ELSEVIER

Pattern Recognition Letters 19 (1998) 237–246

Pattern Recognition Letters

A new cluster validity index for the fuzzy *c*-mean

M. Ramze Rezaee, B.P.F. Lelieveldt, J.H.C. Reiber *

Division of Image Processing, Department of Radiology, Leiden University Medical Center, P.O. Box 9600, 2300 RC Leiden, The Netherlands

Received 17 June 1997; revised 16 October 1997

Abstract

In this paper a new cluster validity index is introduced, which assesses the average compactness and separation of fuzzy partitions generated by the fuzzy *c*-means algorithm. To compare the performance of this new index with a number of known validation indices, the fuzzy partitioning of two data sets was carried out. Our validation performed favorably in all studies, even in those where other validity indices failed to indicate the true number of clusters within each data set.

Keywords: Cluster validity functional; Fuzzy clustering; Fuzzy *c*-mean

1. Introduction

The objective of most clustering methods is to provide useful information by grouping (unlabelled) data in clusters; within each cluster the data exhibits similarity. Similarity is defined by a distance measure, and global objective functional or regional graph-theoretic criteria are optimized to find the optimal partitions of data. The partitions generated by a clustering approach define for all data elements to which class (cluster) they belong. The partitions may define a hard boundary between sub-partitions; this is called hard clustering. In contrast, the boundaries between sub-partitions generated by a fuzzy clustering algorithm are vague. This means that each pattern of object data of a fuzzy partition belongs to different classes with different membership values.

Since clustering algorithms are unsupervised, irrespective of the clustering method (hard or fuzzy), the final partitions of data require some kind of validation in most applications. Backer and Jain (1981) have studied the performance of several clustering techniques based on a fuzzy set decomposition of the data set under consideration. In general, cluster validation may answer, among other aspects, questions such as: How good are the partitions? Is there a better partitioning possible?, etc. In addition, if the number of classes within the data is not known a priori, a validation index may help to find out the optimal number of classes. This is performed in three stages. Firstly, all parameters of the clustering method, except for the number of clusters, are fixed. Secondly, by varying the number of clusters between 2 and an upper value c_{max} and applying a clustering algorithm, for each number of clusters $c_i \in \{2,3,\dots,c_{max}\}$ a different partition of the data is found. In the final stage a validation index is applied

* Corresponding author. E-mail: hreiber@lkeb.azl.nl.

0167-8655/98/$19.00
PII S0167-8655(97)00168-2

to each partition, obtained at the second stage, to define a validation value. The true number of data classes can be determined based on extrema of these validation values for all c_i.

In this paper we will introduce a new validation index, which measures the separation between clusters and the cohesion within clusters, which are obtained by the fuzzy c-mean (FCM). After a brief review of the FCM in Section 2, a number of validation indices is described in Section 3. In addition, in this section the new validation index is introduced. To evaluate the performance of this validation index, one synthetic and one real world data set are used and the optimal number of object classes is assessed by applying a number of validation indices, including the new index. The results of the evaluation study are described in Section 5. Finally, this paper concludes with a discussion section (Section 4).

2. Fuzzy c-mean algorithm

Fuzzy c-mean (FCM) is an unsupervised clustering algorithm that has been applied successfully to a number of problems involving feature analysis, clustering and classifier design. FCM has a wide domain of applications such as agricultural engineering, astronomy, chemistry, geology, image analysis, medical diagnosis, shape analysis and target recognition (Bezdek, 1987). Unlabeled data are classified by minimizing an objective function based on a norm and clusters prototype. Although the description of the original algorithm dates back to 1973 (Bezdek, 1973; Dunn, 1974) derivatives have been described with modified definitions for the norm and prototypes for the cluster centers (Dave and Bhaswan, 1992; Krishnapuram et al, 1992; Man and Gath, 1994).

The FCM minimizes an objective function J_m, which is the weighted sum of squared errors within groups and is defined as follows:

$$J_m(\boldsymbol{U},\boldsymbol{V};X) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^m \| x_k - v_i \|_A^2, \; 1 < m < \infty, \tag{1}$$

where $\boldsymbol{V} = (v_1, v_2, \dots, v_c)$ is a vector of unknown cluster prototype (centers) $v_i \in R^p$. The value of u_{ik} represents the grade of membership of data point x_k of set $X = \{x_1, x_2, \dots, x_n\}$ to the ith cluster. The inner product defined by a norm matrix A defines a measure of similarity between a data point and the cluster prototypes. A nondegenerate fuzzy c-partition of X is conveniently represented by a matrix $\boldsymbol{U} = [u_{ik}]$.

