IEEE Access
Multidisciplinary ⁝ Rapid Review ⁝ Open Access Journal

Received 20 May 2024, accepted 2 June 2024, date of publication 14 June 2024, date of current version 26 June 2024.

*Digital Object Identifier 10.1109/ACCESS.2024.3414415*

RESEARCH ARTICLE

# A New Internal Validity Index for Fuzzy *c*-Means Algorithm

**KHAIRUL NURMAZIANNA ISMAIL[1], ALI SEMAN[2], AND KHYRINA AIRIN FARIZA ABU SAMAH[1], (Member, IEEE)**
[1]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM), Jasin, Melaka 77300, Malaysia
[2]College of Computing, Informatics and Mathematics, Universiti Teknologi MARA (UiTM), Shah Alam, Selangor 40450, Malaysia

Corresponding author: Ali Seman (aliseman@tmsk.uitm.edu.my)

This work was supported by the Universiti Teknologi MARA (UiTM), Malaysia.

**ABSTRACT** Fuzzy *c*-Means (FCM) is a popular clustering algorithm that can partition a set of objects into groups such that objects within a group are similar to each other and dissimilar to those in other groups. A validity index, either external or internal, is required to validate a cluster quality formed by the FCM algorithm. External validations require known class labels for measuring a cluster quality and serve as the clustering ground truth. In real-world data with unknown class labels, a cluster quality can be validated only via internal validations. A variety of internal validation measures with different scoring models have been developed, including minimum model, maximum model, and range model with minimum to maximum scores. No internal validation measure proposed thus far is associated with a model ranging from 0 to 1, like the clustering ground truth (external validation). Therefore, a new internal validation, namely, the fuzzy validity index (FVI), is proposed. Experimental results based on several cluster properties demonstrated that the FVI is highly promising. Overall, the scores of the FVI were comparable to the scores obtained by the external validity index, i.e., F-measure. Statistically, the correlation coefficient between the FVI and F-measure was high (around 0.8 and above), indicating their similarity. Therefore, the FVI could potentially serve as the ground truth for measuring the cluster quality of FCM.

**INDEX TERMS** Fuzzy clustering, fuzzy *c*-means, internal validity index, fuzzy validity index.

## I. INTRODUCTION

Clustering refers to the problem of partitioning a set of objects into groups such that objects within a group are similar to each other and dissimilar to those in other groups [1]. It is an unsupervised learning technique categorized into two broad classes: hard (also known as crisp) and soft (also known as fuzzy) clustering approaches. In the hard-clustering algorithm, an object is assigned to a cluster based on a partition matrix that contains $\{0,1\}$. A value of 1 indicates that the object belongs to the cluster, whereas 0 indicates otherwise. In the soft- or fuzzy-clustering approach, assigning an object to a cluster is based on the fuzzy partition matrix values as well. However, unlike the hard-clustering algorithm, the value of the fuzzy partition matrix ranges from 0 to 1, representing the degree of confidence that an object belongs to a cluster. Therefore, the closer the fuzzy membership value to 1, the stronger the confidence in its membership for that cluster [2], [3]. The FCM algorithm is widely used in clustering numerical data. FCM is derived from the hard $k$-means algorithm, which was introduced by MacQueen [4]. FCM was initially introduced by Dunn [3] and later improved upon by Bezdek [2]. It is based on fuzzy theory, which allows for more flexibility in assigning the degree of membership to each object. The algorithm depends on two parameters: the number of clusters $k$ and a parameter $m$ (also known as the weighting exponent or degree of fuzziness). The parameter $m$ controls the degree of membership assigned to each object, allowing it to differentiate between hard- and soft-clustering approaches. When $m$ is greater than 1, it indicates a greater degree of fuzziness, while $m$ equal to 1 indicates a crisp approach. In cluster analysis, two validations are required in the context of FCM: validation for determining the correct number of clusters $k$ and validation

The associate editor coordinating the review of this manuscript and approving it for publication was Wai-Keung Fung.

for measuring the cluster quality [5], [6]. The first validation aims to automatically detect the correct number of clusters since one of the conditions of FCM is to specify the number of clusters beforehand. The second validation aims to measure the quality of the clusters formed by FCM, which generally relies on three options as follows:

a) The validation is based only on external information (known as external validity index) such as class labels and the number of classes. The Rand index [7], Fowlkes-Mallows index [8], and F-measure [9] are some examples of commonly used external validity indices.
b) The validation is based only on internal information (known as internal validity index). It measures how well the results fit between the data and the expected structure using quantities and features inherent in the dataset [1]. Many internal validity indices have been reported, and the Silhouette index is one of the commonly used internal validity indices to measure the cluster quality formed by FCM.
c) The validation is based on external and internal validity indices measured together.

The last option is more ideal and practical and is extensively used by researchers for a wider spectrum of evaluation. In this context, external and internal validations complement each other for several reasons in particular for known class labels. Furthermore, external validations typically serve as the ground truth for measuring cluster quality [10], [11], [12]. The ground truth refers to the actual nature of clustering groups formed by any clustering model to validate whether it is clustered accurately or inaccurately. It assesses the degree to which the clustering solution matches the external knowledge of the data, such as the class labels assigned by experts in the domain. The ground truth for clustering models is typically based on a confusion matrix, which is introduced to measure classification results. In essence, the matches between the clustering solution and external knowledge (class labels) are calculated based on supervised learning. For obtaining a clustering accuracy, the number of positive objects that are correctly clustered (true positive, TP) and the number of negative objects that are correctly clustered, known as (true negative, TN) are divided by all objects, including the positive objects that are incorrectly clustered (false positive, FP) and the negative objects that are incorrectly clustered (false negative, FN). As a result, the clustering accuracy could be 0% (all objects were incorrectly clustered), 100% (all objects were correctly clustered), or any values in between 0% to 100%. Thus, to benchmark a cluster quality, the external validations typically report a score, which is based on a bounded index from 0 to 1. Thus, this index serves as the ground truth for measuring the cluster quality.

Nevertheless, in real-world applications, most data are devoid of external information with no class labels. In this case, internal validation is the only option for measuring quality clusters, and therefore, the external validation is out of scope. Basically, the internal validation measures a cluster quality based on internal information of the clustering itself, typically object's distance. It measures the degree of intra-cluster similarity, i.e., similarity between objects within a cluster, known as compactness, and inter-cluster similarity, known as separation. For a well-separated cluster, internal validation is characterized by objects that are significantly closer to other objects in the same cluster than to objects in any other cluster. One of the popular internal validity indices, which is associated to FCM is Silhouette index [13]. The Silhouette utilizes objects' average distance in a cluster as its compactness measure and objects' average distance in different clusters as its separation measure. Relatively, the Silhouette generates three scores: -1, 0, and +1, corresponding to incorrect clustering, cluster overlap, and adequate clustering. Thus, the Silhouette also based on the bounded-index model that measures a cluster quality ranging from a minimum of -1 to a maximum of +1.

Notably, external and internal validations are relatively distinctive by nature. External validation is based on external information, which is known class labels and then serves as the ground truth. Internal validation is based on internal information, which is inherent structure of data with unknown class labels, cluster properties, and no ground truth. Therefore, internal validations come with various scoring models and are not standardized. Unlike external validations, internal validations come with both bounded- and unbounded-index models. In fact, the unbounded index comes with either a minimum or maximum score and no boundary. That is, to benchmark a cluster quality, the best clustering results will be based on the highest minimum score obtained such as that proposed by Fukuyama and Sugeno [14], VK index [15], XB index [16], and partition entropy index [17] or the highest maximum score, such as that proposed by Dunn index [3], Calinski-Harabasz index [18], and Partition coefficient and exponential separation [19]. Only a few internal validation indices with the bounded-index model have been proposed that are constrained within a minimum and a maximum score such as the external validity indices, like the Silhouette index, fuzzy Silhouette index [20], and generalized intra-inter Silhouette index [21]. On the contrary, the Davies-Bouldin index [22] is a bounded-index model with values from 1 to 0. To the best of our knowledge, thus far, there are no other internal validation indices for FCM with the bounded-index model from 0 to 1, as characterized by external validations.

