

# Clustering of fMRI data: the elusive optimal number of clusters

Mohamed L. Seghier

Cognitive Neuroimaging Unit, Emirates College for Advanced Education, Abu Dhabi, United Arab Emirates

## ABSTRACT

Model-free methods are widely used for the processing of brain fMRI data collected under natural stimulations, sleep, or rest. Among them is the popular fuzzy  $c$ -mean algorithm, commonly combined with cluster validity (CV) indices to identify the ‘true’ number of clusters (components), in an unsupervised way. CV indices may however reveal different optimal  $c$ -partitions for the same fMRI data, and their effectiveness can be hindered by the high data dimensionality, the limited signal-to-noise ratio, the small proportion of relevant voxels, and the presence of artefacts or outliers. Here, the author investigated the behaviour of seven robust CV indices. A new CV index that incorporates both compactness and separation measures is also introduced. Using both artificial and real fMRI data, the findings highlight the importance of looking at the behavior of different compactness and separation measures, defined here as building blocks of CV indices, to depict a full description of the data structure, in particular when no agreement is found between CV indices. Overall, for fMRI, it makes sense to relax the assumption that only one unique  $c$ -partition exists, and appreciate that different  $c$ -partitions (with different optimal numbers of clusters) can be useful explanations of the data, given the hierarchical organization of many brain networks.

**Subjects** Bioinformatics, Computational Biology, Neuroscience

**Keywords** Functional MRI, Data-driven analysis, Unsupervised fuzzy clustering, Brain networks, Cluster validity, Fuzzy compactness and separation

Submitted 6 May 2018  
Accepted 19 July 2018  
Published 3 October 2018

Corresponding author  
Mohamed L. Seghier,  
mseghier@gmail.com

Academic editor  
Jafri Abdullah

Additional Information and  
Declarations can be found on  
page 23

DOI 10.7717/peerj.5416

© Copyright  
2018 Seghier

Distributed under  
Creative Commons CC-BY 4.0

## OPEN ACCESS

## INTRODUCTION

There are many contexts where model-based methods are inadequate to map brain function, including for instance tasks that cannot be fully controlled (e.g., sleep, learning, natural stimulation, continuous rest; *Bartels & Zeki, 2004; Bartels & Zeki, 2005; Hasson et al., 2004; Lee et al., 2012; Malinen, Hlushchuk & Hari, 2007; Zacks et al., 2001*) or when the hemodynamic correlates of neural activity are altered in unknown ways (e.g., patients with impaired vasculature). In such cases, approaches without *a priori* knowledge, known also as model-free or data-driven methods, are of great help.

Several data-driven methods have previously been used in fMRI (*DonGiovanni & Vaina, 2016; Thirion et al., 2014*), including fuzzy clustering (*Baumgartner, Windischberger & Moser, 1998; Fadili et al., 2000; Golay et al., 1998; Jahanian et al., 2004*) and independent component analysis (*McKeown et al., 1998*). These methods have been used in many scenarios to extract meaningful information from fMRI data in the absence of any prior knowledge (*Aljobouri et al., 2018; Baumgartner et al., 2000;*

*Lange et al., 2004; Ma et al., 2011; Smolders et al., 2007; Tang et al., 2015; Wismuller et al., 2004*). One popular data-driven clustering method is based on the classic fuzzy  $c$ -mean (FCM) algorithm (*Bezdek, 1981*). Although FCM allows high computational flexibility, its robustness may depend on several methodological issues. Specifically, these include the initialisation problem, the choice of similarity or distance metric, and the usually unknown optimal number of classes or prototypes (e.g., *Alexiuk & Pizzi, 2004; Esposito et al., 2002; Fatemizadeh, Taalimi & Davoudi, 2009; Jahanian, Soltanian-Zadeh & Hossein-Zadeh, 2005; Lange et al., 2004; Moller et al., 2002; Quiqley et al., 2002; Soltanian-Zadeh et al., 2004; Windischberger et al., 2003*). This study focuses on the issue of the optimal number of clusters that can be extracted from fMRI data.

It is critical for any reliable clustering method to be able to determine whether: (i) the data contains any structure and (ii) the segregated clusters are ‘true’ representations of the data (*Dubes, 1987; Windham, 1981*). This issue is generally expressed in terms of the ability of the algorithm, here FCM, to cluster the data into an optimal number of clusters ( $c_{\text{opt}}$ ). To do that, previous studies have introduced many measures, called cluster validity (CV) indices, to estimate  $c_{\text{opt}}$  in an unsupervised manner (for a review see *Bezdek & Pal, 1998; Hammah & Curran, 2000; Kim & Ramakrishna, 2005; Maulik & Bandyopadhyay, 2002; Wang & Zhang, 2007; Zhou et al., 2014*). The rationale behind these CV indices is that a good and useful clustering should yield compact and well-separated clusters. Indeed, it is not surprising that many proposed CV indices combine different measures of compactness (cohesiveness) and separation (isolation) among clusters, and would reach their optimal values for the best  $c$ -partition (i.e., data clustered into  $c_{\text{opt}}$  clusters).

A few studies have previously investigated the effectiveness of CV indices in the context of fMRI data clustering (e.g., *Alexiuk & Pizzi, 2004; Fadili et al., 2000; Fadili et al., 2001; Goutte et al., 1999; Moller et al., 2002; Seghier & Price, 2009*). Some known features of fMRI data may make the clustering particularly challenging (*Thirion et al., 2014*), including for instance the huge number of points (i.e., voxels) in a typical fMRI dataset, the poor signal-to-noise ratio in fMRI (noisy data), the small proportion of voxels of interest that might be considered as relevant (i.e., an ill-balanced problem), and the presence of artefacts or outliers (i.e., caused by head motion or signal loss). Given this complexity, it might be the case that reliance on a single CV index might not be enough, in particular when the data are noisy and the expected number of clusters is relatively high. Here, the author compared the identified optimal  $c$ -partition when applying different CV indices to the same datasets. In particular, the author investigated the behaviour of different measures of compactness and separation when using previously published CV indices. The current study also aims to introduce a new CV index that specifically incorporates suitable compactness and separation measures that are useful for data with larger optimal number of clusters.

## METHODS

### Fuzzy clustering

Our clustering method was based on the popular fuzzy  $c$ -mean (FCM) algorithm ([Bezdek, 1981](#); [Bezdek et al., 1997](#)). In the context of fMRI, the FCM algorithm can segregate or cluster  $n$  brain voxels (feature vectors) into  $c$  expected clusters ( $c \geq 2$ ). Each voxel  $i$  is a vector  $X_i$  of  $p$  properties (e.g., number of collected volumes or scans). Each cluster  $j$  is characterised by a centroid  $V_j$ , that represents its characteristic timecourse (prototype). The resemblance between each voxel  $i$  and each centroid  $V_j$  is assessed by the distance  $D_{ij}$  between  $X_i$  and  $V_j$ . The degree of membership  $U_{ij}$  is calculated for each voxel  $i$  by comparing  $D_{ij}$  for each cluster  $j$  to all other clusters.

In brief, the standard FCM algorithm iteratively minimises the following objective function  $J_m$ :

$$J_m = \sum_{i=1}^n \sum_{j=1}^c U_{ij}^m \cdot D_{ij}^2 \quad (1)$$

where “ $m$ ” is the degree of fuzziness.

Degrees of membership  $U$  and centroids  $V$  are updated as following:

$$U_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{D_{ij}}{D_{ik}} \right)^{2/m-1}} \quad (2)$$

$$V_j = \frac{\sum_{i=1}^n U_{ij}^m \cdot X_i}{\sum_{i=1}^n U_{ij}^m}. \quad (3)$$

Optimal clustering depends on the choice of the similarity  $D$ , the degree of fuzziness  $m$  and the optimal number of clusters  $c_{opt}$ , as detailed below.

### Similarity measure $D$

Here I used a modified version of the hyperbolic correlation distance proposed previously by [Golay et al. \(1998\)](#). In their work,  $D$  was defined as ([Golay et al., 1998](#)):

$$D_{ij} = \frac{1 - CC_{ij}}{1 + CC_{ij}}. \quad (4)$$

Where  $CC_{ij}$  is the Pearson correlation coefficient between  $X_i$  and  $V_j$ .

Here, a modified version of  $D$  was used:

$$D_{ij} = \frac{\sqrt{|CC_{ij}|} - CC_{ij}}{\sqrt{|CC_{ij}|} + CC_{ij}}. \quad (5)$$

This new formula uses the square root function, a monotonically increasing function over  $x > 0$  that satisfies the following inequality:  $\sqrt{x} \geq x$ , for  $x \in [0, 1]$ . The rationale here was to increase the difference (i.e., discrimination power) between relatively close correlation values in particular between mid and high correlations (cf. [Fig. S1](#)).

## Optimal number of clusters

A good and robust clustering should yield compact and well-separated clusters. This is assumed to be the case when the number of clusters reaches an optimal value  $c_{\text{opt}}$ . The exact  $c_{\text{opt}}$  value is however unknown in fMRI data. Previous reports have suggested that  $c_{\text{opt}}$  can be found within the interval  $[2, \sqrt{n}]$  (Zahid, Limouri & Essaid, 1999); however the exact  $c_{\text{opt}}$  can only be estimated empirically. Typically, FCM is repeated several times with different  $c$  values (i.e., equivalent to an unsupervised fuzzy clustering analysis Fadili et al., 2001) and the  $c$  value that optimises a given criterion, here a given CV index, is considered as the optimal  $c_{\text{opt}}$ , and that criterion is typically defined as a trade-off between compactness and separation.

Before introducing the different CV indices used here, it might be helpful to define the core measures of compactness and separation using unified mathematical notations. These measures can be seen as building blocks that can be combined into different CV indices. Ultimately, the definition of those measures would help appreciate the inherent links (or similarity) between previously suggested CV indices, before introducing the rationale of the new CV index.

## Compactness and separation measures

Two core quantities, noted  $n_{m,j}$  and  $\sigma_{m,j}$ , were defined as following:

$$n_{m,j} = \sum_{i=1}^n U_{ij}^m \quad (6)$$

$$\sigma_{m,j} = \sum_{i=1}^n U_{ij}^m \cdot D_{ij}^2. \quad (7)$$

The measures  $n_{1,j}$  and  $n_{2,j}$  represent the fuzzy cardinality and the fuzzy partition of cluster  $j$  respectively. The quantity  $\sigma_{m,j}$  denotes the fuzzy variation of cluster  $j$ , though other studies have instead used  $\sigma_{1,j}$  as a measure of fuzzy variation (e.g., Gath & Geva, 1989; Rezaee, Lelieveldt & Reider, 1998; Sun, Wang & Jiang, 2004).

Those core quantities can then be combined into different forms to give away different measures of fuzzy compactness (cohesiveness) for a given  $c$ -partition. Using similar notation as previous studies, quantities called  $\pi_{m,1}$  (Bensaid et al., 1996; Zahid et al., 1999),  $\pi_{m,m}$  (Bouguessa, Wang & Sun, 2006), and FC (Fadili et al., 2001; Zahid et al., 1999) were computed as following:

$$\pi_{m,1} = \sum_{j=1}^c \frac{\sigma_{m,j}}{n_{1,j}} \quad (8)$$

$$\pi_{m,m} = \sum_{j=1}^c \frac{\sigma_{m,j}}{n_{m,j}} \quad (9)$$

$$FC = \frac{\sum_{i=1}^n (\max_j (U_{ij}))^2}{\sum_{i=1}^n \max_j (U_{ij})}. \quad (10)$$

Likewise, the fuzzy separation (isolation) between clusters was previously estimated with several fuzzy separations quantities called  $K_m$  (Fukuyama & Sugeno, 1989),  $FS$  (Fadili et al., 2001; Zahid et al., 1999),  $S$  (Zahid et al., 1999) and  $SS$  (Rezaee, Lelieveldt & Reider, 1998):

$$K_m = \sum_{j=1}^c n_{m,j} \cdot \|V_j - \bar{X}\|^2 \quad (11)$$

$$FS = \sum_{j=1}^{c-1} \sum_{k=1}^{c-j} \frac{\sum_{i=1}^n (\min(U_{ij}, U_{i,k+j}))^2}{\sum_{i=1}^n \min(U_{ij}, U_{i,k+j})} \quad (12)$$

$$S = \frac{1}{c} \sum_{j=1}^c \|V_j - \bar{X}\|^2 \quad (13)$$

$$SS = \sum_{j=1}^c \frac{1}{\sum_{k=1}^c \|V_j - V_k\|} \quad (14)$$

where  $\bar{X}$  stands for the global mean of the whole data.

Interestingly, the ratio  $FS/FC$  (i.e., separation divided by compactness) is known as the fuzzy overlap (FO) coefficient (see Fadili et al., 2000 for more details).

Furthermore, different measures of between-centroid distance have been proposed, including the minimum distance  $V_{d\min}$  (e.g., Schwämmle & Jensen, 2010; Xie & Beni, 1991), the maximum distance  $V_{d\max}$  (e.g., Rezaee, Lelieveldt & Reider, 1998), and the minimum distance  $V_{d\min,j}$  between a cluster  $j$  and the remaining clusters (Wu & Yang, 2005):

$$V_{d\min} = \min_{j,k} (\|V_j - V_k\|) \quad (15)$$

$$V_{d\max} = \max_{j,k} (\|V_j - V_k\|) \quad (16)$$

$$V_{d\min,j} = \min_{k \neq j} (\|V_j - V_k\|). \quad (17)$$

These measures, based on the distance between estimated centroids, can be seen as alternative separation measures. They can be handy when the clustering is showing redundant clusters.