It has been shown by Bezdek (1981) that if $\| x_k - v_i \|_A > 0$ for all i and k, then $(\boldsymbol{U},\boldsymbol{V})$ may minimize J_m only, when $m > 1$ and

$$v_i = \frac{\sum_{k=1}^{n} (u_{ik})^m x_k}{\sum_{k=1}^{n} (u_{ik})^m} \quad \text{for } 1 \leqslant i \leqslant c, \tag{2}$$

$$u_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{\| x_k - v_i \|_A^2}{\| x_k - v_j \|_A^2} \right)^{1/m-1}}$$

$$\text{for } 1 \leqslant i \leqslant c, \; 1 \leqslant k \leqslant n. \tag{3}$$

Among others, J_m can be minimized by the Picard iteration approach. This method minimizes J_m by initializing the matrix $\boldsymbol{U}$ randomly (or predefined) and computing the cluster prototypes (Eq. (2)) and the membership values (Eq. (3)) after each iteration. The iteration is terminated when it reaches a stable condition. This can be defined for example, when the changes in the cluster centers or the membership values at two successive iteration steps is smaller than a predefined threshold value. The FCM algorithm always converges to a local minimum or a saddle point. A different initial guess of u_{ij} may lead to a different local minimum. Finally, to assign each data point to a specific cluster, defuzzification is necessary, e.g., by attaching a data point to a cluster for which the value of the membership is maximal.

3. Validation indices for the fuzzy c-mean

3.1. Validation criteria

Although the FCM can find a partition of data for a fixed number of clusters (objects), one objective of a cluster validity procedure is to determine automatically the optimal number of clusters. This is desired, for example, when the FCM is used for image segmentation purposes and the number of objects in the image is not known a priori. Validation of a generated fuzzy partition by the FCM can be achieved by a cluster validity index. If $c_{\min}$ is defined as the minimum and $c_{\max}$ as the maximum number of clusters, respectively, then for each $c \in [c_{\min}, c_{\max}]$ a partition can be generated by the FCM. The value of a cluster validity index can then be calculated for the partitions of data for each c. By comparing all values of an index for all possible number of clusters, one can determine the optimal number of clusters. This can be achieved, for example, by selecting the number of clusters for which an index is minimized.

There are a number of clusters validation indices available (Bezdek, 1974, 1975; Gunderson, 1978; Windham, 1981, 1982; Libert and Roubens, 1983; Windham et al., 1989; Fukuyama and Sugeno, 1989; Xie and Beni, 1991; Gindy et al., 1995). Some validity methods use only the membership values of a fuzzy partition of data. Among other functional, such indices are the partition coefficient V_{PC} (Bezdek, 1974), the partition entropy V_{PE} (Bezdek, 1975), the proportion exponent (Windham, 1981) and the uniform data functional (Windham, 1982).

Table 1 lists a number of cluster validation indices, which are evaluated in our study. The functional partition coefficient V_{PC} and the partition entropy V_{PE} use only the membership values u_{ik} of a fuzzy partition of data set X. Some empirical studies have shown that maximizing V_{PC} (or minimizing V_{PE}) often leads to a good interpretation of the data (Pal and Bezdek, 1995); the best partition is achieved when the value for V_{PC} has a maximum or the value for V_{PE} has a minimum, for a certain number of clusters. However, a strong criticism against the indices V_{PC} or V_{PE}, is that they are only implicitly a function of the data set X. It means that they do not use the data itself. To overcome this lack of direct connection to the geometrical properties of the data set, Gunderson's separation coefficient (Gunderson, 1978) uses both the data set and the prototypes (cluster centers). This is also the case for the Fukuyama's and Sugeno's index (Fukuyama and Sugeno, 1989) and the Xie–Beni's index (Xie and Beni, 1991). A limited analysis of these indices was performed by Pal and Bezdek (1995, 1997). They also analyzed the reliability of these indices as a function of the weighting exponent m of Eq. (1).

3.2. Proposed validation index

A reliable validation functional for the FCM must consider both the *compactness* and the *separation* of a fuzzy c-partition. The optimal partitions require a maximal compactness for each cluster (partition) in such a way that the clusters are located far from each other. If only the compactness requirement is considered by a validation functional, the best partition is obtained when each data point is considered as a separate cluster; nothing is compacter than a cluster that includes only one data point. On the other hand, if only the optimal separation between the clusters is considered as the single validation criterion, the best partition will be the data set itself; the distance between a cluster (the total data set) and itself is

Table 1
Four validation functionals for the fuzzy c-mean

Validity index	Functional description	Optimal cluster number
Partition coefficient	$V_{PC}(\boldsymbol{U}) = \frac{1}{n}(\sum_{k=1}^{n}\sum_{i=1}^{c}u_{ik}^2)$	$\mathrm{Max}\{V_{PC}(\boldsymbol{U},c_i,m)\}$
Partition entropy	$V_{PE}(\boldsymbol{U}) = -\frac{1}{n}(\sum_{k=1}^{n}\sum_{i=1}^{c}u_{ik}\log_a(u_{ik}))$	$\mathrm{Min}\{V_{PE}(\boldsymbol{U},c_i,m)\}$
Fukuyama and Sugeno (1989)	$V_{FS,m}(\boldsymbol{U},\boldsymbol{V};X) = \sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^m(\Vert x_k - v_i\Vert^2 - \Vert v_i - \bar{v}\Vert_A^2)$	$\mathrm{Min}\{V_{FS}(\boldsymbol{U},c_i,m)\}$
Xie and Beni (1991)	$V_{XB}(\boldsymbol{U};\boldsymbol{V};X) = \dfrac{\sum_{i=1}^{c}\sum_{k=1}^{n}u_{ik}^m\Vert x_k - v_i\Vert^2}{n\left(\min\{v_i - v_j\}\right)}$	$\mathrm{Min}\{V_{XB}(\boldsymbol{U},c_i,m)\}$

x_k is the kth data point, v_i are cluster prototypes (cluster centers), c_i is the number of clusters, $\bar{v}$ is the grand mean of all data x_k and u_{ik} is the membership value of data x_k of class c_i.

zero. Therefore a reliable validation functional will have an optimal value for that partition, which combines both criteria. In an attempt to find such a functional, we have designed a new validation index which includes the *compactness* and *separation* criteria. We called this index V_{CWB} (Compose Within and Between scattering). To define this functional, let us first describe the following requirements and definitions.