Motivated by the aforementioned, a new internal validity index called fuzzy validity index (FVI) is proposed for FCM. The FVI is based on the bounded-index from 0 to 1, similar to external validity indices as it serves as the ground truth for measuring a cluster quality. In summary, this study makes the following four contributions to the domain of clustering, particularly for clustering validation:

a) A novel internal validity index, namely, the fuzzy validity index, or FVI, is introduced by using the bounded-index model, ranging from 0 to 1, likewise the ground truth of the external validations.

b) The FVI is primarily established by a new method called the adjusted fuzzy membership matrix derived from the fuzzy membership of FCM. This is because the fuzzy membership matrix obtained by FCM may not precisely be used for FVI to represent cluster properties, especially for cluster overlap, elliptical or mixture of shapes, unbalanced clusters, and unknown cluster properties, such as real-world data.

c) The FVI is also derived from a new concept of compactness and separation, defined by using fuzzy membership values. This compactness-separation ratio (CSR) is the underlying idea of the FVI and is defined through three possibilities: the perfect case, best case, and worst case.

d) The FVI could potentially serve as the ground truth for internal validations to measure the cluster quality of FCM regardless of missing class labels and cluster properties.

The remainder of the paper is structured as follows. Section II provides a critical review in particular for validation methods commonly used for FCM. Section III elaborates on the formalization of FCM as a preliminary background to propose its validation method. Section IV introduces the concept of CSR and formalizes the new validity index. Section V describes tools, parameter setup, datasets used for the experiments, and the evaluation method chosen for comparison. Section VI demonstrates and discusses comprehensive results for each cluster property. Section VII presents the final discussions of the research, and finally, section VIII discussed possible directions for further improvement.

## II. RELATED WORK

Many validation methods have been reported and used for validating FCM. About six external validity indices have been introduced and used for measuring the cluster quality of FCM. Table 1 enumerates the external validity indices with the bounded-index model from 0 to 1, except adjusted Rand Index (ARI), with the bounded index from −1 to 1.

**TABLE 1. List of proposed external validity indices and used for validating FCM.**

| No. | Validity Index | Year | Range |
|---|---|---|---|
| 1 | Rand Index [7] | 1971 | [0,1] |
| 2 | Fowlkes-Mallow [8] | 1983 | [0,1] |
| 3 | Adjusted Rand Index [15] | 1985 | [-1,1] |
| 4 | Fuzzy Rand Index [16] | 1988 | [0,1] |
| 5 | F-Measure Index [9] | 1999 | [0,1] |
| 6 | Normalized Mutual Info [17] | 2002 | [0,1] |

To the best of our knowledge for the internal validity index, about 16 validity indices have been commonly used for FCM since 1973. Table 2 presents a list of internal validity indices introduced for measuring the cluster quality of FCM. However, most of these indices are based on an unbounded index, whether they use minimum or maximum values to indicate a good cluster quality. Notably, only Silhouette index [13] and Davies-Bouldin index [22] use a bounded-index model to indicate the minimum to maximum cluster quality. However, the ranges for both indices are different. The Silhouette index utilizes a range of −1 to 1, whereas the Davies-Bouldin index indicates that the minimum value is 1 and the maximum value is 0. Thus, no internal validity index has been proposed for FCM that uses the bounded index from 0 to 1 like the external validity indices.

Table 3 lists some studies reported in the past decade that have applied FCM and employed certain validity indices for measuring the quality clusters. Notably, the Silhouette index is the most popular method, particularly for internal validation. This metric is somewhat synonymous with the *k*-mean-type clustering algorithm traditionally used to validate its clustering results. The Silhouette index is motivated by distance-based scores and is derived from the average distance of intra-cluster objects (compactness) and the average distance of inter-cluster objects (separation). A notable advantage of the Silhouette index is that it can also be used to validate FCM by using distance objects and not the fuzzy values.

Heretofore, only two validity indices are found to be related directly to FCM with the attempt to manipulate fuzzy values in order to validate the cluster quality as follows:

a) The fuzzy Silhouette Index (FSI) Campello and Hruschka [13] proposed the FSI based on the framework of the Silhouette index. The authors explicitly implemented the idea of compactness and separation via the fuzzy membership values associated with objects. FSI introduces weightage by using the fuzzy membership degree of the corresponding object to its first and second best matching fuzzy cluster. The FSI aims to improve the performance of its crisp counterpart in detecting regions by focusing on objects in the vicinity of a cluster rather than objects in an overlapped area.

b) The generalize intra-inter Silhouette index (GIIS) The GIIS was introduced by Rawashdeh and Ralescu [14]. It can be applied to both crisp and fuzzy approaches as it combines the Silhouette index and fuzzy partition matrix. In this context, the GIIS uses the intra-distance measure (compactness) to any of the clusters and inter-distance measure (separation) with respect to any of the two clusters. GIIS uses Boolean conjunction and disjunction to determine the intra- and inter-cluster distance from the cluster assignment from the crisp and fuzzy partition matrices.

In conclusion, FS and GIIS are internal validity indices that aim to manipulate the fuzzy partition matrix to measure the quality of clusters of FCM. Both indices inherit the bounded index established by Silhouette, which is from −1 to 1. However, it is found that the overall scores produced by both indices are almost similar to the scores obtained by the Silhouette index. In terms of scoring, both indices produce

**TABLE 2. List of internal validity indices for validating FCM.**

| No. | Validity Index | Year | Range |
|---|---|---|---|
| 1 | Dunn Index [3] | 1973 | Unbounded, max |
| 2 | Calinski-Harabasz Index [18] | 1974 | Unbounded, max |
| 3 | Davies-Bouldin Index [22] | 1979 | Bounded, $[1, 0]$ |
| 4 | Fuzzy Entropy [2] | 1981 | Unbounded, min |
| 5 | Partition Coefficient [2] | 1981 | Unbounded, min |
| 6 | Silhouette Coefficient [13] | 1987 | Bounded, $[-1, 1]$ |
| 7 | Fukuyama and Sugeno [14] | 1989 | Unbounded, min |
| 8 | Xie-Beni Index [16] | 1991 | Unbounded, min |
| 9 | Partition Entropy [17] | 1992 | Unbounded, min |
| 10 | Fuzzy Entropy [25] | 1993 | Unbounded, min |
| 11 | Fuzzy Relative Entropy [26] | 1998 | Unbounded, min |
| 12 | VK Index [15] | 1998 | Unbounded, min |
| 13 | PCAES [27] | 2005 | Unbounded, max |
| 14 | Fuzzy Silhouette Index [20] | 2006 | Bounded, $[-1, 1]$ |
| 15 | Validity Index based on Mountain Function and subtractive algorithm [28] | 2010 | Unbounded, min |
| 16 | The Generalize Intra-Inter Silhouette Index [21] | 2012 | Bounded, $[-1, 1]$ |