This section introduces two new measures of separation and discrimination between voxels by combining different measures of fuzzy cardinality and variation (cf. Eqs. (6) and (7)): a fuzzy intra-cluster ( $ID_{intra}$ ) dissimilarity coefficient and an inter-cluster ( $ID_{inter}$ ) dissimilarity coefficient:

$$ID_{intra} = \max_j \left( \frac{n - n_{1,j}}{n_{1,j}} \cdot \frac{\sigma_{1,j}}{\sum_{k=1, k \neq j}^c \sigma_{1,k}} \right) \quad (18)$$

$$ID_{inter} = \min_j \left( \frac{\min_{k, k \neq j} (\sigma_{1,k})}{\sigma_{1,j}} \right). \quad (19)$$

Small  $ID_{intra}$  values would indicate that, across all clusters, voxels that are close to a given cluster are well-isolated from voxels that are far from that cluster, whereas high  $ID_{inter}$

values indicate well-discriminated voxels (i.e., small fuzzy overlap between clusters). Our initial tests with noisy simulated data showed the need to define new separation measures that are robust to noise and can handle  $c$ -partitions with higher number of clusters, hence the new definitions in Eqs. (18) and (19).

### Cluster validity measures

There are many CV indices in the literature (probably more than 50 indices), hence it is beyond the scope of this study to test all of them. In a preliminary analysis (results not shown here), about 20 selected CV indices were first tested on several simulated datasets (as defined in [Bezdek & Pal, 1998](#); [Bouguessa, Wang & Sun, 2006](#); [Dave, 1996](#); [Fukuyama & Sugeno, 1989](#); [Geva et al., 2000](#); [Kim, Park & Park, 2001](#); [Kim, Lee & Lee, 2003](#); [Kim & Ramakrishna, 2005](#); [Kwon, 1998](#); [Pakhira, Bandyopadhyay & Maulik, 2004](#); [Pakhira, Bandyopadhyay & Maulik, 2005](#); [Pal & Bezdek, 1995](#); [Rezaee, Lelieveldt & Reider, 1998](#); [Rhee & Oh, 1996](#); [Sun, Wang & Jiang, 2004](#); [Tsekouras & Sarimveis, 2004](#); [Wu & Yang, 2005](#); [Xie & Beni, 1991](#); [Yu & Li, 2006](#); [Zahid et al., 1999](#); [Zahid, Limouri & Essaid, 1999](#)). These CV indices were selected from earlier studies (for a similar rationale, see recent comparison study [Zhou et al., 2014](#)), and many of them are well-established indices. Some of these CV indices have been used in previous fMRI studies. More recent CV indices (e.g., see [He, Tan & Fujimoto, 2016](#); [Hu et al., 2011](#); [Lin et al., 2016](#); [Ren et al., 2016](#); [Rezaee, 2010](#); [Yang et al., 2018](#); [Zhang et al., 2014](#)) were not explicitly tested here.

From this preliminary analysis, seven CV indices (out of twenty) were selected according to the following four criteria: CV indices should (i) combine both measures of separation and compactness; (ii) not suffer from monotonic dependency with the number of expected clusters; (iii) not necessitate the categorisation or the binarisation of  $U$  (i.e., crisp degrees of membership) during CV computation; (iv) be fast to compute when  $n$  is expected to be very high (e.g., hundreds of thousands of voxels in the context of fMRI data). The seven selected CV indices that satisfied the different criteria are described below and listed in [Table 1](#).

(1)- The Rezaee-Lelieveldt-Reider index  $CV_{RLR}$  ([Rezaee, Lelieveldt & Reider, 1998](#)):

$$CV_{RLR} = \frac{\sum_{j=1}^c \sigma_{1,j}}{c \cdot \|\sigma_X\|} + \frac{1}{\alpha} \cdot \left( \frac{V_{d\max} \cdot SS}{V_{d\min}} \right). \quad (20)$$

The constant  $\alpha$  is a weighting constant and  $\sigma_X$  is the variance of the whole data set. The best  $c$ -partition is obtained by minimising  $CV_{RLR}$  with respect to the number of clusters  $c$ . In the original definition of  $CV_{RLR}$ , the constant  $\alpha$  was set to 1; however, here  $\alpha$  was set to the value of  $\frac{V_{d\max}}{V_{d\min}} \cdot SS$  at the maximum number of clusters ( $c_{\max}$ ) as suggested previously ([Sun, Wang & Jiang, 2004](#)).

(2)- The Zahid-Limouri-Essaid index  $CV_{ZLE}$  ([Zahid et al., 1999](#); [Zahid, Limouri & Essaid, 1999](#)):

$$CV_{ZLE} = \alpha \cdot \left( \frac{S}{\pi_{m,1}} \right) - \frac{FS}{FC}. \quad (21)$$

The constant  $\alpha$  is independent from  $c$  and was introduced here as a scaling factor to take into account the difference in values between the two subtracted quantities. The constant  $\alpha$

**Table 1** List of the selected cluster validity (CV) indices.

CV index	Proposed by	Range	Value at $c_{opt}$
$CV_{RLR}$	Rezaee-Lelieveldt-Reider index (1998). A modified version was used here (Sun, Wang & Jiang, 2004)	$[0, +\infty[$	Minimal
$CV_{ZLE}$	Zahid-Limouri-Essaid index (1999)	$] -\infty, +\infty[$	Maximal
$CV_{GV}$	Geva index (2000)	$[0, +\infty[$	Maximal
$CV_{KP}$	Kim-Park index (2001)	$[0, +\infty[$	Minimal
$CV_{PBM}$	Pakhira-Bandyopadhyay-Maulik index (2004)	$[0, +\infty[$	Maximal
$CV_{WY}$	Wu-Yang index (2005)	$[-c, c]$	Maximal
$CV_{BWS}$	Bouguessa-Wang-Sun index (2006)	$[0, +\infty[$	Maximal
$CV_{new}$	A new CV index	$[0, +\infty[$	Maximal

was set here to the value of the fuzzy overlap ( $FS/FC$ ) at  $c = c_{max}$  (note that in the original paper of Zahid et al.,  $\alpha$  was set equal to 1). The best  $c$ -partition is obtained by maximising  $CV_{ZLE}$  with respect to  $c$ . This  $CV_{ZLE}$  index has previously been used for fMRI analysis (Fadili et al., 2001).

Note that the ratio  $\left(\frac{S}{\pi_{m,1}}\right)$  in Eq. (21) is also known as the Pal-Bezdek cluster validity index (Pal & Bezdek, 1995).

(3)- Among several CV indices suggested by Geva and colleagues (Geva et al., 2000), the invariant index  $CV_{GV}$  was selected here to measure the ratio of the between-cluster scatter matrix to the within-cluster scatter matrix (Geva et al., 2000):

$$CV_{GV} = \frac{K_1}{c^2 \cdot J_1}. \quad (22)$$

The normalisation with the number of clusters  $c$  minimise the monotonically increase of  $CV_{GV}$  when  $c$  increased. This index should be maximal at the optimal  $c$ -partition.

(4)- The Kim-Park index, noted  $CV_{KP}$  (Kim, Park & Park, 2001):

$$CV_{KP} = \frac{\pi_{1,1}}{c} + \frac{1}{\alpha} \cdot \left( \frac{c}{V_{dmin}} \right). \quad (23)$$

The best  $c$ -partition is obtained by minimising the index  $CV_{KP}$  with respect to the number of clusters  $c$ . This index has previously been used for fMRI analysis (Moller et al., 2002).

(5)- The Pakhira-Bandyopadhyay-Maulik index  $CV_{PBM}$  (Pakhira, Bandyopadhyay & Maulik, 2004; Pakhira, Bandyopadhyay & Maulik, 2005):

$$CV_{PBM} = \frac{\alpha}{c} \cdot \frac{V_{dmax}}{J_m}. \quad (24)$$

With  $\alpha$  as a constant term (e.g.,  $\alpha$  was set here to  $n$ ). The best  $c$ -partition is obtained by maximising  $CV_{PBM}$  with respect to the number of clusters  $c$ .

(6)- The Wu-Yang index  $CV_{WY}$  (Wu & Yang, 2005):

$$CV_{WY} = \sum_{j=1}^c \left( \frac{n_{1,j}}{\max_j(n_{1,j})} - \exp\left(-\frac{V_{dmin,j}^2}{S}\right) \right). \quad (25)$$

This index compared the fuzzy partition of each cluster to its exponential separation, with  $-c < CV_{WY} < c$ , and  $CV_{WY}$  is maximal at  $c_{opt}$ .

(7)- The Bouguessa-Wang-Sun index  $CV_{BWS}$  index (Bouguessa, Wang & Sun, 2006):

$$CV_{BWS} = \frac{K_m}{\pi_{m,m}}. \quad (26)$$

This index  $CV_{BWS}$  should be maximised with respect to  $c$ .

(8)- Our new  $CV$  index, noted  $CV_{new}$ , combined different measures of compactness and separation as following:

$$CV_{new} = K_m \cdot \left( \frac{ID_{inter}}{ID_{intra}} \right) \cdot \left( \frac{FC}{J_1} \right). \quad (27)$$

The best  $c$ -partition should maximise  $CV_{new}$ . The rationale behind incorporating those specific compactness and separation measures ( $ID_{inter}$ ,  $K_m$ ,  $FC$ ,  $ID_{intra}$ ,  $J_1$ ) in the definition of  $CV_{new}$  is illustrated below with simulated (noisy) datasets.

### Simulated data

Twenty-two simulated datasets were generated as following. First, a fixed number  $c$  of time-courses with  $p$  datapoints ( $p = 100$ ) were generated from a unit normal distribution (mean = 0,  $\sigma = 1$ ). The Pearson correlation between these  $c$  time-courses was less than 0.1 for all simulated datasets. Second, each time-course was replicated  $r_j$  times, with  $j = 1 \dots c_{opt}$ , and  $\sum_{j=1}^c r_j = n$  (where  $n$  is the total number of voxels, set here to 1,000). Third,  $n$  random timecourses with  $p$  datapoints, generated from a normal distribution (mean = 0) but with variable noise levels ( $\sigma = 1$  or 4) were added to the replicated time-courses. This would help to test the robustness of FCM at different noise levels (for a similar rationale see Kim, Park & Park, 2001; Wang & Zhang, 2007) and to monitor the behaviour of the different  $CV$  indices when the fuzzy compactness of clusters became very low (i.e., high intra-class dissimilarity in noisy data). This procedure generated a dataset  $X$  of  $n$  voxels, each with  $p$  datapoints, with known and fixed numbers of classes. Multidimensional scaling (MDS) tools were used to visualise the simulated  $c$  clusters.

Specifically, the following 22 datasets were generated: (i) a single-cluster dataset (noted 1-cluster; e.g., the 'null' case, see Tibshirani, Walther & Hastie, 2001) with highly similar voxels ( $c_{opt} = 1$ ; Fig. S2A); (ii) a dataset without any obvious structure (noted  $n$ -cluster data;  $c_{opt}$  near to  $n$ ; Fig. S2B), see Suleman (2017); (iii) ten datasets with known number of clusters  $c_{opt}$  varying from 2 to 11 and low noise level ( $\sigma = 1$ , see illustration in Fig. S2C with  $c_{opt} = 3$ ); (iv) ten datasets with a known number of clusters  $c_{opt}$  varying from 2 to 11 and high noise level ( $\sigma = 4$ , see illustration in Fig. S2D with  $c_{opt} = 3$ ).

All simulated datasets were clustered by FCM with  $c$  (i.e., number of expected clusters) varying between  $c_{min} = 2$  and  $c_{max} = 19$ . All analyses were carried out with homemade Matlab-based scripts (MathWorks, Natick, MA, USA).

### Real fMRI data

Real data consisted of single subject fMRI data with a block paradigm design (freely available at: <http://www.fil.ion.ucl.ac.uk/spm/data/auditory.html>). The block paradigm consisted of



alternated epochs between rest and auditory stimulation. 96 volumes were acquired on a modified 2T Siemens MAGNETOM Vision system (TR = 7s, 64 contiguous slices). To avoid T1 effects in the initial scans, the first 12 scans were discarded, leaving 84 scans for further analysis ( $p = 84$ ). The data were realigned, normalised (voxel size  $2 \times 2 \times 2$ ) and smoothed (FWHM =  $6 \times 6 \times 6$  mm). This dataset was selected because it has been used in many previous studies with clustering techniques including FCM (e.g., [Gu et al., 2005](#); [Lu, Jiang & Zang, 2004](#)). FCM was applied on this real fMRI dataset with  $c$  varying between  $c_{\min} = 2$  and  $c_{\max} = 39$ . To identify relevant FCM cluster(s) with activated auditory regions, the centroids (prototype)  $V_j$  ( $j = 1 \dots c$ ) were correlated with the experimental block design ([Bandettini et al., 1993](#)).

To appreciate the distribution of brain regions' sizes in each  $c$ -partition, a morphological granulometry was applied to all identified clusters after binarization ([Soille, 2003](#)). This analysis estimated the size of each spatially distinct region or blob (26-connected neighbourhood) for a given crisp FCM partition. Given that each voxel belongs to all clusters at different degrees of membership (cf.  $U_{ij}$  in [Eq. \(2\)](#)), the threshold was set to 0.5 so that each voxel belongs maximally to one cluster. Practically, for a given  $c$ -partition and for each binary cluster  $j$  ( $j = 1 \dots c$ ), the size of each region as well as the number of isolated voxels (i.e., single-voxel regions) were calculated.

This dataset was also analysed with SPM12 software package (Wellcome Trust Centre for Neuroimaging, London UK; <http://www.fil.ion.ucl.ac.uk/spm/>) using standard procedures. This allowed auditory activations to be identified using model-based methods.

### Degree of fuzziness $m$

The degree of fuzziness  $m$  might influence the output of clustering (e.g., [Bezdek, 1981](#); [Fadili et al., 2000](#); [Fadili et al., 2001](#); [Krishnapuram & Keller, 1993](#); [Selim & Ismail, 1986](#); [Yu, Cheng & Huang, 2004](#)): when  $m$  tends to 1 the classification becomes crisp and  $U_{ij}$  takes the value 0 (voxel  $i$  is not a member of cluster  $j$ ) or 1 (voxel  $i$  belongs to cluster  $j$ ) but when  $m$  tends to  $+\infty$  the classification is purely fuzzy ( $U_{ij}$  is near to  $1/c$ ). The optimal value of  $m$  may depend on the characteristics of the data. Previous empirical work approximated  $m$  by a nonlinear function of the dimensions of the data ( $n$  and  $p$ ); for example, [Eq. \(5\)](#) of [Schwämmle & Jensen \(2010, Page 2845\)](#) yields  $m$  values of 1.044 and 1.019 for our artificial and real datasets respectively. However, these estimated values are too low compared to typical  $m$  values encountered in neuroimaging studies. Previous studies have explored the influence of  $m$  on the computation of CV indices (e.g., [Zhou, Fu & Yang, 2014](#)), and they found better clustering results with  $m$  between 1.2 and 2.5 for fMRI data ([Fadili et al., 2000](#); [Fadili et al., 2001](#); [Moller et al., 2002](#); [Smolders et al., 2007](#)).