Requirements. A fuzzy c-partition of the data set $\mathrm{X}=\{x_1,x_2,\dots,x_n|x_i\in R^p\}$ with c cluster centers v_i, such that $\boldsymbol{V}=\{v_1,v_2,\dots,v_c\}$ and $\boldsymbol{U}=[u_{ik}\ (i=1,2,\dots,c;\ k=1,2,\dots,n)]$.

Definition 1. The variance of the pattern set X is called $\sigma(X)\in R^p$ with the value of the pth dimension defined as

$$\sigma_x{}^p=\frac{1}{n}\sum_{k-1}^{n}\left(x_k^p-\bar{x}^p\right)^2, \tag{4}$$

where $\bar{x}^p$ is the pth value of the grand mean of $\overline{X}=\sum_{k=1}^{n}x_k/n,\ \forall x_k\in X$.

Definition 2. The fuzzy variation of the cluster i is called $\sigma(v_i)\in R^p$ with the pth value defined as

$$\sigma_{v_i}{}^p=\frac{1}{n}\sum_{k=1}^{n}u_{ik}\left(x_k^p-v_i^p\right)^2. \tag{5}$$

Definition 3. The average scattering for c clusters is defined as

$$\mathrm{Scat}(c)=\frac{\frac{1}{c}\sum_{i=1}^{c}\|\sigma(v_i)\|}{\|\sigma(X)\|}, \tag{6}$$

where $\|x\|=(x^{\mathrm{T}}\cdot x)^{1/2}$.

Definition 4. A distance functional $\mathrm{Dis}(c)$ is defined as

$$\mathrm{Dis}(c)=\frac{D_{\max}}{D_{\min}}\sum_{k=1}^{c}\left(\sum_{z=1}^{c}\|v_k-v_z\|\right)^{-1}, \tag{7}$$

where $D_{\max}=\mathrm{maximum}\{\|v_i-v_j\|\}\quad\forall i,j\in\{2,3,\dots,c\}$ is the maximum distance between the cluster prototypes. The $D_{\min}$ has the same definition as $D_{\max}$, but for the minimum distance between the cluster prototypes.

Our validation index V_{CWB} is now defined by combining the last two equations:

$$V_{\mathrm{CWB}}(\boldsymbol{U},\boldsymbol{V})=\alpha\ \mathrm{Scat}(c)+\mathrm{Dis}(c), \tag{8}$$

where α is a weighting factor equal to $\mathrm{Dis}(c_{\max})$.

The first term of V_{CWB}, i.e. $\mathrm{Scat}(c)$ of Eq. (6), indicates the average of the scattering (variation) *within* the clusters for c number of clusters. A small value for this term indicates a compact partition. As the scattering within the clusters increases, they become less compact, and therefore $\mathrm{Scat}(c)$ is a good indication of the average compactness of clusters. In general, the scattering index $\mathrm{Scat}(c)$ does not take any geometric assumption about the prototype into account. The second term of our validation index, $\mathrm{Dis}(c)$, indicates the total scattering (separation) *between* the clusters. Generally, this term will increase with the number of clusters and is influenced by the geometry of the cluster centers. Since the values of the two terms of V_{CWB} are of a different range, a weighting factor α is needed in order to counterbalance both terms in a proper way.

A cluster number, which minimizes the validation index V_{CWB} can be considered as an optimal value for the number of object classes present in the data.

3.3. FCM validation algorithm

If the minimum and maximum values of the number of clusters be denoted as $c_{\min}$ and $c_{\max}$, respectively, then by the following pseudo algorithm an optimal number of clusters c_{opt} of a data set can be obtained.

1. Initialization: $c_{\mathrm{opt}}=c\leftarrow c_{\max}$;
2. Apply FCM to the data set to update the cluster centers v_i and the membership values μ_{ik}
3. Do iteration and test for convergence; if not goto 2.
4. if $(c=c_{\max})$ $\{\alpha=\mathrm{Dis}(c_{\max})$; indexValue $=V_{\mathrm{cwb}}(c)\}$
   else if $(V_{\mathrm{CWB}}(c)<\mathrm{indexValue})$ $\{c_{\mathrm{opt}}\leftarrow c$; indexValue $=V_{\mathrm{CWB}}(c);\}$
5. $c\leftarrow c-1$, if $(c=c_{\min}-1)$ stop else goto 2.