**TABLE 3. Publications related to FCM application and validity indices used within the last 10 years.**

| No. | Research | Year | Validity Indices used |
|---|---|---|---|
| 1 | Pomponi & Vinogradov [29] | 2013 | External: None<br>Internal: Silhouette Index, Davies-Bouldin and Dunn |
| 2 | Rahman & Islam [30] | 2013 | External: F-Measure, Entropy and Purity<br>Internal: Silhouette |
| 3 | Banerjee, Badr & Al-Shammari [31] | 2014 | External: None<br>Internal: Dunn, Davies-Bouldin and Silhouette |
| 4 | Civicioglu et al., [32] | 2014 | External: None<br>Internal: Davies-Bouldin, Silhouette, Dunn and R-Squared |
| 5 | Banerjee, Choudhary & Pal [33] | 2015 | External: Rand, Jaccard and Fowles-Mallow<br>Internal: Silhouette, Davies-Bouldin and Dunn Indices |
| 6 | Fernandez et al., [34] | 2016 | External: None<br>Internal: Silhouette |
| 7 | Bakar et al., [35] | 2016 | External: F-Measure<br>Internal: Silhouette |
| 8 | Nurjanah, Dewanto & Sari [36] | 2017 | External: None<br>Internal: Silhouette |
| 9 | Renjith, Sreekumar & Jathavedan [37] | 2018 | External: Entropy<br>Internal: Silhouette, Dunn |
| 10 | Rajkumar, Yesubabu & Subrahmanyam [38] | 2019 | External: None<br>Internal: Silhouette and Dunn |
| 11 | Mishra & Bhoi [39] | 2020 | External: Precision, Recall and F-Score (F-Measure)<br>Internal: Davies-Bouldin, Dunn, Silhouette |
| 12 | Eliyanto & Surono [40] | 2021 | External: Accuracy, Purity and Rand Index<br>Internal: Silhouette and Davies-Bouldin |
| 13 | Rochman, Miswanto & Herry [41] | 2022 | External: None<br>Internal: Silhouette and Davies-Bouldin |
| 14 | Altameem & Hafez [42] | 2022 | External: F-Measure and Rand<br>Internal: Davies-Bouldin, Silhouette and Dunn |
| 15 | Wenhao et al., [43] | 2023 | External: None<br>Internal: Davies-Bouldin, Silhouette and Calinski & Harabasz |
| 16 | Zhu et al., [44] | 2023 | External: F-Measure, Accuracy, Precision and Recall<br>Internal: Silhouette |
| 17 | Bhandari & Pahwa [45] | 2023 | External: None<br>Internal: Silhouette, Dunn, Xie-Beni and Davies-Bouldin |

scores far lower than the scores of ground-truths set by any external validity indices.

## III. FCM ALGORITHM

FCM is a generalization of the classical $k$-means algorithm, which moves from a hard partition to a soft partition. In other words, FCM is a partitioning algorithm that offers a fuzzy membership partition matrix in which objects can belong to more than one cluster. In addition, FCM may be better than $k$-means in terms of offering a globally optimized procedure in its cost function. Therefore, to partition the dataset $X$ such that $X = \{X_1, X_2, \ldots, X_n\}$, FCM minimises the cost function in Equation (1):

$$J_{FCM}(\mu, Z) = \sum_{i=1}^{n} \sum_{l=1}^{k} \mu_{il}^{m} d(x_i, z_l) \quad (1)$$

With respect to $\mu$ (a fuzzy $k$-partition matrix) and $Z$ (a set of cluster centers), where $m$ is a weighting exponent that satisfies, $Z_l$ is the centroid of cluster, $\mu$ the degree of membership of object $x_i$ belonging to cluster $l$, $d(x_i, z_l)$ is a distance measure between object $x$ and cluster center $z$, and $k$

is the number of clusters. The fuzzy $k$ partition matrix, $\mu$ is subject to the following constraints:

a) $\mu_{il} = [0, 1] \, \forall \, i, l$.
b) $\sum_{i=1}^{k} \mu_{il} = 1, \forall \, l$ and
c) $0 < \sum_{i=1}^{n} \mu_{il} < n, \ \forall \, i$.

In general, the steps for FCM are as follows:

**Step 1:** Generate the initial center $Z_l$ $(l = 1, 2, \ldots, k)$.

**Step 2:** Compute the fuzzy partition matrix $\mu_{il}$ using Equation (2).

$$\mu_{il} = \frac{1}{\left(\sum_{l=1}^{k}\left(\frac{d(x_i,z_c)}{d(x_i,z_l)}\right)^{\frac{2}{(m-1)}}\right)} \tag{2}$$

**Step 3:** Compute the fuzzy center $Z_l$ using Equation (3).

$$Z_l = \frac{\sum_{i=1}^{n} \mu_{il}^{m} x_i}{\sum_{i=1}^{n} \mu_{il}^{m}} \tag{3}$$

and update the fuzzy membership matrix $\mu_{il}$ to $\tilde{\mu}_{il}$ according to Equation (2).

**Step 4:** If $\max_{il} |\mu_{il} - \tilde{\mu}_{il}| < \varepsilon$, then stop, otherwise go to Step 3, where $\varepsilon$ is a termination criterion between 0 and 1.

## IV. PROPOSED INTERNAL VALIDITY INDEX

### A. COMPACTNESS-SEPARATION RATIO

The idea of the FVI is motivated by the underlying concept of compactness and separation imposed by the Silhouette and its variants (fuzzy validity indices based on the Silhouette index). Unlike the Silhouette and its variants, the proposed internal validity index measures compactness and separation at the individual object level rather than at the cluster level. In this context, compactness and separation are viewed as complementary pairs that work together cohesively. Therefore, this concept is referred to as the compactness-separation ratio, or CSR.

Therefore, an object is considered to have a perfect CSR when the degree of fuzzy membership assigned to its cluster is 1. This means that the CSR is 100:0, which can be interpreted as the percentage of compactness for that object being 100% and the separation being 0%, resulting in well-separated clusters. Consequently, this CSR may vary for each object with regard to the degree of fuzzy membership. In general, if an object has a degree of fuzzy membership of 0.92, it can be interpreted as a CSR of 92:8, meaning that 92% of the object belongs to that cluster and the remaining 8% belongs to other clusters. The ratio is based on a two-cluster basis: the cluster that the object belongs to (in the fuzzy context, the object with the highest degree of fuzzy membership) and the other cluster with the highest degree of membership among the object's remaining membership degree. For example, consider an object $x$ with the degree of its fuzzy membership in three clusters, namely $C_1$, $C_2$, and $C_3$ as 0.15, 0.80, and 0.05, respectively. As $C_2$ has the highest degree of membership, the object belongs to $C_2$, with a compactness ratio of 80%. On the other hand, the highest degree of membership among the remaining clusters ($C_1$ and $C_3$) is 0.15, indicating a CSR of 15%. Thus, its CSR is 80:15. In this context, each object may fall into one of the following ratio cases:

- *Perfect-case* - This occurs when an object has a compact-separation ratio of 100:0. This scenario is not a normal case in fuzzy clustering as it indicates a hard (crisp) clustering.
- *Best-case* - This occurs when an object has a higher percentage of compactness than the percentage of separation. Relatively, the percentage of compactness is greater than 50% but less than 100%, whereas the percentage of separation is greater than 0% but less than 50%.
- *Worst-case* - This occurs when an object has the percentage of compactness which is less than or equal to 50% of the percentage of separation. Relatively, the percentage of compactness may be less than or equal to 50% and its separation may be greater than or equal to 50%. However, the worst-case compact-separation ratio is also dependent on the degree of fuzzy membership of the object in that cluster and it must be the highest among the degree of fuzzy membership of the other clusters. For example, if there are three clusters and an object has the degree of fuzzy membership of 0.40, 0.35, 0.25 for $C_1$, $C_2$ and $C_3$ respectively, the compact-separation ratio is 40:35 and the object is assigned to Cluster 1 because the degree of fuzzy membership is the highest.