More specifically, there are two issues to be considered when selecting  $m$  during the computation of CV indices. First, several CV indices became inadequate with hard  $c$ -partitions (i.e.,  $m$  tends to 1). Specifically, any measures that are based exclusively on the distribution of  $U$  values (e.g.,  $FC$ ,  $FS$ ) would artificially reach their optimal values independently from the number of clusters  $c$ . Second, according to [Eq. \(3\)](#), centroids become close to the mean of the whole data set  $\bar{X}$  when  $m$  tends towards  $+\infty$ . In other words, the  $c$  clusters would have comparable fuzzy cardinality values (i.e., [Eq. \(6\)](#)) for larger

$m$  values, which may be problematic when some clusters are expected to contain a small number of voxels (see illustration in Fig. S3); for more details see (Selim & Ismail, 1986; Tsekouras & Sarimveis, 2004; Yu, Cheng & Huang, 2004). This issue is particularly critical when analysing task-related fMRI data because activated voxels are expected to represent a small fraction of the whole brain.

Here,  $m$  was held to 1.5 throughout this study.

### Voxel selection and the ill-balanced dataset problem in fMRI

One important issue during the clustering of fMRI datasets is the selection of the relevant  $n$  voxels. Because the number of activated voxels is small (i.e., a few percent) compared to the total number of voxels in a typical whole-brain fMRI dataset, previous studies have suggested different approaches to overcome this 'ill-balanced' data problem. For instance, FCM can be limited to relevant voxels within the gray matter, in specific anatomical brain regions, or to voxels with some kind of task-related effects (e.g., see Fadili et al., 2000; Goutte et al., 1999; Gu et al., 2005; Lee et al., 2012; Moller et al., 2002; Seghier, Friston & Price, 2007). Voxel selection might be useful for: (i) reducing the high dimensionality of the problem and improving both computational robustness and speed; (ii) minimising the influence of redundant voxels; and (iii) increasing the accuracy of the clustering by focusing mainly on meaningful voxels. However, the author preferred here to include all brain voxels so that the robustness of the different CV indices can be appreciated when noisy voxels (voxels with no effect of interest) and artefacts are present. FCM was thus applied to all voxels of the real fMRI dataset, yielding a total number of voxels  $n = 227,716$ .

### FCM convergence and the initialisation problem

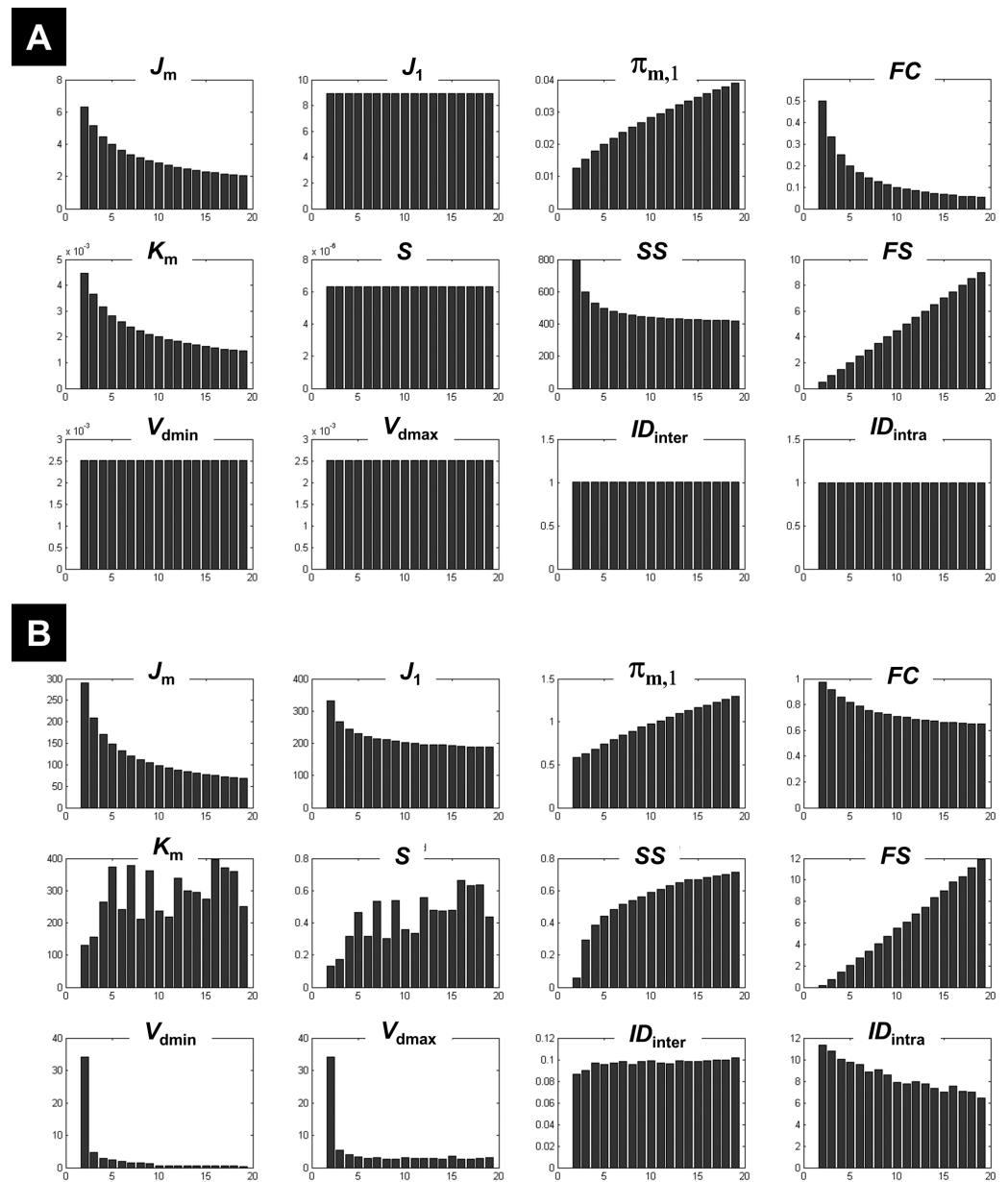
Depending on the initialisation of the degrees of membership  $U$  (Bezdek, 1981), the FCM algorithm may converge to different  $c$ -partitions (e.g., local minima). This problem of initialisation may lead to spurious  $c$ -partitions (Moller et al., 2002) when using CV indices. One possible solution is to repeat the FCM algorithm on the same dataset with several different random initialisations (e.g., Moller et al., 2002; Pena, Lozano & Larranaga, 1999), with the expectation that it is unlikely that different starting conditions will lead to the same local minima. Accordingly, for each  $c$  value, the FCM algorithm was re-run on the real fMRI dataset ten times with random initialisations (for a similar procedure see Chuang et al., 1999).

## RESULTS

### FCM on simulated data

#### The 1-cluster dataset

Clustering the 1-cluster dataset ( $c_{\text{opt}} = 1$ ) showed how compactness and separation measures behave when data cannot be clustered any further. In this context of high redundancy, it is expected to observe: (i) high similar or identical centroids  $V$ , (ii) degrees of membership  $U$  near to the fuzziest value  $1/c$ , and (iii) comparable fuzzy cardinality across clusters. As illustrated in Fig. 1A, the fuzzy compactness  $FC$  decreased



**Figure 1** Illustration of the behaviour of different measures of compactness and separation. FCM on the one-cluster (A) and the  $n$ -cluster (B) dataset. The number of clusters varied between 2 and 19. See full definition of the different measures in the ‘Methods’.

Full-size DOI: 10.7717/peerj.5416/fig-1

monotonically with  $c$  (i.e.,  $FC = 1/c$ ) whereas fuzzy separation  $FS$  increased linearly with  $c-1$  (i.e.,  $FS = (c - 1)/2$ ), suggesting a high fuzzy overlap  $FO$  between clusters. Likewise, as expected, the fuzzy compactness  $\pi_{m,1}$  and separation  $K_m$  showed monotonic dependency with  $c^{m-1}$  and  $c^{1-m}$  respectively, suggesting that the product  $\pi_{m,1} \cdot K_m$  remained constant (independent from  $c$ ) when data were classified into pure fuzzy clusters. Interestingly, measures of separation based on centroids  $V$  (e.g.,  $V_{dmin}$ ,  $V_{dmax}$ ,  $S$ ) and distances  $D$  (e.g.,

$ID_{intra}$  and  $ID_{inter}$ ) were independent from  $c$ , suggesting highly similar (i.e., identical) centroids  $V$ .

### **The $n$ -cluster dataset**

Clustering the  $n$ -cluster dataset (i.e.,  $c_{opt}$  towards  $n$ ) tested the robustness of the different measures of compactness and separation when data is patternless with high dispersion. Compactness coefficients showed similar behaviour as above when clustering the 1-cluster dataset, except for  $J_1$  and  $SS$  measures. Interestingly, separation measures based on centroids  $V$  and distances  $D$  showed more complex dependencies with  $c$  (Fig. 1B) as compared to the 1-cluster case (Fig. 1A), in particular when using the two new coefficients  $ID_{intra}$  and  $ID_{inter}$ .

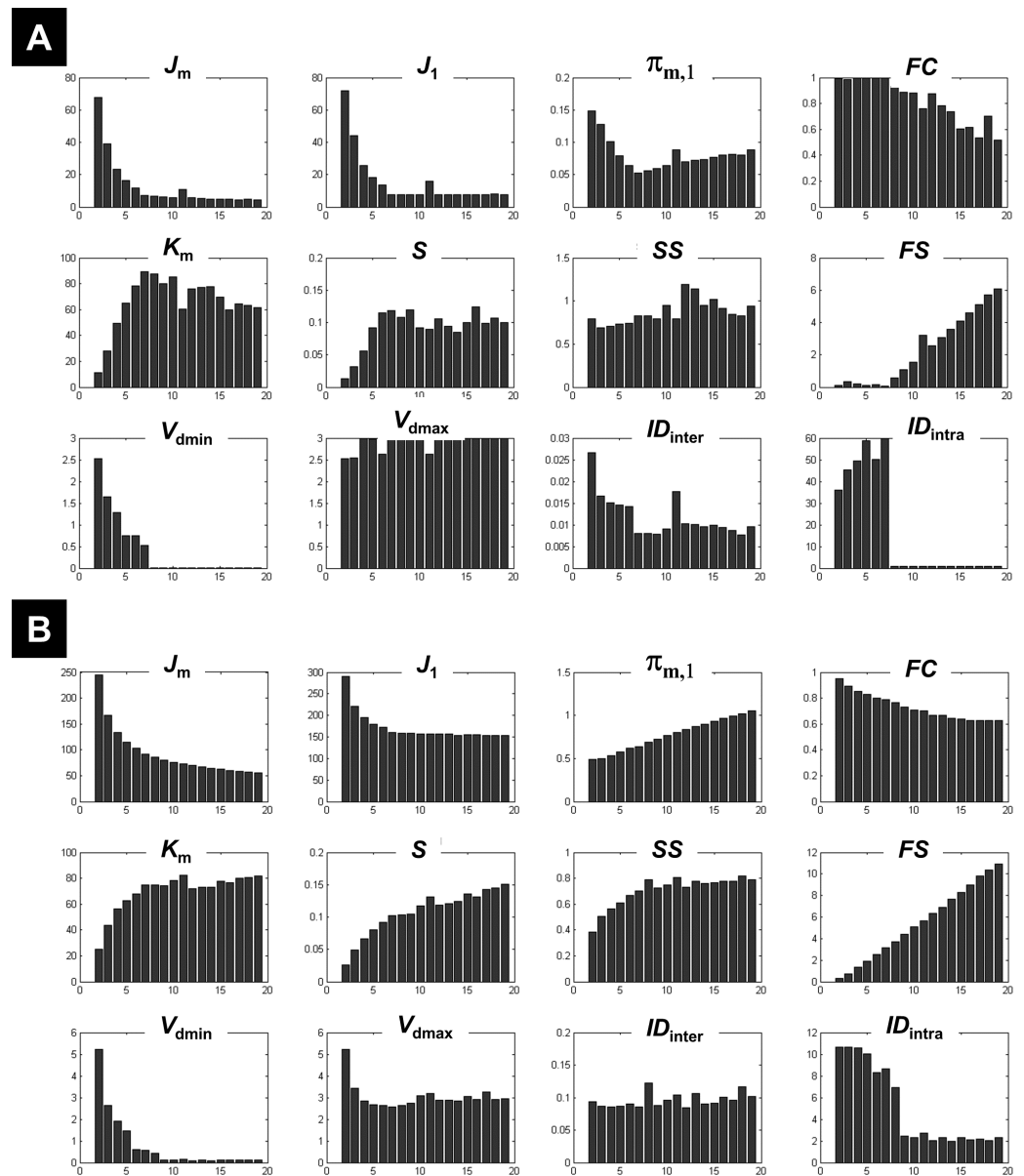
What emerged from above is that  $K_m$ ,  $ID_{intra}$ ,  $ID_{inter}$ , and  $J_1$  behaved differently on 1-cluster and  $n$ -clusters datasets, which is highly desirable when clustering fMRI data that have complex structure. These results motivated the rationale of including them in the computation of the new  $CV_{new}$  index (as defined in Eq. (27)).