In step 3 the convergence of FCM can be tested, e.g., by the condition $|u_{ik}(t+1)-u_{ik}(t)|<\epsilon\quad\forall i\in$

$[c_{\min}, c_{\max}] \wedge \forall k \in \{1,2,\ldots,n\}$, where t and $t+1$ represent two successive iteration steps. After step 3 the fuzzy partition is generated. This partition is validated by the V_{CWB} in step 4. The parameter ''indexValue'' stores the minimum value of V_{CWB} so far. After step 5 the optimal number of clusters c_{opt} will be found that corresponds to the minimum value of the index V_{CWB} within the range of $[c_{\min}, c_{\max}]$.

Also, by varying the fuzzy parameter $m \in (1,\infty)$ of the objective function of Eq. (1) and by applying the above algorithm, the optimal number of clusters is found that is the function of $V_{\mathrm{CWB}}(m,c)$. This optimum minimizes the cluster validity V_{CWB}.

4. Data acquisition and FCM parameters

To compare the performance of the proposed validation index with a number of known validation indices, an evaluation study was carried out. This study included two data sets.

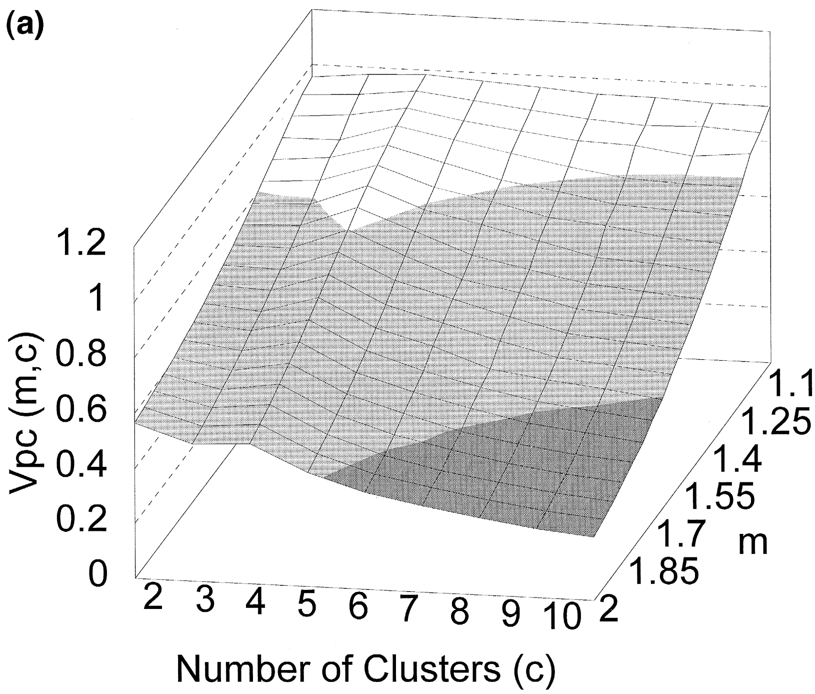
Normal-4. The first data set called Normal-4 was described by Pal and Bezdek (1995) as follows: Normal-4 includes 800 data points consisting of clusters of 200 points, each of the four components of a mixture of $c=4$, $p=4$-variate normal. The population mean vector and covariance matrix for each component of the normal mixture were $\mu_i = 3e_i$ and $\Sigma_i = I_4$, $i=1,2,3,4$, and $e_i = (0,\ldots,1,\ldots,0)$ (the 1 is on the ith position).

IRIS. The second data set of the evaluation study was the Iris data set of Anderson–Fisher (Anderson, 1935; Fisher, 1936). This is a biometric data set consisting of 150 measurements belonging to three flower varieties: Setosa, Versicolor and Virginica, generally known as the Iris data set. Each class includes 50 observations, in which two variables, length and width of the petal and sepal, are measured. Since the length and the width of each variable are measured, each individual measurement (one flower out of 150 samples) is represented as a point in $p=4$-dimensional measurement space. Pal and Bezdek (1997) have indicated that since two of the three classes have a substantial overlap, one can argue in favor of both $c=2$ and $c=3$. Halgamuge and Glesner (1994) have shown that a very good classification can be obtained by using only two features. As demonstrated in the next section, a very good classification rate can be obtained when only the feature ''petal length'' is used.

Parameters of the study. In our study the parameters of the FCM were set as follows: the stop criterion for the iteration was $\epsilon = 0.001$. The norm $\|{}^*\|_A$ was the Euclidean and all iterations began with a randomly initialized u_{ik}. Two parameters of the FCM, the number of clusters c and the weighting exponent m, were varied for both data sets. The weighting exponent m was varied from 1.1 to 2.0 with a step value of 0.05 (total of 19 samples). To obtain the optimal number of clusters for each fixed value of m, nine different partitions were made by varying the number of clusters from two to ten and applying the FCM. All these nine partitions were then used to calculate the value of five validation indices: the partition coefficient V_{PC}, the partition entropy V_{PE}, Xie and Beni's index V_{XB}, Fukuyama's and Sugeno's index V_{FS} and the V_{CWB}. By comparing the values of the validation index of a specific functional $V(m=\text{fixed},\ 2 \leqslant c \leqslant 9)$, the optimal number of clusters was defined by that functional. In total 171 partitions (19 samples for $m \times 9$ possibilities for the number of clusters) were generated for each data set by applying the FCM. The optimal number of clusters for each specific m value was then registered.