Figure 1 depicts the concept of compactness and separation at the object level for a particular CSR. For simplicity, this example is based on three clusters, and the sum of fuzzy membership values assigned to these three clusters must be 1 and distributed to each cluster according to Equation (2). The circle denotes a single object, which is divided into three sections based on its fuzzy membership values for each cluster. In the crisp clustering context, the object belongs to Cluster 1, which has the highest degree of fuzzy membership. The CSR is directly adapted from the fuzzy membership value with respect to the crisp cluster and the highest fuzzy membership values for the remaining clusters. Therefore, Figures 1(a) and 1(b) illustrate the best CSR in two cases with the ratios of 95:3 and 75:15, respectively, while Figures 1(c) and 1(d) illustrate the worst CSR in two cases with the ratios of 50:40 and 40:35, respectively.

### B. NEW FUZZY VALIDITY INDEX

The new FVI is primarily based on the concept of CSR, as discussed above. This ratio can be computed for each object as each has its own fuzzy membership that contributes to cluster formation. FCM applies fuzzification by using the reciprocal of the distance between objects and cluster center to determine the degree of fuzzy membership for each object and cluster, which is shown in Equation (2). By using the distance between object and center as the base of the power and $\frac{2}{(m-1)}$ as the exponent, the fuzzy membership

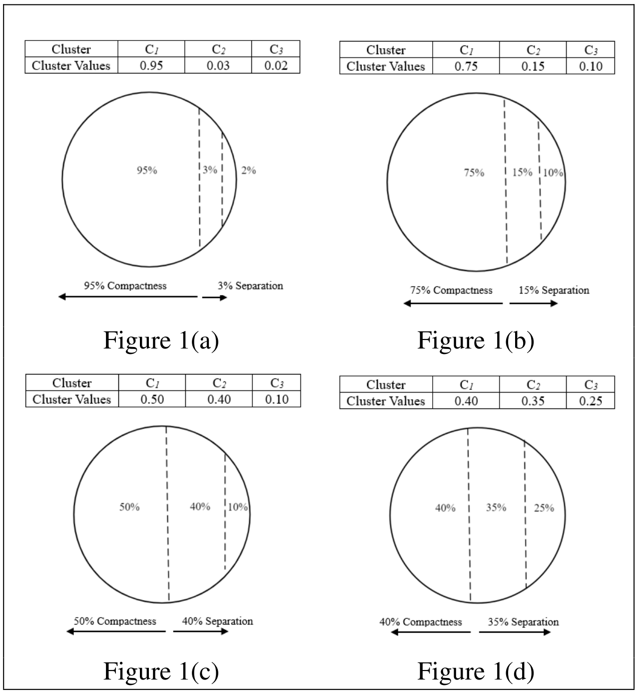


**FIGURE 1. Overall concept of CSR. (a) best compactness–separation case with a ratio of 95:3. (b) best compactness-separation case with the ratio of 75:15. (c) worst compactness-separation case with the ratio of 50:40, and (d) worst compactness-separation case with the ratio of 40:35.**

values decline rapidly when $m$ increases. In this context, two significant consequences are observed:

a) The fuzzy membership values obtained may only be representative of that cluster if it is a well separated cluster, but not for other cluster properties such as cluster overlap, elliptical or mixture of shapes, unbalance cluster and unknown cluster properties like real-world data.

b) When the weighting exponent $m$ increases toward $\infty$, the fuzzy membership values become fuzzier and gradually move away from 1 (crisp cluster). Thus, it is possible that the objects may not belong to a particular cluster even if their fuzzy membership values indicate otherwise. Therefore, it may not be ideal to rely solely on fuzzy membership when formulating a new internal cluster validity index.

Thus, an empirically adjusted fuzzy membership value is required to overcome these issues. Here, the adjusted fuzzy membership value is proposed to be used as the primary input for formulating the FVI. The adjusted fuzzy membership is obtained by using the fuzzy membership $\mu$ as the base of the power and $\frac{m}{(m-1)^{-1}}$ as the exponent, $m-1^{-1}$ is the exponent introduced by fuzzy $k$-modes [46] and the $k$-approximate modal haplotypes clustering algorithm [47]. Figure 2 illustrates the effect of exponents that decline rapidly for the exponent: $\frac{2}{(m-1)}$ (used by FCM) and also $\frac{1}{(m-1)}$ (used by fuzzy k-modes). The exponent $\frac{m}{(m-1)^{-1}}$, proposed for the adjusted fuzzy membership, increases steadily. By imposing $m$ over $m-1^{-1}$, such that $\frac{m}{(m-1)^{-1}}$, the adjusted fuzzy membership values are now more representative of whether an object belongs to its cluster or other clusters, and therefore, can resolve the two issues mentioned above. Although the empirical analysis of the experimental results is very promising, the theoretical aspect of this scenario needs to be further investigated.

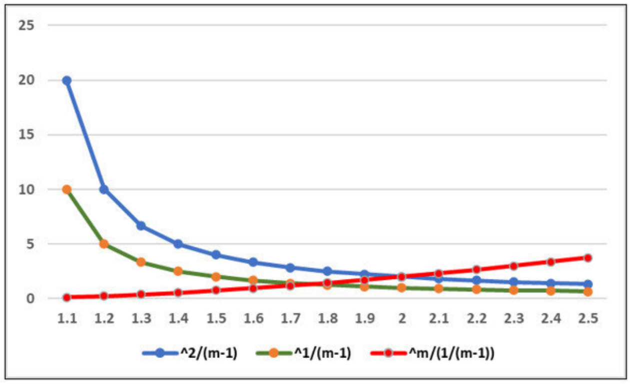


**FIGURE 2. The effect of exponents declined rapidly for the exponents: $\frac{2}{(m-1)}$, $\frac{1}{(m-1)}$, and $(m-1)^{-1}$.**

As a result of above explanation, the overall steps to obtain the FVI shown as Figure 3.

Roughly, the steps based on Figure 3 describe as follows.

**Step 1:** Calculate adjusted fuzzy membership (AFM). The AFM is a new method that is derived from the fuzzy membership of FCM.

**Step 2:** Apply the concept of CSR by finding the highest and the second highest AFM scores. The highest score represents the degree of compactness, whereas the second highest score represents the degree of separation. In the case of the perfect-case CSR, the degree of compactness is 1.0, and the degree of separation will be 0. It follows for the other scenarios of CSR discussed earlier.

**Step 3:** Calculate a standard deviation based on the CSR above. The score is called standard deviation of compactness–separation (SDCS).

**Step 4:** Assign a weightage score for each SDCS score. If the SDCS score is greater or equal to 0.5 and less than or equal to 1.0, the weightage is 1.0. If the SDCS score is greater or equal to 0.1 and less than 1.5, the weightage is the highest score from AFM for that particular object. Otherwise, the weightage is 0.

**Step 5:** Calculate the average of the weightage scores.

**Step 6:** Obtain the FVI, which is the average weightage score.

The formalization of the FVI is described as follows.