### **CV indices on data with known numbers of clusters**

The different measures of compactness and separation are shown in Fig. 2A for the 7-clusters data set. Several measures showed different values over the number of clusters as compared to the clustering of the 1-cluster and  $n$ -cluster datasets. For instance, the coefficient  $J_1$  decreased in the interval  $c = 2$  to  $c = 7$ , consistent with the fact that data can be clustered further (as seen for the  $n$ -cluster data); then it reached a plateau for higher number of classes, consistent with the fact that the data cannot be segregated any further (as the case of the 1-cluster data). The limit between the two behaviours was indeed at the true number of clusters ( $c = 7$ ). This observation is valid for the other measures of compactness (e.g.,  $FC$ ,  $\pi_{m,1}$ ,  $ID_{intra}$ ) and separation (e.g.,  $FS$ ,  $ID_{inter}$ ,  $S$ ,  $K_m$ ). When the data became noisy, some measures were less sensitive to the structure of the data (i.e., the presence of seven clusters). As illustrated in Fig. 2B, fuzzy separation  $FS$  and compactness  $\pi_{m,1}$  showed comparable behaviour as in the clustering of the  $n$ -cluster dataset, which reflects the influence of noisy distant points (low within-cluster compactness and between-cluster separation). Interestingly, in addition to  $V_{dmin}$ , quantities  $ID_{intra}$ ,  $K_m$  and  $J_1$  were more robust to noise and showed high discriminability with an optimal value around the expected number of classes (Fig. 2B). This observation further motivated their inclusion in the definition of the new CV index.

Figure 3 illustrates all CV indices for the 3-cluster, 7-cluster, and 11-cluster datasets with low noise level ( $\sigma = 1$ ). All CV indices indicated the best  $c$ -partition for the expected number of clusters (maximum value for  $CV_{ZLE}$ ,  $CV_{GV}$ ,  $CV_{PBM}$ ,  $CV_{WY}$ ,  $CV_{BWS}$ ,  $CV_{new}$ ; minimum value for  $CV_{RLR}$  and  $CV_{KP}$ ). Note that the new index  $CV_{new}$  is highly discriminative in pointing to the optimal  $c$ -partition. When the data became noisy ( $\sigma = 4$ ), all CV indices, except  $CV_{BWS}$  and  $CV_{new}$ , failed to indicate the optimal  $c$ -partition (Fig. 4). However, for data with higher  $c_{opt}$  (e.g.,  $c_{opt} > 9$ ), only the new index  $CV_{new}$  identified the true number of clusters, albeit with lower discriminability (e.g., compare Figs. 3B to 4B).

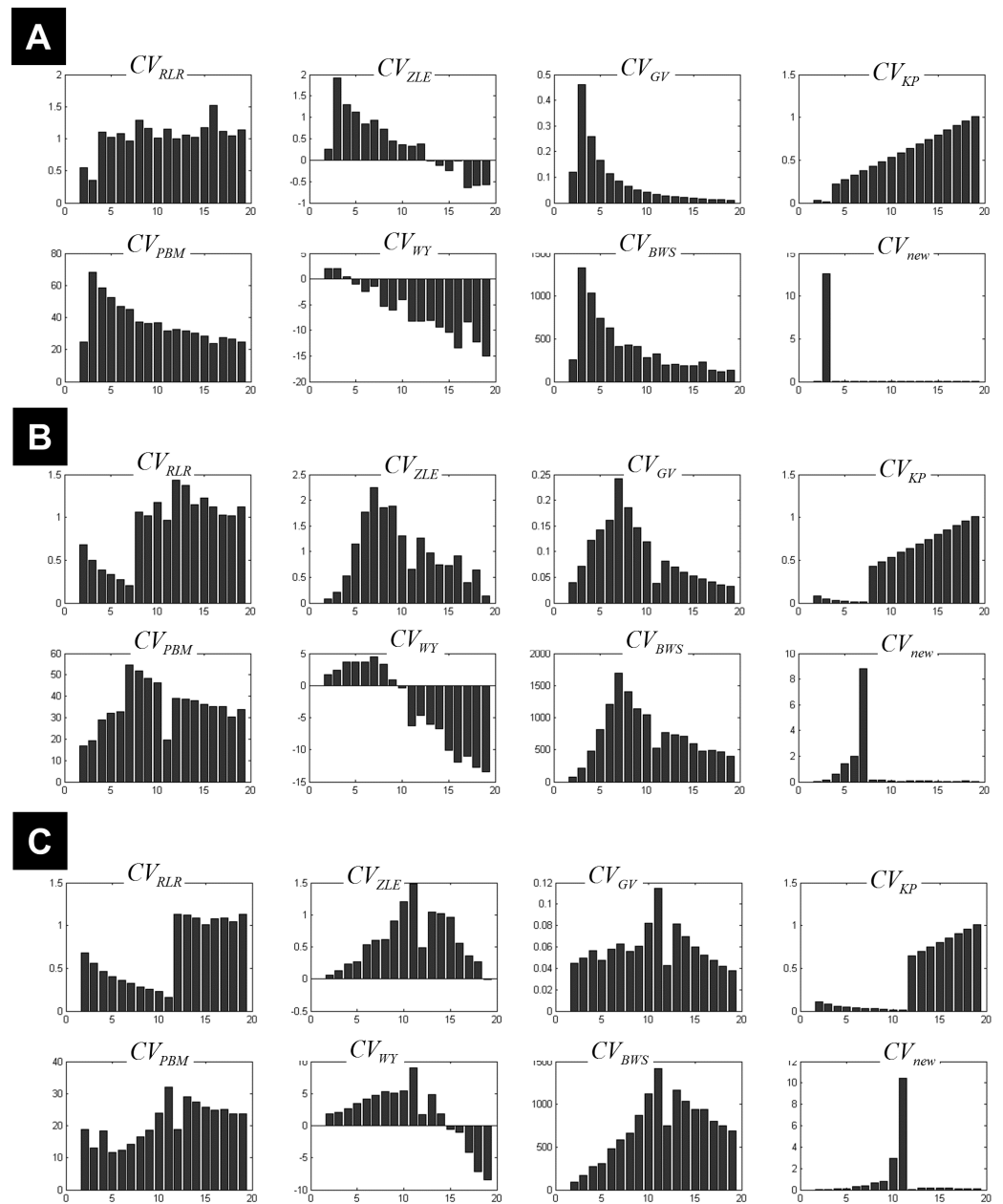
An ad hoc analysis was conducted to monitor the behaviour of  $CV_{new}$  over different degrees of fuzziness  $m$  ( $m$  varying between 1.2 and 2.5), for a similar rationale see



**Figure 2** Illustration of the behaviour of different measures of compactness and separation at different noise levels. The behaviour of different measures of compactness and separation during FCM of the 7-cluster dataset with low (A,  $\sigma = 1$ ) and high (B,  $\sigma = 4$ ) noise levels. See full definition of the different measures in the ‘Methods’.

Full-size  DOI: [10.7717/peerj.5416/fig-2](https://doi.org/10.7717/peerj.5416/fig-2)

(Schwämmle & Jensen, 2010). This analysis showed that  $CV_{new}$  correctly identified the true number of clusters  $c_{opt}$  in almost all simulated datasets for  $m \in [1.2, 2.5]$ , except for datasets with both high noise level ( $\sigma = 4$ ) and high number of true clusters ( $c_{opt} > 9$ ) where  $CV_{new}$  failed to identify  $c_{opt}$  when  $m \geq 2$  (i.e.,  $CV_{new}$  underestimated  $c_{opt}$  at higher  $m$  values including the popular value of  $m = 2$ ). This ad hoc analysis confirmed the initial choice of  $m = 1.5$ .

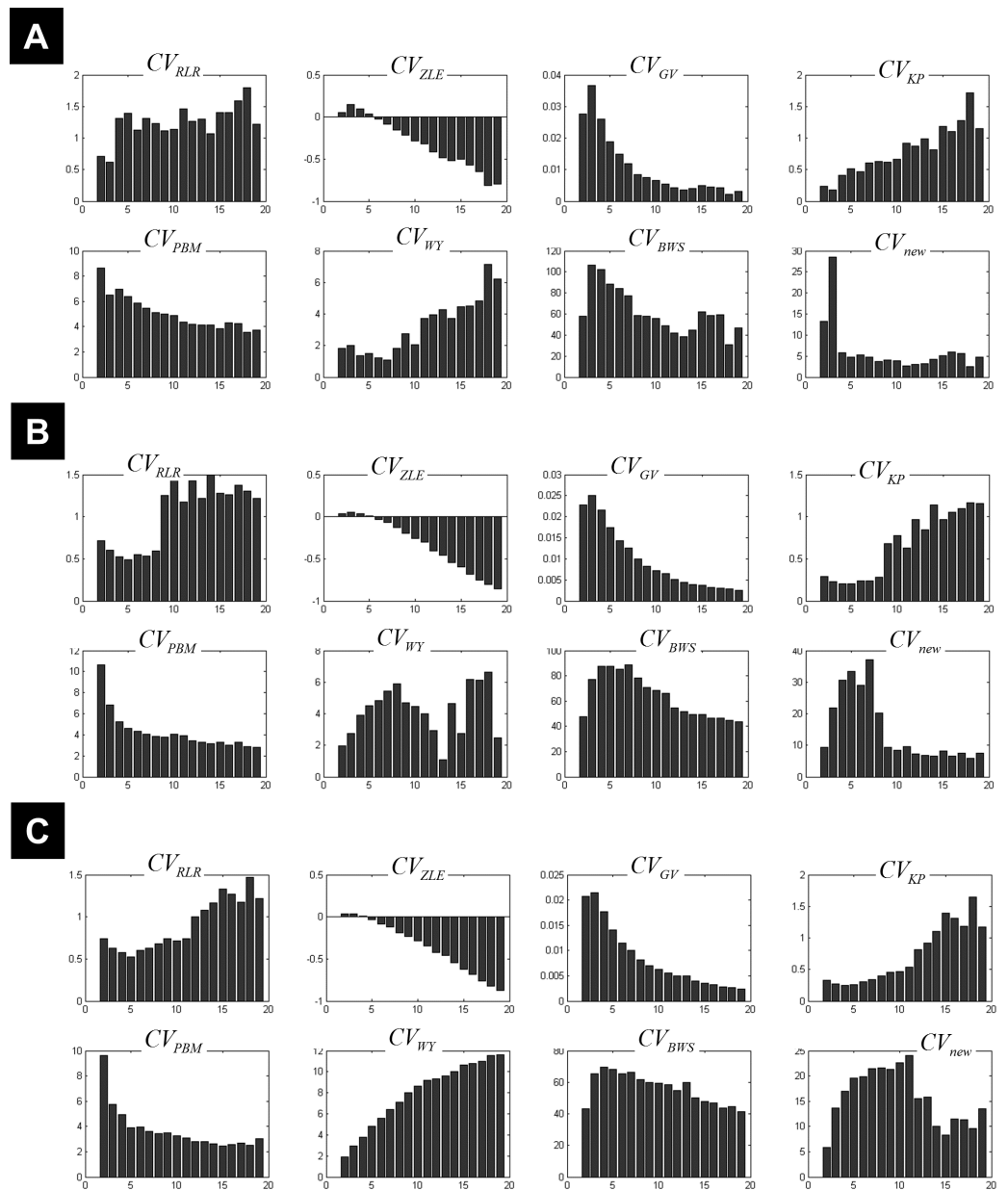


**Figure 3** Plots of the CV indices for the simulated data at low noise level. Plots of the CV indices for the simulated 3-cluster (A), 7-cluster (B) and 11-cluster (C) datasets with low noise level ( $\sigma = 1$ ), when number of clusters increased from 2 to 19. All CV indices successfully indicated the expected number of clusters ( $c_{\text{opt}} = 3$  in a,  $c_{\text{opt}} = 7$  in b and  $c_{\text{opt}} = 11$  in c). See full definition of these indices in the ‘Methods’.

Full-size DOI: [10.7717/peerj.5416/fig-3](https://doi.org/10.7717/peerj.5416/fig-3)

### FCM real fMRI data

As expected, the number of iterations for the convergence of the FCM algorithm varied across the 10 different initialisations. However, for a given  $c$  value and across the ten runs, the obtained  $c$ -partitions were very similar and the function  $J_m$  (Eq. (1)) reached the same



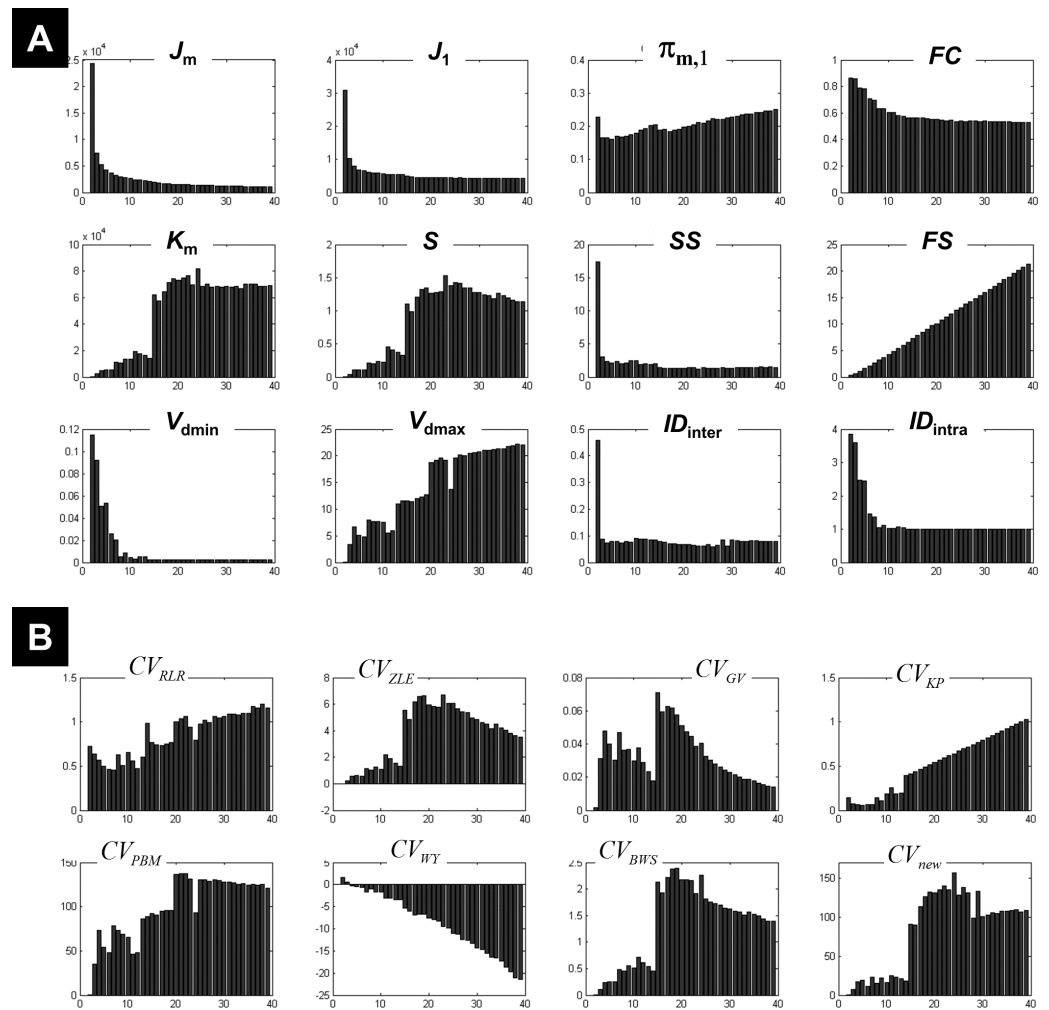
**Figure 4** Plots of the CV indices for the simulated data at high noise level. Plots of the CV indices for the simulated 3-cluster (A), 7-cluster (B) and 11-cluster (C) datasets with high noise levels ( $\sigma = 4$ ), when the number of clusters increased from 2 to 19. Only the new CV index identified the correct 11-partition at this level of noise.