5. Results

Normal-4. For 4 number of clusters the average classification error of FCM was 3.5% for all 19 samples of m. Fig. 1(a–e) shows the validation index values of all indices as a function of the weighting exponent m and the number of clusters c. As Fig. 1(a) demonstrates, by increasing the value of the weighting exponent m the partition coefficient $V_{\mathrm{PC}}(m,c)$ decreases. Up to the value of $m=1.75$, for each fixed m the maximum of $V_{\mathrm{PC}}(m,c)$ indicates 4 as the true numbers of clusters correctly. For $m \geqslant 1.8$ the maximum value of $V_{\mathrm{PC}}(m,c)$ appears at $V_{\mathrm{PC}}(m,c=2)$, indicating 2 as the optimal number of clusters. Fig. 1(b) shows the validation index entropy



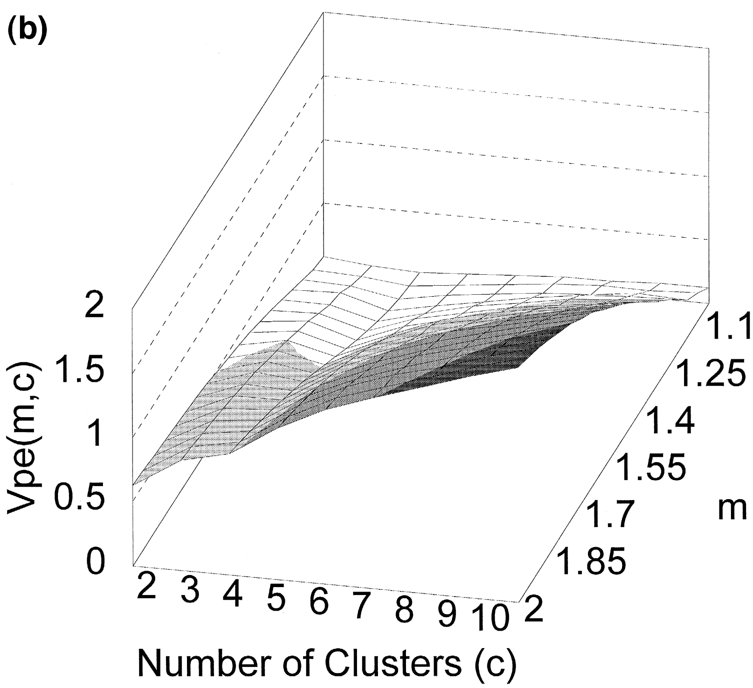


Fig. 1. The validation index of 4-Normal data set as a function of the number of clusters c and the weighting exponent m for: (a) partition coefficient V_{PC}; (b) partition entropy V_{PE}; (c) Fukuyama and Sugeno's index V_{FS}; (d) Xie and Beni's index V_{XB}; (e) the index V_{CWB}. The grey scales in the figures indicate the different validating index values.