Let $\delta = \delta_1, \delta_2, \ldots, \delta_n$ be a set of $n \times k$ AFM matrix,and $\delta_i$ be represented with its clusters as $[\delta_{i1}, \delta_{i2}, \ldots, \delta_{ik}]$; therefore, to obtain $\delta_{il}$ as in Equation (4):

$$\delta_{il} = \frac{\mu_{il}^{\frac{m}{(m-1)^{-1}}}}{\sum_{l=1}^{k} \mu_{il}^{\frac{m}{(m-1)^{-1}}}} \tag{4}$$

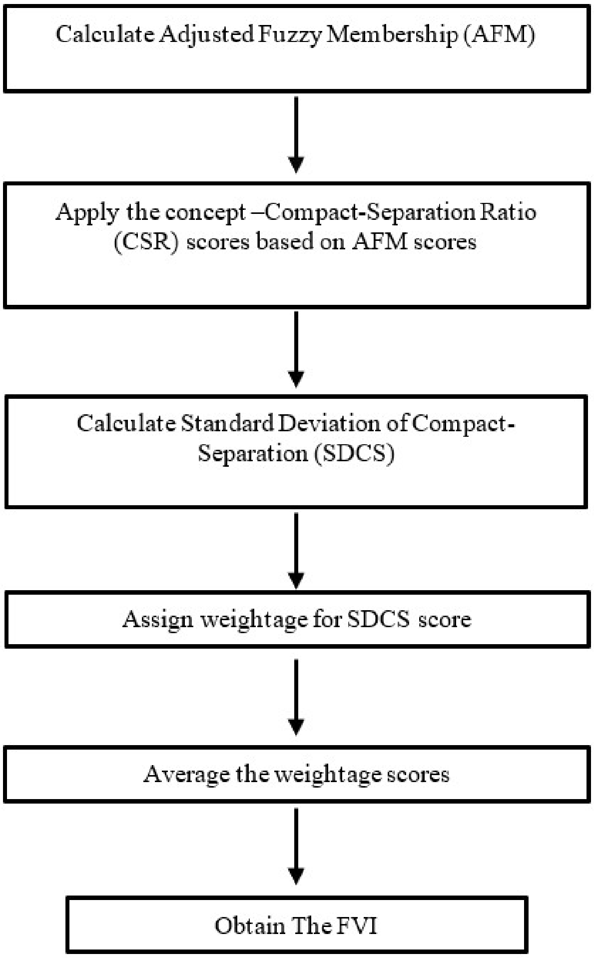


**FIGURE 3.** Steps for obtaining the FVI.

where, $\mu_{il}$ is the fuzzy partition matrix and $m$ is the weighting exponent of FCM and is subject to the constraints $\delta_{il} \epsilon\ [0, 1]$ and $\sum_{l=1}^{k} \delta_{il} = 1$, for all $i$.

Next, the FVI can be obtained in Equation (5):

$$FVI = \frac{1}{n}\sum_{i=1}^{n} \gamma_i,\ where\ 1 \leq i \leq n \tag{5}$$

where, $\gamma$ is a FVI weightage given for each object, $x_i$ as denoted in Equation (6) and is subject to SDCS as Equation (7):

$$\gamma_i = \begin{Bmatrix} 1, & \text{if } SDCS_i \geq 0.5 \text{ and } SDCS_i \leq 1.0; \\ \max^{\delta_{il} \text{ for } 1 \leq l \leq k}, & \text{if } SDCS_i \geq 0.1 \text{ and } SDCS_i < 0.5; \\ 0, & \text{Otherwise;} \end{Bmatrix} \tag{6}$$

$$SDCS_i = \sqrt[2]{\sum_{f=1}^{2} (q_f - \bar{q})^2} \tag{7}$$

$q_1 = max^{\delta_{il}}$ (the highest AFM score that represents the compactness), $q_2 = max^{\delta_{it}}; l \neq t$ for $1 \leq l \leq k$ (the second highest AFM score that represents the separation), and $\bar{q}$ is the mean for $q_1$ and $q_2$.

In this context, the CSR is represented by $q_1$:$q_2$. This principle follows the Silhouette index, particularly the principle of separation. The Silhouette index compares the average distance of an object from a cluster, which is usually called $a$, with the average distance of other objects from the other clusters, which is called $b$, based on the minimum distance. In the FVI context, compactness and separation are based on the maximum fuzzy membership values. Table 4 presents a simulation of adjusted fuzzy membership values that represent the scenario regarding SDCS and the weightage given based on cutoff points of 0.5 (represented by cases 1–9), 0.1 (represented by cases 10–43), and 0 (represented by cases 44–52). The example is based on three clusters, where Cluster 1 has the highest degree of membership, and Cluster 2 has the second highest degree of membership.

The weightage $\gamma$ of FVI is derived from the concept of compact-separation explained above, which is represented by $q_1$ (the highest value of the adjusted fuzzy membership) and $q_2$ (the highest value of the adjusted fuzzy membership of the remaining clusters).

**TABLE 4.** Simulated values of SDCS and their cutoff values.

| No. | Cluster 1 | Cluster 2 | Cluster 3 | SUM | $q_1$:$q_2$ | SDCS |
|---|---|---|---|---|---|---|
| 1 | 1.00 | 0.00 | 0.00 | 1 | 100:0 | 0.71 |
| 2 | 0.90 | 0.07 | 0.03 | 1 | 90:7 | 0.59 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 9 | 0.83 | 0.11 | 0.06 | 1 | 83:11 | 0.51 |
| 10 | 0.82 | 0.12 | 0.06 | 1 | 82:12 | 0.49 |
| 11 | 0.81 | 0.13 | 0.06 | 1 | 81:13 | 0.48 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 43 | 0.49 | 0.34 | 0.17 | 1 | 49:34 | 0.11 |
| 44 | 0.48 | 0.35 | 0.17 | 1 | 48:35 | 0.09 |
| 45 | 0.47 | 0.35 | 0.18 | 1 | 47:35 | 0.08 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 52 | 0.40 | 0.40 | 0.20 | 1 | 40:40 | 0.00 |

Table 5 shows an example of how these values are obtained.

**TABLE 5.** An example of $\mu$, $\delta$, SDCS, and $\gamma$ of an object.

| Object | $\mu$ (Fuzzy membership) | | | $\delta$ (Adjusted fuzzy membership) | | | SDCS | $\gamma$ |
|---|---|---|---|---|---|---|---|---|
| | $C_1$ | $C_2$ | $C_3$ | $C_1$ | $C_2$ | $C_3$ | | |
| $x_1$ | 0.478 | 0.476 | 0.046 | 0.500 | 0.495 | 0.005 | 0.004 | 0 |
| $x_2$ | 0.541 | 0.404 | 0.035 | 0.657 | 0.340 | 0.003 | 0.224 | 0.657 |
| $x_3$ | 0.832 | 0.091 | 0.077 | 0.980 | 0.012 | 0.008 | 0.685 | 1 |

Table 5 provides a comprehensive example for each case and shows the value of fuzzy membership, adjusted fuzzy membership based on Equation (4), SDCS explained in Equation (7) and $\gamma$ that is obtained via Equation (6). It should be noted that the $m$ in this example is 2.0. For simplicity, all objects belong to Cluster 1; therefore, $q_1$ represents Cluster 1 and $q_2$ represents Cluster 2. For object $x_1$, the highest AFM (0.500) belongs to $q_1$ and the highest AFM of the remaining clusters (0.495) belongs to $q_2$, resulting in a CSR of roughly 50:50. The value of SDSC is 0.004 (based on Equation (7)).

When the value of SDCS is less than 0.1, it falls under the worst-case CSR and is assigned a weightage of 0. For object $x_2$, the highest AFM (0.657) is $q_1$ and the highest AFM of the remaining clusters (0.340) is $q_2$, resulting in a CSR of roughly 65:34. The value of SDSC is 0.224. When the value of SDCS is greater than or equal to 0.1, it falls under the best-case CSR and is assigned a weightage of 0.657. For object $x_3$ the highest AFM (0.980) is $q_1$ and the highest AFM of the remaining clusters (0.012) is $q_2$, resulting in a CSR of roughly 98:1. The value of SDSC is 0.685. When the value of SDCS is greater than or equal to 0.5, it falls under the perfect-case CSR and is assigned a weightage of 1.

