tsiachristas A, Hoedemakers M, Verbeek N, Islam K, Askildsen JE, de Bont A, Bal R, Rutten-van Mölken M (2017) Development of an analytical framework to perform a comprehensive evaluation of integrated care. 53–56

Liu Y, Li Z, Xiong H, Gao X, Wu J (2010) Understanding of internal clustering validation measures. In: 2010 IEEE International Conference on Data Mining. pp 911–916. https://doi.org/10.1109/ICDM.2010.35

Lorr M (1983) Cluster Analysis for Social Scientists. Jossey-Bass Inc.,U.S., San Francisco, USA

Martinez WL, Martinez AR, Solka JL (2010) Exploratory Data Analysis with MATLAB. Chapman and Hall/CRC Computer Science and Data Analysis, Boca Raton

McInnes L, Healy J, Astels S (2017) hdbscan: Hierarchical density based clustering. J Open Source Softw 2(11):205

Metwalli SA (2020) Clustering 101: How to Choose the Right Algorithm for Your Application - An Introduction to different types of clustering algorithms. towards data science, Toronto, Canada. https://towardsdatascience.com/clustering-101-how-to-choose-the-right-algorithm-for-your-application-fb1521ea13fc. Accessed 01 Mar 2021

Oberschmidt J (2010) Multikriterielle Bewertung von Technologien zur Bereitstellung von Strom und Wärme. Universität Göttingen, Göttingen

Pictet J, Bollinger D (2005) The silent negotiation or How to elicit collective information for group MCDA without excessive discussion. J Multi-Criteria Decis Anal 13:199–211. Lausanne

Puzicha J, Hofmann T, Buhmann JM, Letters P (2000) A theory of proximity based clustering: structure detection by optimization. Pattern Recogn 33(4):617–634

Rendón E, Abundez I, Arizmendi A, Quiroz EM (2011) Internal versus external cluster validation indexes. Int J Comput Commun 5(1):27–34

Rendón E, Abundez IM, Gutierrez C, Zagal SD, Arizmendi A, Quiroz EM, Arzate HE (2011) A comparison of internal and external cluster validation indexes. In: Proceedings of the 2011 American conference on applied mathematics and the 5th WSEAS international conference on Computer engineering and applications Vol. 29. pp 158–163

Rendón E, Garcia R, Abundez I, Gutierrez C, Gasca E, Del Razo F, Gonzalez A (2008) Niva: a robust cluster validity. In: Proceedings of the 12th WSEAS International Conference on Communications (ICCOM'08). pp 241–248

Samweber F (2017) Systematischer Vergleich Netzoptimierender Maßnahmen zur Integration elektrischer Wärmeerzeuger und Fahrzeuge in Niederspannungsnetze (Doctoral Thesis). Technical University of Munich. p 104

Samweber F, Köppl S, Bogensperger A, Böing F, Bruckmeier A, Estermann T, Müller M, Zeiselmair A (2017) Abschlussbericht Einsatzreihenfolgen - Projekt MONA 2030: Ganzheitliche Bewertung Netzoptimierender Maßnahmen gemäß technischer, ökonomischer, ökologischer, gesellschaftlicher und rechtlicher Kriterien. pp 125–127

Schmuck P (2012) Transdisciplinary Evaluation of Energy Scenarios for a German Village Using Multi-Criteria Decision Analysis. Sustainability 4(4):604–629

Schütz T, Schraven MH, Fuchs M, Remmen P, Müller D (2018) Comparison of clustering algorithms for the selection of typical demand days for energy system synthesis. Renew Energy 129:570–582

Sharma P (2020) What is Predictive Power Score (PPS) – Is it better than Correlation ? [With Python Code]. Machine Learning Knowledge, Carlsbad. https://machinelearningknowledge.ai/predictive-power-score-vs-correlation-with-python-implementation/. Accessed 01 Mar 2021

Siala K, Mahfouz MY (2019) Impact of the choice of regions on energy system models. Energy Strateg Rev 25:75–85

Simos J (1990) Evaluer l'impact sur l'environnement. Une approche originale par l'analyse multicritère et la négociation. Presses polytechniques et universitaires romandes, Lausanne

Syakur MA, Khotimah BK, Rochman EMS, Satoto BD (2018) Integration k-means clustering method and elbow method for identification of the best customer profile cluster. In: IOP Conference Series: Materials Science and Engineering Vol. 336. p 012017

Tanwar AK, Crisostomi E, Raugi M, Tucci M, Giunta G (2015) Clustering analysis of the electrical load in european countries. In: 2015 International Joint Conference on Neural Networks (IJCNN). pp 1–8

Tomasini C, Emmendorfer L, Borges EN, Machado K (2016) A methodology for selecting the most suitable cluster validation internal indices. In: Tomasini C, et al (eds). SAC '16: Proceedings of the 31st Annual ACM Symposium on Applied Computing. Association for Computing Machinery, New York

Toussaint W, Moodley D Comparison of clustering techniques for residential load profiles in South Africa? In: Toussaint W, et al (eds). Proceedings of the South African Forum for Artificial Intelligence Research Fair. CEUR Workshop Proceedings, 2019

Tukey JW (1977) Exploratory Data Analysis. Addison-Wesley Publishing Company, Boston

Van Mechelen I, Hampton JA (1993) Categories and Concepts: Theoretical Views and Inductive Data Analysis. Academic Press, New York, USA

Vendramin L, Campello RJGB, Hruschka ER (2010) Relative clustering validity criteria: A comparative overview. Stat Anal Data Mining: ASA Data Sci J 3(4):209–235

Wang J, You-Yin J, Chun-Fa Z, Jun-Hong Z (2009) Review on multi-criteria decision analysis aid in sustainable energy decision-making. Renew Sust Energ Rev 13(9):2263–2278. Amsterdam

Wilkens I (2012) Multikriterielle Analyse zur Nachhaltigkeitsbewertung von Energiesystemen - Von der Theorie zur praktischen Anwendung. Technische Universität Berlin, Berlin

Xu Z (2015) Uncertain Multi-Attribute Decision Making: Methods and Applications. Springer, Berlin; Heidelberg

Wang X, Xu Y (2019) An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index. In: IOP Conference Series: Materials Science and Engineering. 569. 052024. https://doi.org/10.1088/1757-899X/569/5/052024

Yang J, Ning C, Deb C, Fan Z, Cheong D, Eang Lee S, Sekhar C, Tham KW (2017) K-Shape clustering algorithm for building energy usage patterns analysis and forecasting model accuracy improvement. Energy Build 146:27–37

Zardari NH, Ahmed K, Shirazi SM, Yusop ZB (2015) Weighting Methods and their Effects on Multi-Criteria Decision Making Model Outcomes in Water Resources Management. Springer, Basel

Zhou K, Yang S, Shao Z (2017) Household monthly electricity consumption pattern mining: A fuzzy clustering-based model and a case study. J Clean Prod 141:900–908

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.