Full-size DOI: [10.7717/peerj.5416/fig-4](https://doi.org/10.7717/peerj.5416/fig-4)

minimum value (except for  $c$  values between 12 and 15 where one initialisation reached a different minimal  $J_m$  value compared to the other nine initialisations).

### Identified clusters

Figure 5 plots the different coefficients and CV indices against the number of expected clusters  $c$  varying from 2 to 39. Measures such as  $ID_{inter}$ , and  $V_{dmin}$  showed an interesting



**Figure 5** Illustration of the results using real fMRI data. (A) Different measures of compactness and separation and (B) the different CV indices. The number of clusters varied from 2 and 39.

Full-size [DOI: 10.7717/peerj.5416/fig-5](https://doi.org/10.7717/peerj.5416/fig-5)

pattern when  $c$  increased, with high and decreasing values for small number of clusters ( $c < 10$ ) and low and fixed values when  $c$  increased (a comparable behaviour was also seen for  $FC$ ). This mirrored their behaviour during the clustering of the 1-cluster and  $n$ -cluster datasets. The change in the nature of the dependency occurred around  $c = 13$ , indicating the maximum  $c$  value that ensured different centroids  $V$ . For a number of expected clusters bigger than 13, the  $c$ -partition contained a few redundant classes (identical centroids  $V$ ). However, for  $c < 13$  clusters, although the obtained classes were compact (e.g., high  $FC$  values), the separation between clusters was not optimal (see for instance  $V_{dmax}$ ,  $S$ , and  $K_m$ ). More specifically, the fuzzy separation measures  $S$  and  $K_m$  showed optimal values for higher numbers of expected clusters at  $c$  larger than 17 clusters. At this range, the  $c$ -partition contained at least three similar centroids.

**Figure 5B** illustrates the dependency of different CV indices with  $c$ . Some CV indices (e.g.,  $CV_{RLR}$  and  $CV_{KP}$ ) showed optimal values for low  $c$  values (maximal fuzzy compactness),



whereas other CV indices (e.g.,  $CV_{ZLE}$ ,  $CV_{BWS}$ ,  $CV_{GV}$ , and  $CV_{PBM}$ ) showed optimal values at an intermediate number of expected clusters (i.e., maximal fuzzy separation). Interestingly, the new index  $CV_{new}$  went through different phases (i.e., different plateaus), depending on the weight of fuzzy separation and compactness (a change of behaviour visible at  $c = 15$ ). The new index  $CV_{new}$  reached its maximum value at  $c = 24$  clusters, ensuring a good compromise between separation and compactness of the  $c$ -partition of this real fMRI dataset.

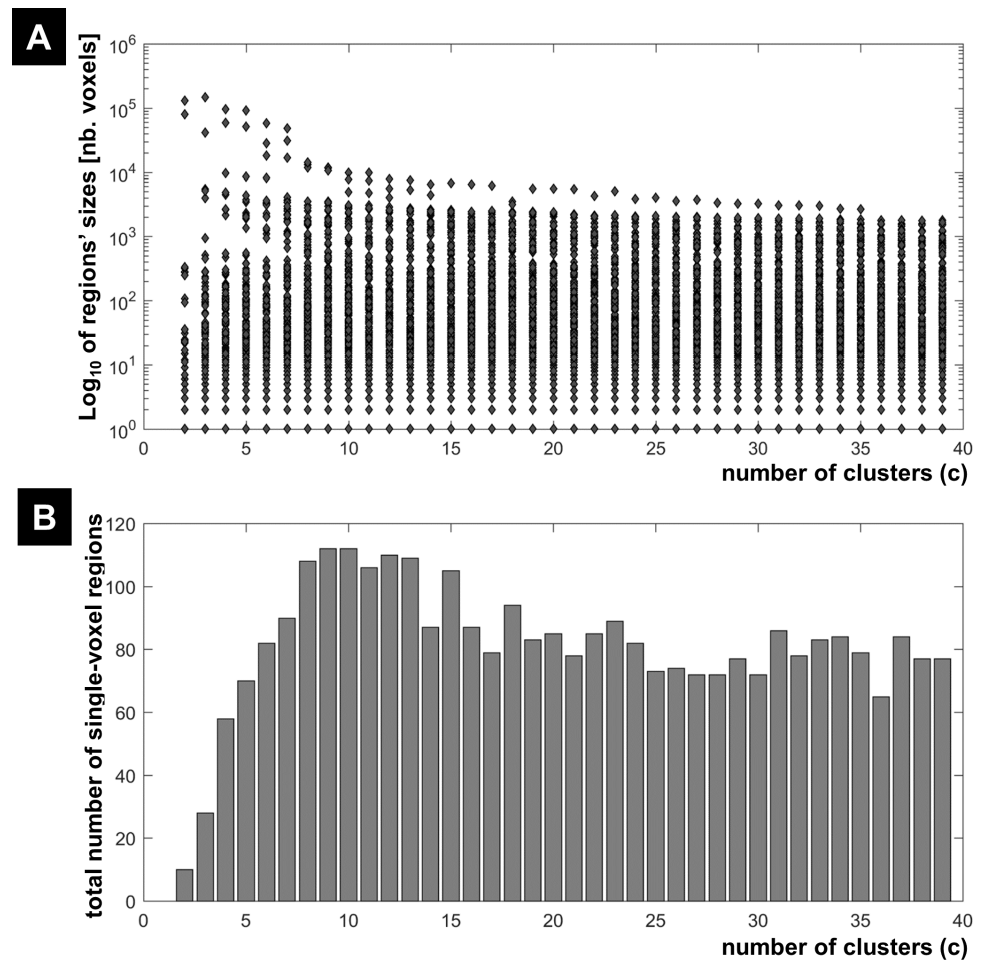
The results of the morphological granulometry at  $U > 0.5$  are illustrated in Fig. 6. As expected, the size of very large regions tend to decrease with the number of expected clusters, as large regions were subdivided further into smaller regions at higher  $c$  values. Interestingly, for each  $c$ -partition, the total number of single-voxel regions over all clusters was less than 0.04% of  $n$  (Fig. 6). Given the spatial smoothness of the fMRI data, there was no cluster containing exclusively single-voxel regions.

Figure 7 illustrates all obtained clusters for  $c$ -partitions with low fuzzy separation ( $c = 8$ ), without redundant clusters ( $c = 13$ ), at high fuzzy separation ( $c = 18$ ), and at the optimal  $c$  value that maximised  $CV_{new}$  ( $c = 24$ ). Identified voxels within the auditory cortex (i.e., voxels of interest) are shown in the first axial slice of each  $c$ -partition. Voxels in the auditory cortex were grouped with those in the occipital lobe at small  $c$  values ( $c = 8$ ), but they became clearly segregated at larger  $c$  values (e.g.,  $c = 18$  and  $c = 24$ ). Interestingly, identified voxels within the auditory cortex in the  $c$ -partition with 24 clusters were remarkably similar to those identified with model-based SPM methods (e.g., SPM map at  $p < 0.05$  FWE-corrected, Fig. 8). Last but not least, the centroid of the relevant cluster with activations in auditory regions (Cluster “1” of the 24-partitions in Fig. 7) was strongly correlated with the experimental block design ( $r = 0.7$ ,  $p < 0.001$ ).

## DISCUSSION

Using both simulated and real fMRI data, this study explored the usefulness of CV indices in identifying the best  $c$ -partition with FCM. This study also examined the behaviour of different compactness and separation measures, defined here as building blocks of the different CV indices. The optimal number of clusters varied with different CV indices, given that measures of compactness and separation were influenced by different features of the fMRI data (e.g., the expected high number of clusters, noise, and the amount of artefacts). A new CV index ( $CV_{new}$ ) was introduced here and it showed relatively good robustness when clustering noisy data with high number of classes. Our study also highlighted the importance of analysing different measures of separation and compactness in order to get a better understating of the complex structure of the data.

The typical low signal-to-noise ratio in fMRI might be the most challenging issue that can hinder the success of clustering techniques. Here, simulated data were based on Gaussian-like noise distributions, and the success of different CV indices depended on the level of noise in the data. Our findings are in line with previous studies that compared several CV indices on different simulated datasets and found that CV indices may fail to indicate the true number of clusters in noisy data that have high number of

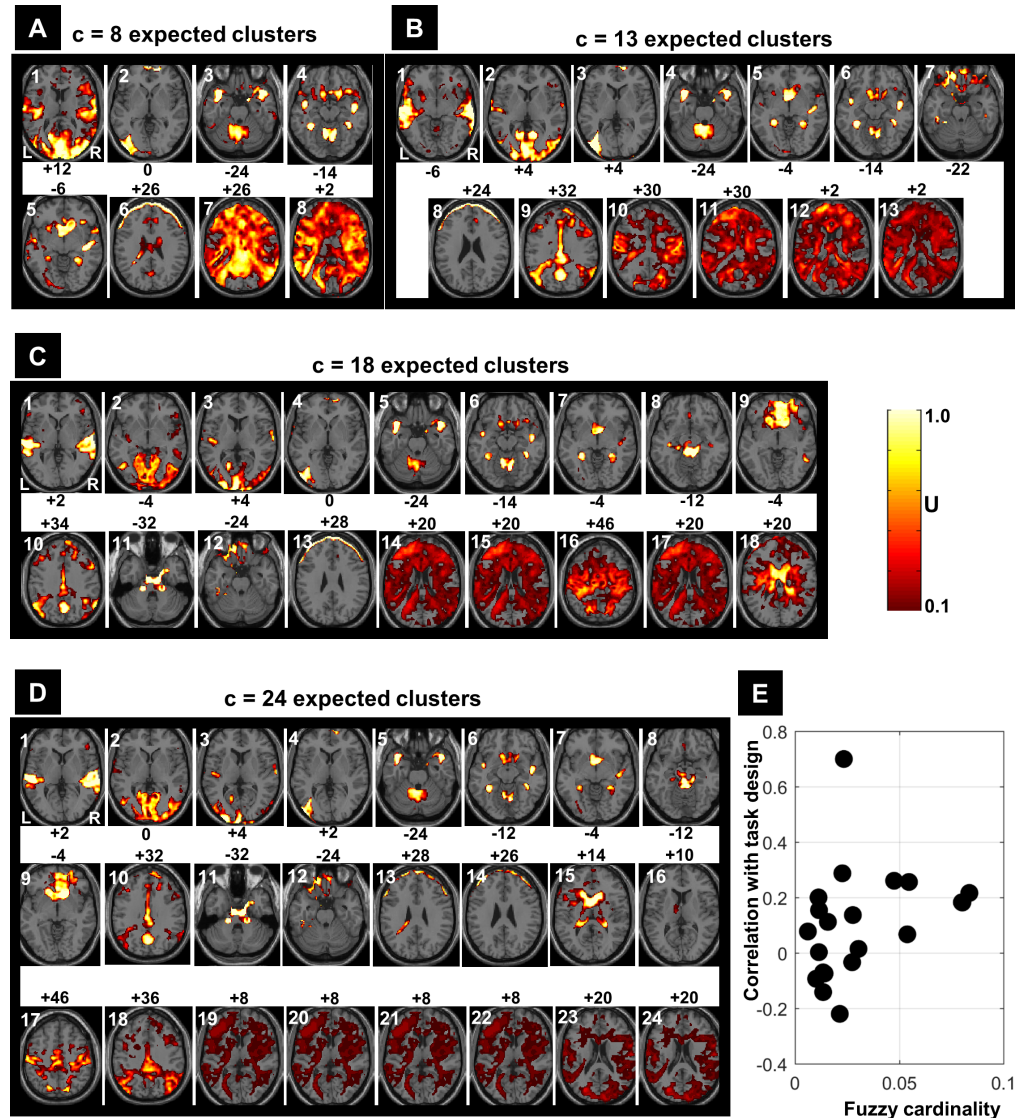


**Figure 6** FCM results at different  $c$  values. (A) Regions' sizes (in number of voxels) for each crisp  $c$ -partition (at an arbitrary threshold of  $U > 0.5$ ). Each dot (diamond shape) represents the size of one region in any cluster of the  $c$ -partition ( $c$  varying between 2 and 39). A base 10 logarithmic scale is used for the y-axis. (B) Total number of single-voxel regions for each  $c$ -partition. For the winning FCM partition ( $c = 24$  with  $CV_{\text{new}}$ ), there was less than 4 single-voxel regions per cluster on average. Total number of voxels  $n = 227,716$ ; voxel size = eight  $\text{mm}^3$ .