The pseudo code for obtaining the FVI is described as in Algorithm 1.

**Algorithm 1** FVI (X, k, m, $\mu$)

**Input:** Dataset $X$, number of clusters $k$, weighting exponent $m$, and fuzzy membership values $\mu$.
**Output:** A new validity index, FVI.

```
for each C_l do
    for each X_i do
        Calculate δ_il as Equation 4
        Calculate SDSC as Equation 7
        Calculate γ_i as Equation 6
        Calculate FVI as Equation 5
    end for
end for
```

## V. EXPERIMENTAL SETUP

The experiments were carried out using the MATLAB built-in FCM with default clustering termination conditions, except for the maximum number of iterations that were set to 5000. The weighting exponent $m$ was set for each experiment from 1.1 to 2.5. The number of iterations was set to 5000 to ensure that optimum clustering results were obtained for each experiment. This option was initially suggested by Steinley [48] and was recently applied by Franti and Sieranoja [49] for their six clustering benchmark experiments. The experiments were rigorously carried out for $m$ values ranging from 1.1 to 2.5. To simplify reports and discussions, this study focused on a specific range of values for $m$, which were 1.1, 1.5, 2.0, and 2.5. The results for the other values of $m$ are also provided in Appendix. For the test cases, two categories of datasets, namely, synthetic and real-world datasets, were used for the experiments. During the experiments, FVI was calculated from the FCM partition matrix, $\mu$ and evaluated along with F-measure (FMI), Silhouette (SIL), and GIIS.

### A. DATASET

Overall, 35 datasets were used in the experiments, with Datasets 1 to 21 falling under the synthetic data category and Datasets 24 to 35 falling under the real-world data category. Table 6 provides a summary of these datasets. The synthetic category comprises datasets proposed by [49] for benchmarking the performance of $k$-mean clustering algorithm. These datasets represent six basic properties that can be applied methodologically for FCM to measure its cluster quality. In total, seven cluster properties, including the real-world datasets, were used for the experiments, which are shown as follows:

a) Number of clusters (Datasets 1 - 13). For the number of cluster property, an additional set of 10 datasets (Datasets 1-10) was purposely created and derived from the original dataset A1 by gradually increasing the number of clusters by 1, starting from a minimum of 2 clusters to 10 clusters, and another dataset had 15 clusters.
b) Cluster overlap (Dataset 14 - 17)
c) Dimensions and cluster overlap (Dataset 18 - 20)
d) Dimensions (Dataset 21)
e) Cluster structure (Dataset 22)
f) Unbalance data (Dataset 23)
g) Real-world datasets, 12 datasets (Dataset 24 - 35) were chosen from UCI repository

### B. EVALUATION METHOD

As discussed in the previous section, external validity indices, such as FMI and others were used as ground truth to measure cluster quality where known class labels are assigned to each object and the number of clusters is set by FCM beforehand. However, for unknown class labels, internal validity indices such as the popular Silhouette index were used to measure cluster quality. Since FVI is based on an internal validity index, the Silhouette index was used as the baseline benchmark for comparison. Another internal validity index, which was the GIIS, was also included in the comparison owing to its design based on fuzzy membership values obtained by FCM that were similar to FVI. Ideally, the scores of the internal cluster indices should be identical to the ground truth benchmark of external validity scores. However, this may not always be possible, and the aim is to close the gap between the two.

Therefore, three benchmarks were set for comparing FVI. The first benchmark was the ground truth benchmark represented by the FMI. The FMI was chosen because (1) It is the newest external validity index compared with others such as Rand index, adjusted Rand index, and Fowlkes–Mallows index and (2) It provides a balanced evaluation by combining precision and recall with harmonic mean in a single measure [57], [58]. The Fowlkes–Mallows index also combines precision and recall but with geometric mean, and our recent report [59] showed that FMI is marginally higher than Fowlkes–Mallows index. The aim was for the scores obtained by FVI to get closer to those obtained by the FMI. The second benchmark was the baseline benchmark represented by the Silhouette index. The aim was for the scores obtained by FVI to be farther from the scores obtained by the Silhouette index. The third benchmark was the fuzzy-based validity benchmark represented by the GIIS. The aim was for the scores obtained by FVI to be better than

**TABLE 6. Summary of datasets.**

| Dataset | Dataset Name | Properties | No. of instances | No. of classes | No. of instances per class | Source |
|---|---|---|---|---|---|---|
| 1 | A1_2 | | 300 | 2 | | |
| 2 | A1_3 | | 450 | 3 | | |
| 3 | A1_4 | | 600 | 4 | | |
| 4 | A1_5 | | 750 | 5 | | |
| 5 | A1_6 | | 900 | 6 | | |
| 6 | A1_7 | Number of clusters | 1050 | 7 | 150 | [50] |
| 7 | A1_8 | | 1200 | 8 | | |
| 8 | A1_9 | | 1350 | 9 | | |
| 9 | A1_10 | | 1500 | 10 | | |
| 10 | A1_15 | | 2250 | 15 | | |
| 11 | A1 | | 3000 | 20 | | |
| 12 | A2 | | 5250 | 35 | | |
| 13 | A3 | | 7500 | 50 | | |
| 14 | S1 | | | | | |
| 15 | S2 | | | | | |
| 16 | S3 | Overlap | 5000 | 15 | 333 | [51] |
| 17 | S4 | | | | | |
| 18 | G2_2_30 | | | | | |
| 19 | G2_2_50 | Dimensions and Overlap | 2048 | 2 | 1024 | [52] |
| 20 | G2_2_70 | | | | | |
| 21 | Dim32 | Dimension | 1024 | 16 | 64 | [53] |
| 22 | Birch1 | Structure | 100,000 | 100 | 100 | [54] |
| 23 | Unbalance | Unbalance | 6500 | 8 | 100-2000 | [55] |
| 24 | Cervical Cancer(Risk factors) | | 72 | 2 | 21,51 | |
| 25 | WDBC | | 569 | 2 | 212,357 | |
| 26 | Parkinson | | 196 | 2 | 48,147 | |
| 27 | Cervical Cancer (Behavior Risk) | | 72 | 2 | 21,51 | |
| 28 | Iris | Unknown/Real World datasets | 150 | 3 | 50,50,50 | [56] |
| 29 | Seeds | | 210 | 3 | 70,70,70 | |
| 30 | User Knowledge | | 255 | 4 | 24,63,88,83 | |
| 31 | HCV | | 589 | 5 | 526,7,20,12,24 | |
| 32 | Segmentation | | 2310 | 7 | 330,330,330,330,330,330,330 | |
| 33 | e-coli | | 336 | 7 | 143,77,2,2,35,20,5,52 | |
| 34 | Haberman | | 306 | 2 | 225, 81 | |
| 35 | Glass | | 214 | 6 | 70, 76, 17, 13, 9, 29 | |

the scores obtained by GIIS. The description for each validity index was shown as follows:

a) FMI

The FMI is based on the combination of precision and recall shown in Equation (8) and (9):

$$recall, R(b, c) = n_{bc}/n_b \tag{8}$$

$$precision, P(b, c) = n_{bc}/n_c \tag{9}$$

where $b$ is a class of the $q$ classes, $c$ is a cluster of the $k$ clusters, $n$ is the number of objects, $n_bc$ is the number of objects in class $b$, $n_c$ is the number of objects in cluster $c$, and $n_bc$ is number of objects within class **b** in cluster $c$. The FMI of cluster $c$ and class $b$ is then given in Equation (10):

$$FMI(b, c) = \frac{(2R(b, c))\ .\ P(b, c)}{R((b, c) + P(b, c)} \tag{10}$$

b) SIL

The Silhouette index is the distance between intra-cluster objects (compactness) and inter-cluster objects (separation). The Silhouette score, i.e., SIL, is calculated based on the Equation (11):

$$SIL = \frac{b_i - a_i}{max(b_i - a_i)} \tag{11}$$

where $a_i$ is the average distance of an object to other objects in the same cluster and $b_i$ is calculated by finding the minimum average distance from the object to other objects in other clusters.