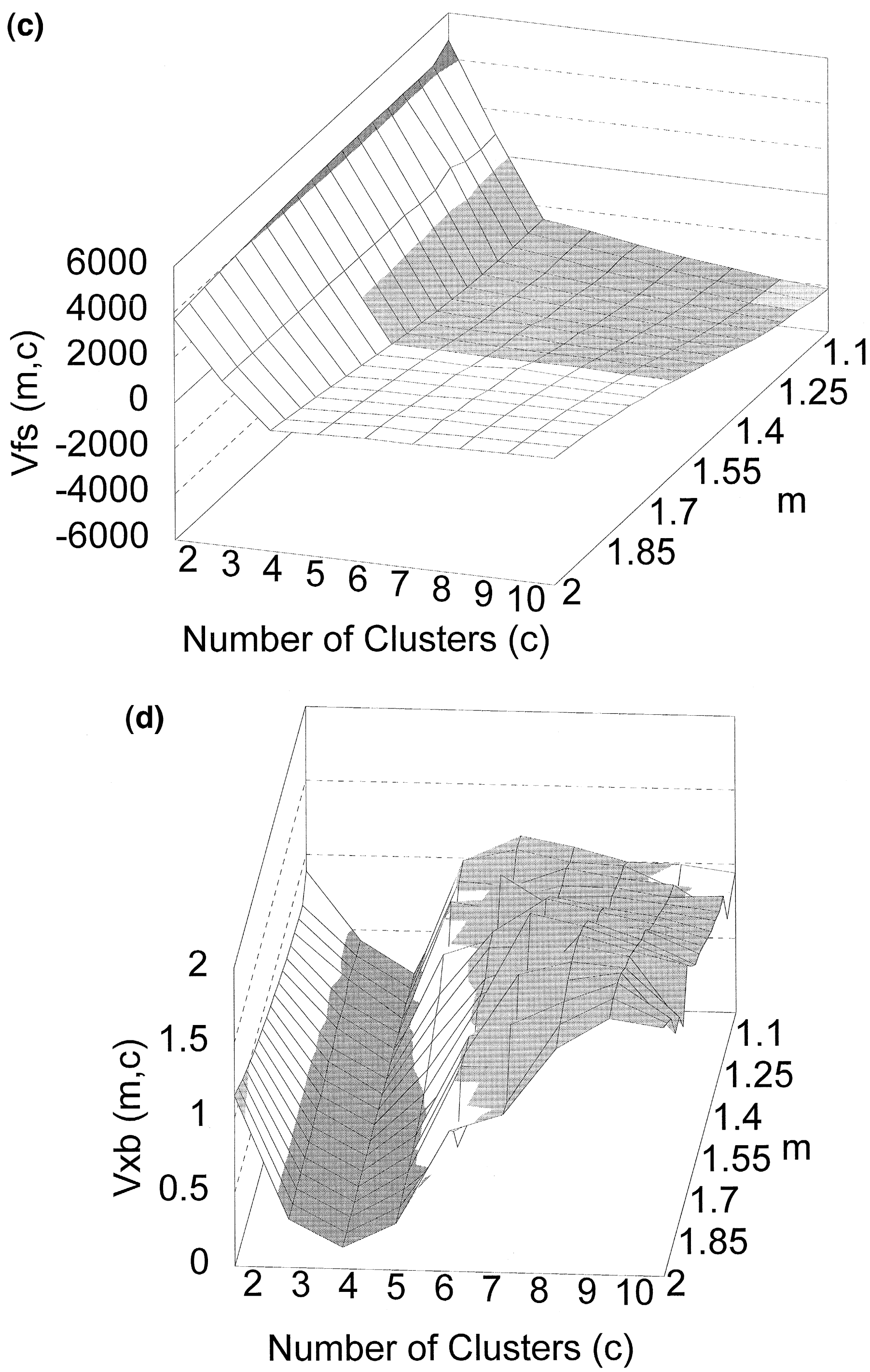


Fig. 1 (continued).

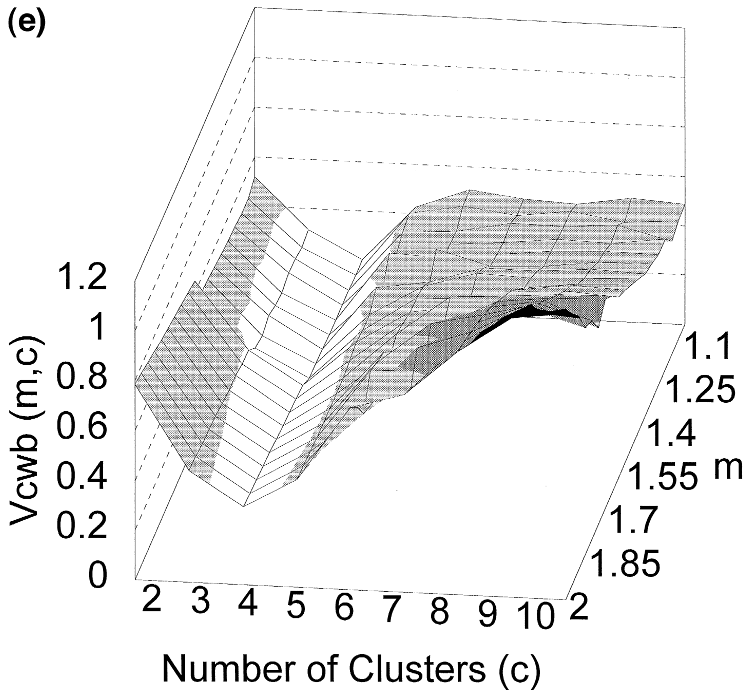


Fig. 1 (continued).

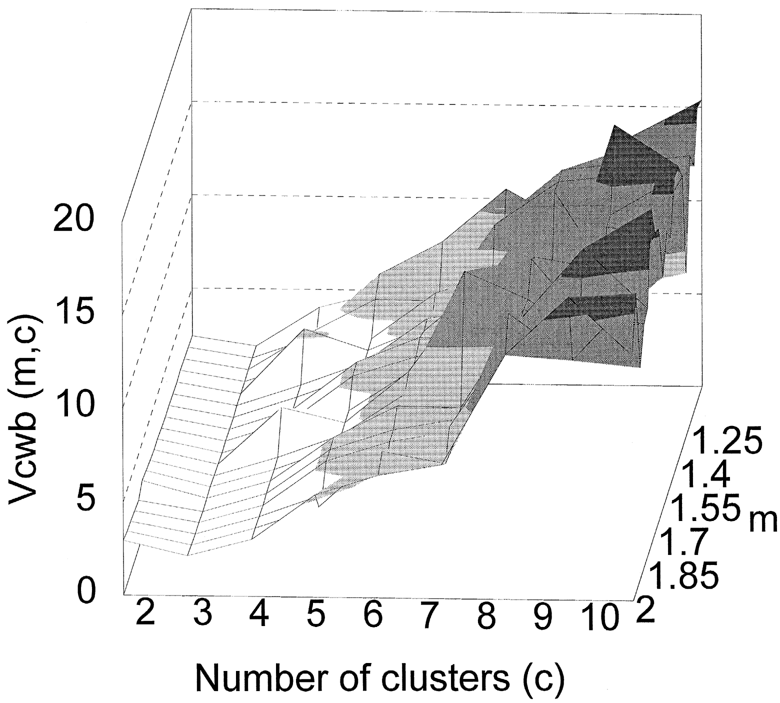


Fig. 2. The validation index V_{CWB} for the IRIS data set.