Full-size [DOI: 10.7717/peerj.5416/fig-6](https://doi.org/10.7717/peerj.5416/fig-6)

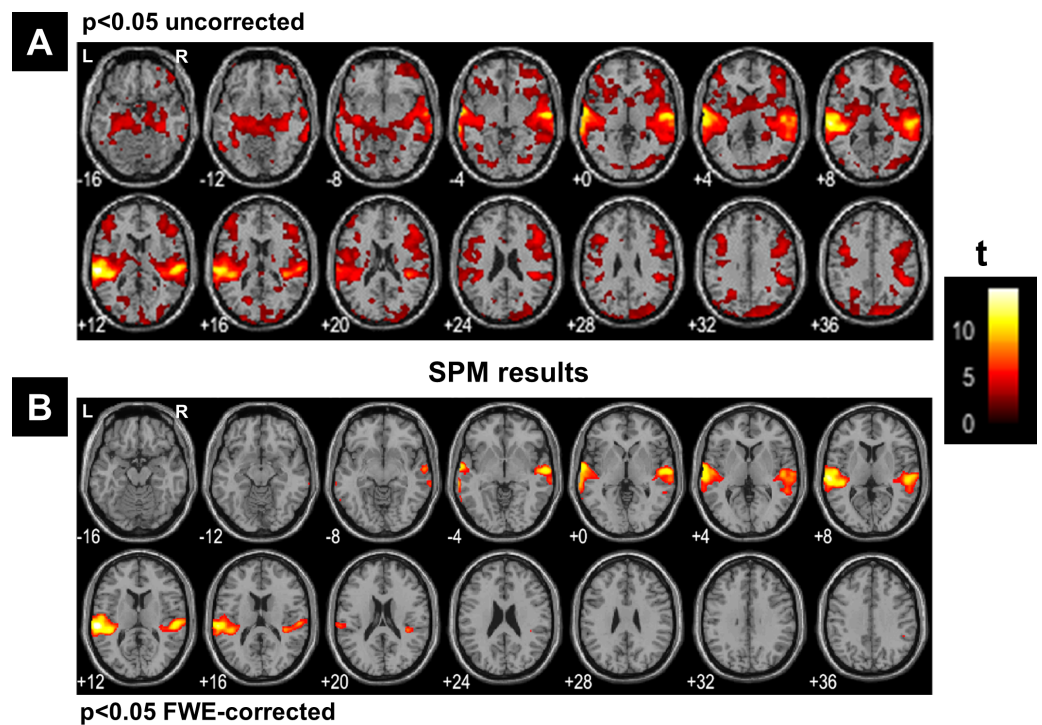
classes (Suleman, 2017; Wang & Zhang, 2007; Zhou et al., 2014). It might be the case their effectiveness might even be lesser given the complex nature of noise in MRI images with significant correlations between voxels (Gudbjartsson & Patz, 1995; Parrish et al., 2000). To ensure better data input to FCM, it is thus recommended to use different pre-processing techniques that can reduce the impact of noise and improve data quality (Caballero-Gaudes & Reynolds, in press). The usefulness of such techniques with FCM on fMRI data warrants further studies.

Perhaps more importantly, the results stressed the importance of reading the behaviour of different separation and compactness measures, defined here as building blocks of CV indices, in order to depict an accurate description of the fMRI data (cf. Fig. 5). This is because it is most likely that there are different meaningful  $c$ -partitions depending on the



**Figure 7** FCM results at different  $c$  values (A:  $c = 8$ , B:  $c = 13$ , C:  $c = 18$  and D:  $c = 24$ ). Each obtained cluster (a 3D image) of each  $c$ -partition is illustrated by its most representative axial slice, with  $U$  values varying from 0.1 to 1.0. Cluster label is shown at the top-left corner of each axial slice (in white) and the MNI-z coordinate is indicated in black. For illustration purposes, the cluster that contained the expected activated voxels within the auditory cortex is labelled as Cluster '1'. The scatter plot (E) illustrates the correlations between the centroids of the 24-partition and the experimental block design ( $y$ -axis) against the fuzzy cardinality (cf. Eq. (6)) of each cluster ( $x$ -axis). Only one cluster showed significant correlation ( $p < 0.001$ ) with the experimental design ( $r = 0.7$ ). The fuzzy cardinality was divided by the total number of voxels, which would approximately reflect the 'proportion' of voxels contained in each cluster (average proportion around 4% ( $=1/c$ )). Using the spatial location of the clustered voxels, one can potentially interpret the results of the FCM 24-partition (D). For example, Cluster 1 is showing auditory activations (cluster of interest) that highly correlated with the experimental block design ( $r = 0.7$ ); Clusters 2–4 illustrate voxels in the visual system; Clusters 5–8 illustrate cerebellar and subcortical regions; Clusters 9–10 illustrate different medial parts of the default mode network; Clusters 11–12 contain voxels in ventral brain regions that are prone to MR signal loss; Clusters 13 and 14 are dominated by motion artefacts; Cluster 15 mainly shows CSF voxels; Clusters 19–24 contain white matter voxels. L, left hemisphere; R, right hemisphere.

Full-size [DOI: 10.7717/peerj.5416/fig-7](https://doi.org/10.7717/peerj.5416/fig-7)



**Figure 8** SPM's results. SPM results illustrated with the function 'montage' of SPM12, with axial slices varying between MNI- $z = -16$  mm to MNI- $z = +36$  mm. (A) Results at a very liberal threshold of  $p < 0.05$  uncorrected, (B) at  $p < 0.05$  FWE-corrected. L, left hemisphere; R, right hemisphere.

Full-size [DOI: 10.7717/peerj.5416/fig-8](https://doi.org/10.7717/peerj.5416/fig-8)

scale at which the different clusters (i.e., brain networks) are segregated. Accordingly, it is not always useful to bias the analysis towards one elusive single  $c$ -partition, but rather appreciate that fMRI data might encompass different plausible patterns or networks at different spatio-temporal scales (Orban et al., 2015). Put another way, users need to relax the assumption that  $c_{\text{opt}}$  must be unique, and look instead for complementary explanations of the data at different  $c_{\text{opt}}$  values. For instance, using fuzzy clustering on resting-state fMRI data, Lee and colleagues (Lee et al., 2012) identified two optimal  $c$ -partitions with seven and eleven clusters that minimised a cluster dispersion measure (used as a CV index). Interestingly, the  $c$ -partition with 11 clusters further subdivided some of the clusters identified in the  $c$ -partition with seven clusters (Lee et al., 2012), most probably due to the known hierarchical organization of the brain networks. Our results of the clustering of real fMRI data also showed similar trends with clusters being further segregated with increasing number of expected clusters (e.g., compare clusters with  $c = 8$  to clusters with  $c = 18$  in Fig. 7).

Previous work suggested that, when CV indices fail to agree on the true number of clusters for high-dimensional datasets, a combination of different indices into a single index should be considered (Sheng et al., 2005; Zhou et al., 2014). Specifically, by using a weighted sum of several normalized CV indices, it has been shown that this weighted sum can improve the confidence of clustering solutions. Ultimately, this approach aims to

force an agreement between CV indices so that one optimal single  $c$ -partition is selected. However, this approach may not be applicable to all contexts because: (i) the number and types of CV indices to be combined are arbitrary, (ii) there is no objective procedure to set optimal weights, and previous empirical work showed that such weights are data-dependent ([Zhou et al., 2014](#)), (iii) the weighted sum does not properly deal with redundant information, given that CV indices are likely to share similar compactness or separation measures, (iv) the relationships of some CV indices with the number of expected clusters can take any arbitrary shape (e.g., [Fig. 5B](#)), hence linear combinations may not be suitable, and (v) this approach implicitly assumes that there must be one unique ‘true’ explanation of the data. Here I argue that summation of different CV indices might not be useful for fMRI data clustering, because it ignores the possibility that different plausible explanations (different  $c$ -partitions) exist for the same data. Differences between CV indices should not be overlooked because they tend to highlight different existing features in the data.

The existence of different plausible explanations ( $c$ -partitions) of the same fMRI data can be further illustrated when examining the different compactness and separation measures used in the definition of the new CV index. More specifically, as illustrated in [Fig. 5B](#),  $CV_{new}$  went through three different phases: (i) low values for  $c < 15$ , (ii) a plateau with high optimal values for  $15 < c < 28$ , and (iii) another plateau for  $c > 28$ . The three phases indicated different segregated data structures depending on the predominance of either compactness or separation measures ([Fig. 5A](#)). For example, high fuzzy separation with well-isolated clusters was only achieved at  $c > 15$ , as reflected in the behaviour of  $K_m$  and  $ID_{intra}$ ; however, when  $c$  increased the  $c$ -partitions became less compact (see  $FC$ ), with higher fuzzy overlap and over-classification when  $c$  increased beyond 28 clusters (see  $ID_{inter}$ ). Given the expected small proportion of task-related activations in the auditory cortex, a segregation of relevant auditory voxels was only achieved with  $c > 15$  clusters, for a similar rationale see ([Chuang et al., 1999](#)). In sum, looking at different compactness and separation measures, in addition to  $CV_{new}$  index, can provide a richer representation of the clustering results so that users can select the most useful  $c$ -partition among many potential possibilities.

Other methodological issues warrant further investigations. For instance, it might be interesting to test these CV indices with other varieties of FCM algorithms that incorporated spatial constraints during the minimisation of the objective function  $J_m$  (e.g., [Ahmed et al., 2002](#); [Liew, Leung & Lau, 2000](#)), which can take into account the inherent spatial dependencies between neighbouring voxels (e.g., dependencies inflated by the spatial resampling and smoothing in fMRI). This would for instance penalise implausible solutions ( $c$ -partitions) with isolated voxels (e.g., [Fig. 6](#)). In addition, if outlier voxels existed in a dataset, this would artificially yield optimal CV values for  $c$ -partitions with a small number of clusters. In this context, it is useful to combine these CV indices with robust clustering techniques (for a review see [Dave & Krishnapuram, 1997](#)), adaptive distance measures ([Tang et al., 2015](#)), or other modified fuzzy clustering algorithms (e.g., [Dik et al., 2014](#); [Kao & Huang, 2013](#); [Keller, 2000](#); [Seghier, Friston & Price, 2007](#)). Another challenging issue is to give meaning to the different identified clusters. Typically, users have to set objective criteria to distinguish relevant clusters from noise or artefact-driven clusters.

For instance, for task-related fMRI data, clusters of interest are expected to have centroids similar (highly correlated) to the paradigm (*Chuang et al., 1999; Fadili et al., 2000; Goutte et al., 1999; Jahanian, Soltanian-Zadeh & Hossein-Zadeh, 2005*), as illustrated in *Fig. 7*. For task-free fMRI data, irrelevant clusters should be discarded, including clusters that are less consistent across sessions (*Levin & Uftring, 2001*) or when they include irrelevant brain voxels (e.g., in the white matter, ventricles, cerebrospinal fluid, arteries) (*Ma et al., 2011*).

Although FCM can provide useful data-driven explanations, deciding which clustering method is best for fMRI data remains an open question (*Derntl & Plant, 2016*). Typically, selecting a specific clustering algorithm entails a trade-off between different criteria (e.g., accuracy versus stability *Thirion et al., 2014*), with different methods may yield different clustering solutions. Many previous fMRI studies for instance have compared FCM against other data-driven methods, but findings varied considerably across studies, probably due to differences in fMRI data features in particular in terms of contrast-to-noise ratio and the level of physiological noise (*Baumgartner et al., 2000; Dimitriadou et al., 2004; Lange et al., 2006; Wismuller et al., 2004*). One popular data-driven method in the current literature is independent component analysis (ICA). ICA allows the detection of unexpected brain responses to stimuli, dissociation of functional networks and can be used as a powerful denoising tool (*Stone, 2002*). Previous work (*Meyer-Baese, Wismuller & Lange, 2004; Smolders et al., 2007*) have shown that FCM may outperform ICA when analyzing task-related fMRI data with good contrast-to-noise ratio. Nonetheless, it is fair to say that any comparison between ICA and FCM is an empirical question that is contingent on the nature of the fMRI data, the exact parametrization of FCM (*Schwämmle & Jensen, 2010*), the type of ICA algorithm, and the number of independent components (*McKeown, Hansen & Sejnowsk, 2003*).

## CONCLUSIONS

Unsupervised FCM with different CV indices is a useful tool for analysing model-free fMRI datasets, an alternative to the widely used independent component analysis methods. It is recommended to combine different CV indices in order to draw a complete picture of the structure of the data. The assumption here is that different CV indices may point to different optimal  $c$ -partitions, given the heterogeneous behaviour of many measures of compactness and separation. Rather than discarding discrepancies between CV indices, such discrepancies should be appreciated because they reflect the hierarchical organization of brain networks. This was clearly visible for instance when analysing the different phases of the plot of the new CV index against the number of clusters. Overall, the existence of different  $c$ -partitions for the same fMRI data should not be overlooked in future clustering studies.

## ADDITIONAL INFORMATION AND DECLARATIONS

### Funding

This work was supported by the ECAE's Research Office. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### Grant Disclosures

The following grant information was disclosed by the author:  
ECAE's Research Office.

### Competing Interests

The authors declare there are no competing interests.

### Author Contributions

- Mohamed L. Seghier conceived and designed the experiments, analyzed the data, contributed reagents/materials/analysis tools, prepared figures and/or tables, authored or reviewed drafts of the paper, approved the final draft.

### Data Availability

The following information was supplied regarding data availability:

The fMRI dataset used in this paper is freely available on SPM's website (<http://www.fil.ion.ucl.ac.uk/spm/data/auditory.html>) as explicitly mentioned in the 'Methods'. A code that describes how artificial data were generated and the core steps of the classic FCM algorithm is provided as [Supplemental File](#).

### Supplemental Information

Supplemental information for this article can be found online at <http://dx.doi.org/10.7717/peerj.5416#supplemental-information>.