c) GIIS

GIIS index is combination of SIL (intra and inter clusters distance) and fuzzy partition matric $\mu$. The intra-matrices for each cluster $\mu_i$ is calculated in Equation (12) and (13):

$$IntraDist_i = [intra_i(d_{jg})]|i = 1, \ldots, c, \tag{12}$$

$$intra_i d_{jg} = \mu_{ij}\ AND\ \mu_{ig} \tag{13}$$

The inter-matrices for each cluster $\mu_r$ and $\mu_s$ are calculated in Equation (15) and (16):

$$InterDist_{rs} = [inter_{rs}(d_{jg})]|r, s = 1, \ldots, c; r < s, \tag{14}$$

$$inter_{rs}(d_{jg}) = (\mu_{rj}\ AND\ \mu_{sg})\ OR\ (\mu_{sj}\ AND\ \mu_{rg}) \tag{15}$$

where GIIS index score can be calculated and defined as $a_j$ and $b_j$ as shown in Equation (17) and (18):

$$a_j = \min\left\{\frac{\sum_{k=1}^{n} \text{intra}_i(j, g) \cdot d(j, g)}{\sum_{k=1}^{n} \text{intra}_i(j, g)}\right\} \tag{16}$$

$$b_j = \min\left\{\frac{\sum_{k=1}^{n} \text{intra}_{rs}(j, g) \cdot d(j, g)}{\sum_{k=1}^{n} \text{intra}_{rs}(j, g)}\right\} \tag{17}$$

## VI. RESULTS AND DISCUSSION

This section demonstrates that the FVI is comparable to the external validity index, i.e., FMI, in the current context, which serves as the ground truth for benchmarking and measuring cluster quality. Consequently, the scores obtained by the FVI should also be higher than those obtained by the SIL, which serves as the baseline benchmark. Another benchmark is that the scores obtained by the FVI should also be better than those obtained by the GIIS. These scenarios provide evidence that the FVI could be used independently without relying on the external validity index to measure the cluster quality specifically for FCM. In the following sub-sections, the results obtained by the FVI based on the six clustering properties proposed by Franti and Sieranoja [49] and additional case studies using real world datasets are discussed. Line graphs are presented to facilitate the comparison of FVI with the other validity indices mentioned earlier, with the Y-Axis representing the scores obtained by all indices ranging from 0 to 1 and the X-Axis representing the datasets used in the case studies based on the aforementioned categories. The results are discussed based on the weighting exponent values of $m = 1.1$, 1.5, 2.0, and 2.5 only, and the results for the rest of the weighting exponents can be found in Appendix.

### A. NUMBER OF CLUSTERS

Figure 4 presents the experimental results obtained for the "number of clusters" property from a minimum of two clusters to a maximum of 50 clusters using 13 datasets, namely A1_2, A1_3, A1_4, A1_5, A1_6, A1_7, A1_8, A1_9, A1_10, A1_15, A1, A2 and A3. Overall, the scores obtained by the FVI (in red color) were markedly higher than the scores obtained by the baseline benchmark, which was the SIL index (in green color) for each $m = 1.1$, 1.5, 2.0, and 2.5. Furthermore, the scores obtained by FVI were slightly lower than the scores obtained by FMI (in blue color). These scenarios indicated that the FVI was effective when the scores obtained by FVI were farther from the scores obtained by the baseline benchmark (SIL) and were closer to the ground truth benchmark (FMI). It could be concluded that the FVI is comparable to the FMI for various numbers of clusters. In fact, for each $m$, the FVI gradually approached the patterns of FMI, beginning from $m = 1.5$ (Figure 4(b)) and 2.0 (Figure 4(c)). Finally, for $m = 2.5$, the pattern was almost identical to the FMI (Figure 4(d)). The GIIS (in yellow color) was found to be irrelevant for comparison because its scores were much lower than those obtained by the FVI.

### B. CLUSTER OVERLAP

Figure 5 presents the experimental results for the "cluster overlap" property, where four datasets (S1, S2, S3, and S4) were used to represent various levels of overlap ranging from 9% to 44%. It was found that the last dataset (S4) had a strong overlap. Overall, the scores obtained by FVI (in red color) were higher than the scores obtained by the SIL scores (in green) for each $m = 1.1$, 1.5, 2.0, and 2.5.

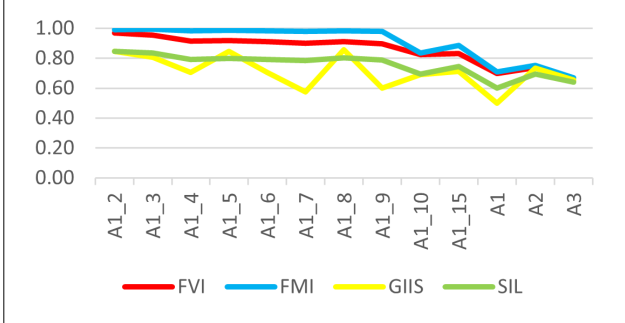


Figure 4(a)

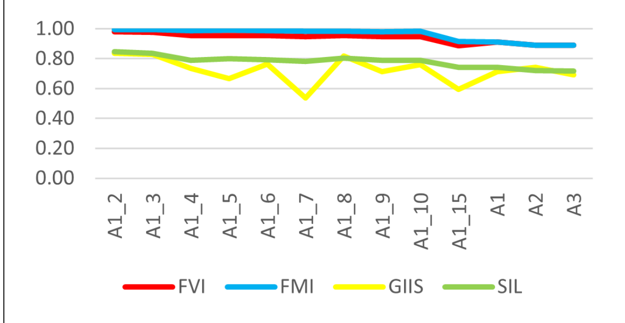


Figure 4(b)

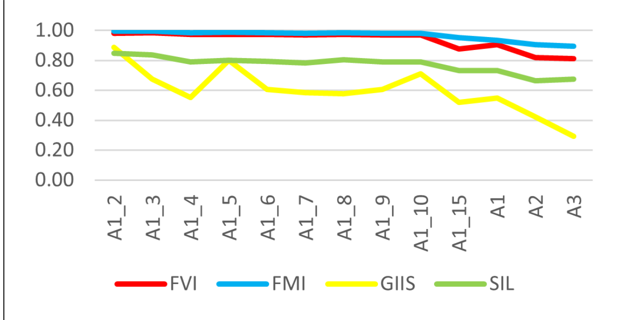


Figure 4(c)

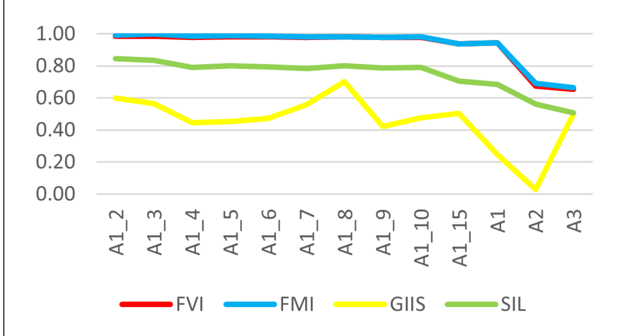


Figure 4(d)