$V_{PE}(c,m)$. In contrast to V_{PC} by increasing the value of the weighting exponent m this index increases. Up to the value of $m = 1.55$ the minimum of this index for a specific sample of m and the variable numbers of clusters $(c \in \{2,3,\dots,10\})$ appears at $V_{PE}(m,c=4)$, indicating 4 as the optimal number of clusters. For all samples of m with a value $m \geqslant 1.55$, the $V_{PE}(m,c=2)$ is the minimum of this index denoting 2 as the optimal number of clusters incorrectly. The minimal values of the validation index of the Fukuyama–Sugeno index $V_{FS}(m,c)$ of Fig. 1(c) indicates 4 as the optimal number of clusters for $m \geqslant 1.7$. The minimum of this index appears, however, at $V_{FS}(m,10)$ for $1.1 \leqslant m \leqslant 1.65$, again denoting 10 as the true number of clusters incorrectly. Fig. 1(d) shows the validation index $V_{XB}(m,c)$. The minimum values of this index appear at $V_{XB}(m,4)$ for all 19 samples of m, which indicates the optimal number of clusters to be equal to 4. Also the minimum of our validation index $V_{CWB}(m,c)$, presented in Fig. 1(e), appears at $V_{CWB}(m,4)$ for all samples of the weighting exponent m.

IRIS data set. For $c = 3$ all feature combinations $(2^4 - 1 = 15)$ were used to assess which one has the smallest classification error. This was the case for the petal width. When this feature was used for all iris samples, the average classification error of the FCM was 6% (9 wrongly classified samples).

From Fig. 2 it is clear that for all samples of the weighting exponent m, the minimum appears at $V_{CWB}(m,3)$ indicating 3 to be the true number of clusters.

Table 2 presents the results for the IRIS data set. To obtain this table for each sample of weighting exponent m the optimal value of validation index $V(m,c)$ of all five indices was found. Since the weighting exponent m had 19 samples, a reliable validation index must find 19 times the true number of clusters. This is indeed the case for our validation index V_{CWB} on the last row of Table 2: for all samples of m, three is indicated as the optimal number of clusters. This is the minimum value of $V_{CWB}(m,c)$ for each sample of m. The V_{XB} indicated 3 to be the correct number of clusters for 9 samples of $1.30 \leqslant m$ and $m \geqslant 1.85$. This can be found in the second column of the third row of the same table. For all other ten samples this index indicated 2 to be the correct numbers of clusters (the first column of the third row). Fukuyama's and Sugeno's index V_{FS} indicates the higher number of clusters as the optimal value for the number of clusters.

Table 2
Results of five validation indices by using the petal width of IRIS data

	Number of clusters								
	2	3	4	5	6	7	8	9	10
Partition coefficient	6	7	0	0	0	0	2	2	2
Partition entropy	9	5	0	0	0	0	2	1	2
Xie and Beni (1991)	10	9	0	0	0	0	0	0	0
Fukuyama and Sugeno (1989)	0	0	0	1	2	0	2	4	10
V_{CWB}	0	19	0	0	0	0	0	0	0

6. Discussion

In this study we have used two data sets to evaluate the performance of a number of known validation functional. In the context of this study, a validation index is called reliable when the optimal number of clusters indicated by a validation index was equal to the true number of clusters of a data set. To find this number, nine fuzzy partitions for each data set were obtained, by fixing the weighting exponent m, varying the number of clusters from 2 to 10 and applying the FCM. These partitions were then used to obtain the validation values $V(m,c)$ for each of the five validation indices. The results of the first data set of 4-Normal classes show the effect of the weighting exponent m on each validation index (Fig. 1(a–e)). Some validation indices such as $V_{PC}(m,c)$ or $V_{FS}(m,c)$ decrease when the weighting exponent m increases. In contrast, other indices increase with an increasing value of m. This directly effects the local extrema (minimum or maximum) of three validation indices V_{PC}, V_{PE} and V_{FS} for a fixed m value: in case of V_{PC} for sample values of $m \geqslant 1.8$ the maximum of $V_{PC}(m,c)$ shifted from the point $V_{PC}(m,c=4)$ to $V_{PC}(m,c=2)$. Therefore, for $m \geqslant 1.8$ the maximum of V_{PC} did not indicate the true numbers of clusters. This was also the case for $V_{PE}(m,c)$, for each sample of $m \geqslant 1.55$ the minimum was shifted from $V_{PE}(m,c=4)$ to $V_{PE}(m,c=2)$. The minimum of $V_{FS}(m,c)$ shifted from $V_{FS}(m,c=4)$ to $V_{FS}(m,c=10)$. On the other hand, the minimum of the V_{XB} and our validation V_{CWB} remained unchanged at $V_{FS}(m,c=4)$ for all 19 samples of m.

The results of the IRIS data set also demonstrated that V_{CWB} was the most reliable index. According to this index, for all samples of m within the range [1.1,2], three classes exist within the IRIS data.