## REFERENCES

- Ahmed MN, Yamany SM, Mohamed N, Farag AA, Moriarty T. 2002.** A modified fuzzy *c*-mean algorithm for bias field estimation and segmentation of MRI data. *IEEE Transactions on Medical Imaging* **21**:193–199 DOI [10.1109/42.996338](https://doi.org/10.1109/42.996338).
- Alexiuk MD, Pizzi NJ. 2004.** Cluster validation indices for fMRI data: fuzzy *c*-means with feature partitions versus cluster merging strategies. In: *Fuzzy information, 2004 processing NAFIPS '04, Alberta, Canada*. Piscataway: IEEE.
- Aljobouri HK, Jaber HA, Koçak OM, Algin O, Çankaya I. 2018.** Clustering fMRI data with a robust unsupervised learning algorithm for neuroscience data mining. *Journal of Neuroscience Methods* **299**:45–54 DOI [10.1016/j.jneumeth.2018.02.007](https://doi.org/10.1016/j.jneumeth.2018.02.007).
- Bandettini PA, Jesmanowicz A, Wong EC, Hyde JS. 1993.** Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine* **30**:161–173 DOI [10.1002/mrm.1910300204](https://doi.org/10.1002/mrm.1910300204).
- Bartels A, Zeki S. 2004.** Functional brain mapping during free viewing of natural scenes. *Human Brain Mapping* **21**:75–85 DOI [10.1002/hbm.10153](https://doi.org/10.1002/hbm.10153).

- Bartels A, Zeki S. 2005.** The chronoarchitecture of the cerebral cortex. *Philosophical Transactions of the Royal Society B: Biological Sciences* **360**:733–750 DOI [10.1098/rstb.2005.1627](https://doi.org/10.1098/rstb.2005.1627).
- Baumgartner R, Ryner L, Richter W, Summers R, Jarmasz M, Somorjai R. 2000.** Comparison of two exploratory data analysis methods for fMRI: fuzzy clustering vs. principal component analysis. *Magnetic Resonance Imaging* **18**:89–94 DOI [10.1016/S0730-725X\(99\)00102-2](https://doi.org/10.1016/S0730-725X(99)00102-2).
- Baumgartner R, Windischberger C, Moser E. 1998.** Quantification in functional magnetic resonance imaging: fuzzy clustering vs. correlation analysis. *Magnetic Resonance Imaging* **16**:115–125 DOI [10.1016/S0730-725X\(97\)00277-4](https://doi.org/10.1016/S0730-725X(97)00277-4).
- Bensaid AM, Hall LO, Bezdek JC, Clarke LP, Silbiger ML, Arrington JA, Murtaqh RF. 1996.** Validity-guided (re)clustering with applications to imagesegmentation. *IEEE Transactions on Fuzzy Systems* **4**:112–123 DOI [10.1109/91.493905](https://doi.org/10.1109/91.493905).
- Bezdek JC. 1981.** *Pattern recognition with fuzzy objective functions algorithms*. New York: Plenum Press.
- Bezdek JC, Hall LO, Clark MC, Goldgof DB, Clarke LP. 1997.** Medical image analysis with fuzzy models. *Statistical Methods in Medical Research* **6**:191–214 DOI [10.1177/096228029700600302](https://doi.org/10.1177/096228029700600302).
- Bezdek JC, Pal NR. 1998.** Some new indexes of cluster validity. *IEEE Transactions on Systems, Man, and Cybernetics B* **28**:301–315 DOI [10.1109/3477.678624](https://doi.org/10.1109/3477.678624).
- Bouguessa M, Wang S, Sun H. 2006.** An objective approach to cluster validation. *Pattern Recognition Letters* **27**:1419–1430 DOI [10.1016/j.patrec.2006.01.015](https://doi.org/10.1016/j.patrec.2006.01.015).
- Caballero-Gaudes C, Reynolds RC. 2017.** Methods for cleaning the BOLD fMRI signal. *NeuroImage* **154**:128–149 DOI [10.1016/j.neuroimage.2016](https://doi.org/10.1016/j.neuroimage.2016).
- Chuang KH, Chiu MJ, Lin CC, Chen JH. 1999.** Model-free functional MRI analysis using kohonen clustering neural network and fuzzy *c*-means. *IEEE Transactions on Medical Imaging* **18**:1117–1128 DOI [10.1109/42.819322](https://doi.org/10.1109/42.819322).
- Dave RN. 1996.** Validating fuzzy partition obtained through *c*-shells clustering. *Pattern Recognition Letters* **17**:613–623 DOI [10.1016/0167-8655\(96\)00026-8](https://doi.org/10.1016/0167-8655(96)00026-8).
- Dave RN, Krishnapuram R. 1997.** Robust clustering methods: a united view. *IEEE Transactions on Fuzzy Systems* **5**:270–293 DOI [10.1109/91.580801](https://doi.org/10.1109/91.580801).
- Derntl A, Plant C. 2016.** Clustering techniques for neuroimaging applications. *WIREs Data Mining and Knowledge Discovery* **6**:22–36 DOI [10.1002/widm.1174](https://doi.org/10.1002/widm.1174).
- Dik A, Jebari K, Bouroumi A, Ettouhami A. 2014.** A new fuzzy clustering by outliers. *Journal of Engineering and Applied Sciences* **9**:372–377 DOI [10.3923/jeasci.2014.372.377](https://doi.org/10.3923/jeasci.2014.372.377).
- Dimitriadou E, Barth M, Windischberger C, Hornik K, Moser E. 2004.** A quantitative comparison of functional MRI cluster analysis. *Artificial Intelligence in Medicine* **31**:57–71 DOI [10.1016/j.artmed.2004.01.010](https://doi.org/10.1016/j.artmed.2004.01.010).
- DonGiovanni D, Vaina LM. 2016.** Select and cluster: a method for finding functional networks of clustered voxels in fMRI. *Computational Intelligence and Neuroscience* **2016**:4705162 DOI [10.1155/2016/4705162](https://doi.org/10.1155/2016/4705162).



- Dubes RC. 1987.** How many clusters are best? An experiment. *Pattern Recognition* 20:645–663 DOI 10.1016/0031-3203(87)90034-3.
- Esposito F, Formisano E, Seifritz E, Goebel R, Morrone R, Tedeschi G, Di Salle F. 2002.** Spatial independent component analysis of functional MRI time-series: to what extent do results depend on the algorithm used? *Human Brain Mapping* 16:146–157 DOI 10.1002/hbm.10034.
- Fadili MJ, Ruan S, Bloyet D, Mazoyer B. 2000.** A multistep unsupervised fuzzy clustering analysis of fMRI time series. *Human Brain Mapping* 10:160–178 DOI 10.1002/1097-0193(200008)10:4<160::AID-HBM20>3.0.CO;2-U.
- Fadili MJ, Ruan S, Bloyet D, Mazoyer B. 2001.** On the number of clusters and the fuzziness index for unsupervised FCA application to BOLD fMRI time series. *Medical Image Analysis* 5:55–67 DOI 10.1016/S1361-8415(00)00035-9.
- Fatemizadeh E, Taalimi A, Davoudi H. 2009.** Extracting activated regions of fMRI data using unsupervised learning. In: *IEEE-INNS-ENNS international joint conference on neural networks, Atlanta, GA, USA*. Piscataway: IEEE, 641–645.
- Fukuyama Y, Sugeno M. 1989.** A new method for choosing the number of clusters for fuzzy *c*-means method. In: *5th international fuzzy systems symposium*. Ankara: TOBB University of Economics & Technology, 247–250.
- Gath I, Geva AB. 1989.** Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 11:773–781 DOI 10.1109/34.192473.
- Geva AB, Steinberg Y, Bruckmair S, Nahum G. 2000.** A comparison of cluster validity criteria for a mixture of normal distributed data. *Pattern Recognition Letters* 21:511–529 DOI 10.1016/S0167-8655(00)00016-7.
- Golay X, Kollias S, Stoll G, Meier D, Valavanis A, Boesiger P. 1998.** A new correlation-based fuzzy logic clustering algorithm for fMRI. *Magnetic Resonance in Medicine* 40:249–260 DOI 10.1002/mrm.1910400211.
- Goutte C, Toft P, Rostrup E, Nielsen FA, Hansen LK. 1999.** On clustering fMRI time series. *NeuroImage* 9:298–310 DOI 10.1006/nimg.1998.0391.
- Gu J, Cao Z, Zheng X, Aihua C. 2005.** Treatment of ill-balanced datasets of fMRI with modified fuzzy *c*-means method. In: *2005 IEEE engineering in medicine and biology 27th annual conference*. Piscataway: IEEE, 1411–1414 DOI 10.1109/IEMBS.2005.1616694.
- Gudbjartsson H, Patz S. 1995.** The Rician distribution of noisy MRI data. *Magnetic Resonance in Medicine* 34:910–914 DOI 10.1002/mrm.1910340618.
- Hammah RE, Curran JH. 2000.** Validity measures for the fuzzy cluster analysis of orientations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22:1467–1472 DOI 10.1109/34.895981.
- Hasson U, Nir Y, Levy I, Fuhrmann G, Malach R. 2004.** Intersubject synchronization of cortical activity during natural vision. *Science* 303:1634–1640 DOI 10.1126/science.1089506.
- He H, Tan Y, Fujimoto K. 2016.** Estimation of optimal cluster number for fuzzy clustering with combined fuzzy entropy index. In: *2016 IEEE international conference on fuzzy systems (FUZZ-IEEE), Vancouver, BC, Canada*. Piscataway: IEEE.

- Hu Y, Zuo C, Yang Y, Qu F. 2011.** A cluster validity index for fuzzy *c*-means clustering. In: *2011 international conference on system science, engineering design and manufacturing informatization (ICSEM), Guiyang, China*. Piscataway: IEEE.
- Jahanian H, Hossein-Zadeh GA, Soltanian-Zadeh H, Ardekani BA. 2004.** Controlling the false positive rate in fuzzy clustering using randomization: application to fMRI activation detection. *Magnetic Resonance Imaging* **22**:631–638 DOI [10.1016/j.mri.2004.01.035](https://doi.org/10.1016/j.mri.2004.01.035).
- Jahanian H, Soltanian-Zadeh H, Hossein-Zadeh GA. 2005.** Functional magnetic resonance imaging activation detection: fuzzy cluster analysis in wavelet and multiwavelet domains. *Journal of Magnetic Resonance Imaging* **22**:381–389 DOI [10.1002/jmri.20392](https://doi.org/10.1002/jmri.20392).
- Kao LJ, Huang YP. 2013.** A novel fuzzy clustering method with no outliers influence. *Applied Mechanics and Materials* **300–301**:735–739 DOI [10.4028/www.scientific.net/AMM.300-301.735](https://doi.org/10.4028/www.scientific.net/AMM.300-301.735).
- Keller A. 2000.** Fuzzy clustering with outliers. In: *19th international conference of the North American fuzzy information processing society (NAFIPS), Atlanta, GA, USA*. Piscataway: IEEE DOI [10.1109/NAFIPS.2000.877373](https://doi.org/10.1109/NAFIPS.2000.877373).
- Kim DW, Lee KH, Lee D. 2003.** Fuzzy cluster validation index based on inter-cluster proximity. *Pattern Recognition Letters* **24**:2561–2574 DOI [10.1016/S0167-8655\(03\)00101-6](https://doi.org/10.1016/S0167-8655(03)00101-6).
- Kim DJ, Park YW, Park DJ. 2001.** A novel validity index for determination of the optimal number of clusters. *IEICE Transactions on Information and Systems* **84-D**:281–285.
- Kim M, Ramakrishna RS. 2005.** New indices for cluster validity assessment. *Pattern Recognition Letters* **26**:2353–2363 DOI [10.1016/j.patrec.2005.04.007](https://doi.org/10.1016/j.patrec.2005.04.007).
- Krishnapuram R, Keller JM. 1993.** A possibilistic approach to clustering. *IEEE Transactions on Fuzzy Systems* **1**:98–110 DOI [10.1109/91.227387](https://doi.org/10.1109/91.227387).
- Kwon SH. 1998.** Cluster validity index for fuzzy clustering. *Electronics Letters* **34**:2176–2177 DOI [10.1049/el:19981523](https://doi.org/10.1049/el:19981523).
- Lange O, Meyer-Baese A, Hurdal M, Foo S. 2006.** A comparison between neural and fuzzy cluster analysis techniques for functional MRI. *Biomedical Signal Process Control* **1**:243–252 DOI [10.1016/j.bspc.2006.11.002](https://doi.org/10.1016/j.bspc.2006.11.002).
- Lange O, Meyer-Baese A, Wismueller A, Hurdal M, Sumners D, Auer D. 2004.** Model-free functional MRI analysis using improved fuzzy cluster analysis techniques. In: Priddy KL, ed. *SPIE intelligent computing: theory and applications*. Bellingham, WA: SPIE, 19–28.
- Lee MH, Hacker CD, Snyder AZ, Corbetta M, Zhang D, Leuthardt EC, Shimony JS. 2012.** Clustering of resting state networks. *PLOS ONE* **7**:e40370 DOI [10.1371/journal.pone.0040370](https://doi.org/10.1371/journal.pone.0040370).
- Levin DN, Uftring SJ. 2001.** Detecting brain activation in fMRI data without prior knowledge of mental event timing. *NeuroImage* **13**:153–160 DOI [10.1006/nimg.2000.0663](https://doi.org/10.1006/nimg.2000.0663).