The results of this study suggest that the new validation index can achieve the optimal result for any possible data set. This is, however, not claimed here. Every validation index may fail, for example, when the numerical representation chosen to describe the different object features do not properly discriminate between different classes. Also the Euclidean norm (used in V_{CWB}) may be unreliable for a specific data set. In addition, the weighting exponent m and also the random initial fuzzy partition of data may affect the reliability of a validation index.

Because of practical implications, we have applied the FCM by taking only a few samples of the weighting exponent m up to the value $m=2$. This parameter is, however, a continuous variable in the range $(1,\infty)$. Although a careful analysis of V_{CWB} is required for very large values of m, many application of FCM have been realized with a weighting exponent $m=2$, which is included in our study.

However, we realize that a more rigorous and theoretical study of V_{CWB} is necessary, especially to analyze the effect of the weighting factor α in the performance of V_{CWB}. In addition, to establish the performance of this index, testing on other challenging synthetic and real world data seems necessary.

Acknowledgements

We wish to thank the referees for their valuable comments and suggestions. This research is supported by the Science Foundation, The Netherlands, project LGN.22.2781.

References

Anderson, E., 1935. The irises of the Gaspe Peninsula. Bull. Amer. Iris Soc. 59, 2–5.

Backer, E., Jain, A.K., 1981. A clustering performance measure based on fuzzy set decomposition. IEEE Trans. Pattern Anal. Machine Intell. 3 (1), 66–75.

Bezdek, J.C., 1973. Fuzzy mathematics in pattern classification. Ph.D. dissertation, Cornell University, Ithaca, NY.

Bezdek, J.C., 1974. Cluster validity with fuzzy sets. J. Cybernet. 3 (3), 58–72.

Bezdek, J.C., 1975. Mathematical models for systematics and taxonomy. In: Estabrook, G. (Ed.), Proc. 8th Internat. Conf. Numerical Taxonomy. Freeman, San Francisco, CA, pp. 143–166.

Bezdek, J.C., 1981. Pattern Recognition with Fuzzy Objective Function Algorithms. Plenum, New York.

Bezdek, J.C., 1987. Partition structures: A tutorial. In: Bezdek, J.C. (Ed.), The Analysis of Fuzzy Information. CRC Press, Boca Raton, FL.

Dave, R.N., Bhaswan, K., 1992. Adaptive fuzzy c-shells clustering and detection of ellipses. IEEE Trans. Neural Networks 3 (5), 643–662.

Dunn, J.C., 1974. A fuzzy relative of the ISODATA process and its use in detecting compact, well-separated clusters. J. Cybernet. 3, 32–57.

Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics 7 (II), 179–188.

Fukuyama, Y., Sugeno, M., 1989. A new method of choosing the number of clusters for the fuzzy *c*-mean method. In: Proc. 5th Fuzzy Syst. Symp., pp. 247–250 (in Japanese).

Gindy, N.N.Z., Ratchev, T.M., Case, K., 1995. Component grouping for GT applications – a fuzzy clustering approach with validity measure. Internat. J. Prod. Res. 33 (9), 2493–2509.

Gunderson, R., 1978. Application of fuzzy ISODATA algorithms to star-tracker pointing systems. In: Proc. 7th Triannual World IFAC Congr., Helsinki, pp. 1319–1323.

Halgamuge, S.K., Glesner, M., 1994. Neural networks in designing fuzzy systems for real world applications. Fuzzy Sets and Systems 65 (1), 1–12.

Krishnapuram, R., Nasraoui, O., Keller, J., 1992. The fuzzy c spherical shells algorithm: A new approach. IEEE Trans. Neural Networks 3 (5), 663–671.

Libert, G. Roubens, M. 1983. New experimental results in cluster validity of fuzzy clustering algorithms. In: Janssen, J., Macrotorchino, J.F., Proth, J.M. (Eds.), New Trends in Data Analysis and Applications. North-Holland, Amsterdam, pp. 205–218.

Man, Y., Gath, I., 1994. Detection and separation of ring-shaped clusters using fuzzy clustering. IEEE Trans. Pattern Anal. Machine Intell. 16 (8), 855–861.

Pal, N.R., Bezdek, J.C., 1995. On cluster validity for the fuzzy *c*-means model. IEEE Trans. Fuzzy Syst. 3 (3), 370–379.

Pal, N.R., Bezdek, J.C., 1997. Correction to on cluster validity for the fuzzy *c*-means model. IEEE Trans. Fuzzy Syst. 5 (1), 152–153.

Windham, M.P., 1981. Cluster validity for fuzzy clustering algorithms. Fuzzy Sets Syst. 5, 177–185.

Windham, M.P., 1982. Cluster validity for the fuzzy *c*-means clustering algorithm. IEEE Trans. Pattern Anal. Machine Intell. 4 (4), 357–363.

Windham, M.P., Bock, H., Walker, H.F., 1989. Clustering information from convergence rate. In: Proc. 2nd Conf. Internat. Federation Classification Soc., Washington, DC, p. 143.

Xie, X.L., Beni, G.A., 1991. A validity measure for fuzzy clustering. IEEE Trans. Pattern Anal. Machine Intell. 13 (8), 841–847.