- Liew AWC, Leung SH, Lau WH. 2000.** Fuzzy image clustering incorporating spatial continuity. *IEE Proceedings-Vision, Image and Signal Processing* **147**:185–192 DOI [10.1049/ip-vis:20000218](https://doi.org/10.1049/ip-vis:20000218).
- Lin PL, Huang PW, Wu CH, Huang SM. 2016.** An efficient validity index method for datasets with complex-shaped clusters. In: *2016 international conference on machine learning and cybernetics (ICMLC)*, Jeju, South Korea. Piscataway: IEEE DOI [10.1109/ICMLC.2016.7872948](https://doi.org/10.1109/ICMLC.2016.7872948).
- Lu Y, Jiang T, Zang Y. 2004.** A split-merge-based region-growing method for fMRI activation detection. *Human Brain Mapping* **22**:271–279 DOI [10.1002/hbm.20034](https://doi.org/10.1002/hbm.20034).
- Ma S, Correa NM, Li XL, Eichele T, Calhoun VD, Adalı T. 2011.** Automatic identification of functional clusters in fMRI data using spatial dependence. *IEEE Transactions on Biomedical Engineering* **58**:3406–3417 DOI [10.1109/TBME.2011.2167149](https://doi.org/10.1109/TBME.2011.2167149).
- Malinen S, Hlushchuk Y, Hari R. 2007.** Towards natural stimulation in fMRI-Issues of data analysis. *NeuroImage* **35**:131–139 DOI [10.1016/j.neuroimage.2006.11.015](https://doi.org/10.1016/j.neuroimage.2006.11.015).
- Maulik U, Bandyopadhyay S. 2002.** Performance evaluation of some clustering algorithms and validity indices. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**:1650–1654 DOI [10.1109/TPAMI.2002.1114856](https://doi.org/10.1109/TPAMI.2002.1114856).
- McKeown MJ, Hansen LK, Sejnowsk TJ. 2003.** Independent component analysis of functional MRI: what is signal and what is noise? *Current Opinion in Neurobiology* **13**:620–629 DOI [10.1016/j.conb.2003.09.012](https://doi.org/10.1016/j.conb.2003.09.012).
- McKeown MJ, Makeig S, Brown GG, Jung TP, Kindermann SS, Bell AJ, Sejnowski TJ. 1998.** Analysis of fMRI data by blind separation into independent spatial components. *Human Brain Mapping* **6**:160–188 DOI [10.1002/\(SICI\)1097-0193\(1998\)6:3<160::AID-HBM5>3.0.CO;2-1](https://doi.org/10.1002/(SICI)1097-0193(1998)6:3<160::AID-HBM5>3.0.CO;2-1).
- Meyer-Baese A, Wismueller A, Lange O. 2004.** Comparison of two exploratory data analysis methods for fMRI: unsupervised clustering versus independent component analysis. *IEEE Transactions on Information Technology in Biomedicine* **8**:387–398 DOI [10.1109/TITB.2004.834406](https://doi.org/10.1109/TITB.2004.834406).
- Moller U, Ligges M, Georgiewa P, Grunling C, Kaiser WA, Witte H, Blanz B. 2002.** How to avoid spurious cluster validation? A methodological investigation on simulated and fMRI data. *NeuroImage* **17**:431–446 DOI [10.1006/nimg.2002.1166](https://doi.org/10.1006/nimg.2002.1166).
- Orban P, Doyon J, Petrides M, Mennes M, Hoge R, Bellec P. 2015.** The richness of task-evoked hemodynamic responses defines a pseudohierarchy of functionally meaningful brain networks. *Cerebral Cortex* **25**:2658–2669 DOI [10.1093/cercor/bhu064](https://doi.org/10.1093/cercor/bhu064).
- Pakhira MK, Bandyopadhyay S, Maulik U. 2004.** Validity index for crisp and fuzzy clusters. *Pattern Recognition* **37**:487–501 DOI [10.1016/j.patcog.2003.06.005](https://doi.org/10.1016/j.patcog.2003.06.005).
- Pakhira MK, Bandyopadhyay S, Maulik U. 2005.** A study of some fuzzy cluster validity indices, genetic clustering and application to pixel classification. *Fuzzy Sets and Systems* **155**:191–214 DOI [10.1016/j.fss.2005.04.009](https://doi.org/10.1016/j.fss.2005.04.009).
- Pal NR, Bezdek JC. 1995.** On cluster validity for the fuzzy *c*-means model. *IEEE Transactions on Fuzzy Systems* **3**:370–379 DOI [10.1109/91.413225](https://doi.org/10.1109/91.413225).

- Parrish TB, Gitelman DR, LaBar KS, Mesulam MM. 2000.** Impact to signal-to-noise on functional MRI. *Magnetic Resonance in Medicine* **44**:925–932 DOI [10.1002/1522-2594\(200012\)44:6<925::AID-MRM14>3.0.CO;2-M](https://doi.org/10.1002/1522-2594(200012)44:6<925::AID-MRM14>3.0.CO;2-M).
- Pena JM, Lozano JA, Larranaga P. 1999.** An empirical comparison of four initialization methods for the K-means algorithm. *Pattern Recognition Letters* **20**:1027–1040 DOI [10.1016/S0167-8655\(99\)00069-0](https://doi.org/10.1016/S0167-8655(99)00069-0).
- Quiqley MA, Haughton VM, Carew J, Cordes D, Moritz CH, Meyerand ME. 2002.** Comparison of independent component analysis and conventional hypothesis-driven analysis for clinical functional MR image processing. *American Journal of Neuroradiology* **23**:49–58.
- Ren M, Liu P, Wang Z, Yi J. 2016.** A self-adaptive fuzzy *c*-means algorithm for determining the optimal number of clusters. *Computational Intelligence and Neuroscience* **2016**:2647389 DOI [10.1155/2016/2647389](https://doi.org/10.1155/2016/2647389).
- Rezaee B. 2010.** A cluster validity index for fuzzy clustering. *Fuzzy Sets and Systems* **161**:3014–3025 DOI [10.1016/j.fss.2010.07.005](https://doi.org/10.1016/j.fss.2010.07.005).
- Rezaee MR, Lelieveldt BPF, Reider JHC. 1998.** A new cluster validity index for the fuzzy *c*-mean. *Pattern Recognition Letters* **19**:237–246 DOI [10.1016/S0167-8655\(97\)00168-2](https://doi.org/10.1016/S0167-8655(97)00168-2).
- Rhee HS, Oh KW. 1996.** A performance measure for the fuzzy cluster validity. In: *Proceedings of the 1996 Asian fuzzy systems symposium on soft computing in intelligent systems and information processing. 11–14 December 1996, Kenting, Taiwan*. Piscataway: IEEE, 364–369 DOI [10.1109/AFSS.1996.583633](https://doi.org/10.1109/AFSS.1996.583633).
- Schwämmle V, Jensen ON. 2010.** A simple and fast method to determine the parameters for fuzzy *c*-means cluster analysis. *Bioinformatics* **56**:1841–2848 DOI [10.1093/bioinformatics/btq534](https://doi.org/10.1093/bioinformatics/btq534).
- Seghier ML, Friston KJ, Price CJ. 2007.** Detecting subject-specific activations using fuzzy clustering. *NeuroImage* **36**:594–605 DOI [10.1016/j.neuroimage.2007.03.021](https://doi.org/10.1016/j.neuroimage.2007.03.021).
- Seghier ML, Price CJ. 2009.** Dissociating functional brain networks by decoding the between-subject variability. *NeuroImage* **45**:349–359 DOI [10.1016/j.neuroimage.2008.12.017](https://doi.org/10.1016/j.neuroimage.2008.12.017).
- Selim AZ, Ismail MA. 1986.** On the local optimality of the fuzzy ISODATA clustering algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8**:284–288 DOI [10.1109/TPAMI.1986.4767783](https://doi.org/10.1109/TPAMI.1986.4767783).
- Sheng W, Swift S, Zhang L, Liu X. 2005.** A weighted sum validity function for clustering with a hybrid niching genetic algorithm. *IEEE Transactions on Systems, Man, and Cybernetics, Part B* **35**:1156–1167 DOI [10.1109/TSMCB.2005.850173](https://doi.org/10.1109/TSMCB.2005.850173).
- Smolders A, De Martino F, Staeren N, Scheunders P, Sijbers J, Goebel R, Formisano E. 2007.** Dissecting cognitive stages with time-resolved fMRI data: a comparison of fuzzy clustering and independent component analysis. *Magnetic Resonance Imaging* **25**:860–868 DOI [10.1016/j.mri.2007.02.018](https://doi.org/10.1016/j.mri.2007.02.018).
- Soille P. 2003.** *Morphological image analysis: principles and applications*. New York: Springer-Verlag.

- Soltanian-Zadeh H, Peck DJ, Hearshen DO, Lajiness-O'Neill RR. 2004.** Model-independent method for fMRI analysis. *IEEE Transactions on Medical Imaging* 23:285–296 DOI [10.1109/TMI.2003.823064](https://doi.org/10.1109/TMI.2003.823064).
- Stone JV. 2002.** Independent component analysis: an introduction. *Trends in Cognitive Sciences* 6:59–64 DOI [10.1016/S1364-6613\(00\)01813-1](https://doi.org/10.1016/S1364-6613(00)01813-1).
- Suleman A. 2017.** Measuring the congruence of fuzzy partitions in fuzzy  $c$ -means clustering. *Applied Soft Computing* 52:1285–1295 DOI [10.1016/j.asoc.2016.06.037](https://doi.org/10.1016/j.asoc.2016.06.037).
- Sun H, Wang S, Jiang Q. 2004.** FCM-based model selection algorithms for determining the number of clusters. *Pattern Recognition* 37:2027–2037 DOI [10.1016/j.patcog.2004.03.012](https://doi.org/10.1016/j.patcog.2004.03.012).
- Tang X, Zeng W, Wang N, Yang J. 2015.** An adaptive RV measure based fuzzy weighting subspace clustering (ARV-FWSC) for fMRI data analysis. *Biomedical Signal Process Control* 22:146–154 DOI [10.1016/j.bspc.2015.07.006](https://doi.org/10.1016/j.bspc.2015.07.006).
- Thirion B, Varoquaux G, Dohmatob E, Poline JB. 2014.** Which fMRI clustering gives good brain parcellations? *Frontiers in Neuroscience* 8:167 DOI [10.3389/fnins.2014.00167](https://doi.org/10.3389/fnins.2014.00167).
- Tibshirani R, Walther G, Hastie T. 2001.** Estimating the number of clusters in a data set via the gap statistic. *Journal of the Royal Statistical Society. Series B, Statistical Methodology* 63:411–423 DOI [10.1111/1467-9868.00293](https://doi.org/10.1111/1467-9868.00293).
- Tsekouras GE, Sarimveis H. 2004.** A new approach for measuring the validity of the fuzzy  $c$ -means algorithm. *Advances in Engineering Software* 35:567–575 DOI [10.1016/j.advengsoft.2004.05.001](https://doi.org/10.1016/j.advengsoft.2004.05.001).
- Wang W, Zhang Y. 2007.** On fuzzy cluster validity indices. *Fuzzy Sets and Systems* 158:2095–2117 DOI [10.1016/j.fss.2007.03.004](https://doi.org/10.1016/j.fss.2007.03.004).
- Windham MP. 1981.** Cluster validity for fuzzy clustering algorithms. *Journal of Fuzzy Sets and Systems* 5:177–185 DOI [10.1016/0165-0114\(81\)90015-4](https://doi.org/10.1016/0165-0114(81)90015-4).
- Windischberger C, Barth M, Lamm C, Schroeder L, Bauer H, Gur RC, Moser E. 2003.** Fuzzy cluster analysis of high-field functional MRI data. *Artificial Intelligence in Medicine* 29:203–223 DOI [10.1016/S0933-3657\(02\)00072-6](https://doi.org/10.1016/S0933-3657(02)00072-6).
- Wismuller A, Meyer-Base A, Lange O, Auer D, Reiser MF, Sumners D. 2004.** Model-free functional MRI analysis based on unsupervised clustering. *Journal of Biomedical Informatics* 37:10–18 DOI [10.1016/j.jbi.2003.12.002](https://doi.org/10.1016/j.jbi.2003.12.002).
- Wu KL, Yang MS. 2005.** A cluster validity index for fuzzy clustering. *Pattern Recognition Letters* 26:1275–1291 DOI [10.1016/j.patrec.2004.11.022](https://doi.org/10.1016/j.patrec.2004.11.022).
- Xie XL, Beni G. 1991.** A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13:841–847 DOI [10.1109/34.85677](https://doi.org/10.1109/34.85677).
- Yang SL, Li K, Liang Z, Li W, Xue Y. 2018.** A novel cluster validity index for fuzzy  $c$ -means algorithm. *Soft Computing* 22:1921–1931 DOI [10.1007/s00500-016-2453-y](https://doi.org/10.1007/s00500-016-2453-y).
- Yu J, Cheng Q, Huang H. 2004.** Analysis of the weighting exponent in the FCM. *IEEE Transactions on Systems, Man, and Cybernetics B* 34:634–639 DOI [10.1109/TSMCB.2003.810951](https://doi.org/10.1109/TSMCB.2003.810951).
- Yu J, Li CX. 2006.** Novel cluster validity index for FCM algorithm. *Journal of Computer Science and Technology* 21:137–140 DOI [10.1007/s11390-006-0137-2](https://doi.org/10.1007/s11390-006-0137-2).

- Zacks JM, Braver TS, Sheridan MA, Donaldson DI, Snyder AZ, Ollinger JM, Buckner RL, Raichle ME. 2001.** Human brain activity time-locked to perceptual event boundaries. *Nature Neuroscience* **4**:651–655 DOI [10.1038/88486](https://doi.org/10.1038/88486).
- Zahid N, Aboulala O, Limouri M, Essaid A. 1999.** Unsupervised fuzzy clustering. *Pattern Recognition Letters* **20**:123–129 DOI [10.1016/S0167-8655\(98\)00142-1](https://doi.org/10.1016/S0167-8655(98)00142-1).
- Zahid N, Limouri M, Essaid A. 1999.** A new cluster-validity for fuzzy clustering. *Pattern Recognition* **32**:1089–1097 DOI [10.1016/S0031-3203\(98\)00157-5](https://doi.org/10.1016/S0031-3203(98)00157-5).
- Zhang D, Ji M, Yang J, Zhang Y, Xie F. 2014.** A novel cluster validity index for fuzzy clustering based on bipartite modularity. *Fuzzy Sets and Systems* **253**:122–137 DOI [10.1016/j.fss.2013.12.013](https://doi.org/10.1016/j.fss.2013.12.013).
- Zhou K, Ding S, Fu C, Yang SL. 2014.** Comparison and weighted summation type of fuzzy cluster validity indices. *International Journal of Computers Communications* **9**:370–378 DOI [10.15837/ijccc.2014.3.237](https://doi.org/10.15837/ijccc.2014.3.237).
- Zhou K, Fu C, Yang SL. 2014.** Fuzziness parameter selection in fuzzy *c*-means: the perspective of cluster validation. *Science China* **57**:1–8 DOI [10.1007/s11432-014-5146-0](https://doi.org/10.1007/s11432-014-5146-0).