**FIGURE 4. FVI scores by comparison with FMI, GIIS and SIL scores on "number of cluster" properties for a weighting exponent of: (a) $m = 1.1$, (b) $m = 1.5$, (c) $m = 2.0$ and (d) $m = 2.5$.**

Furthermore, the scores obtained by FVI were also slightly lower than those obtained by the FMI scores (in blue). This indicated that the FVI scores were significantly better than

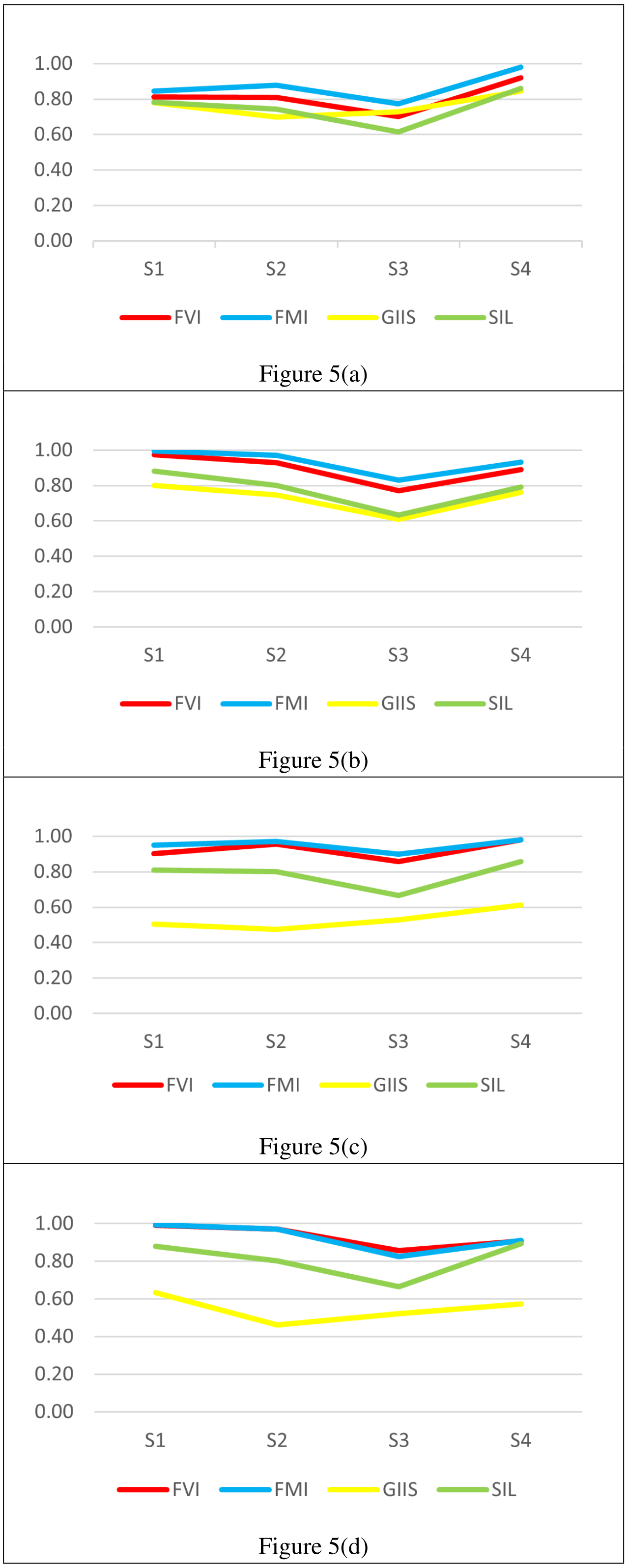


**FIGURE 5.** FVI scores by comparison with FMI, GIIS and SIL scores on "cluster overlap" properties for a weighting exponent of: (a) $m = 1.1$, (b) $m = 1.5$, (c) $m = 2.0$ and (d) $m = 2.5$.

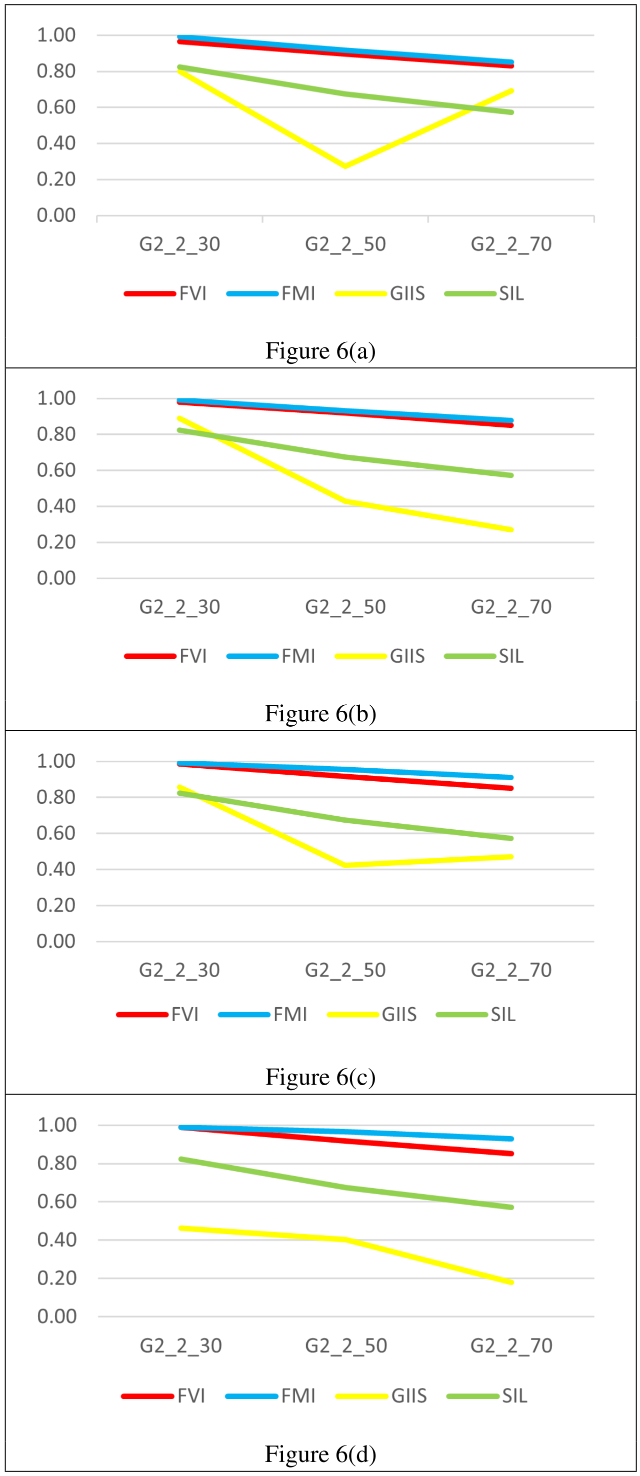


**FIGURE 6.** FVI scores by comparison with FMI, GIIS and SIL scores on "dimensionality" and "cluster overlap" properties for a weighting exponent of: (a) $m = 1.1$, (b) $m = 1.5$, (c) $m = 2.0$ and (d) $m = 2.5$.

the SIL scores and were close to the FMI scores. It could be concluded that the FVI scores were almost similar to the scores obtained by the FMI for the "cluster overlap" properties. For $m = 2.0$, the scores obtained by FVI were almost similar to the FMI scores (Figure 5(c)). However, for $m = 2.5$, the scores dropped slightly for Datasets S3 and S4, indicating that more cluster overlaps might slightly affect the FVI scores. The GIIS scores (in yellow) were found to

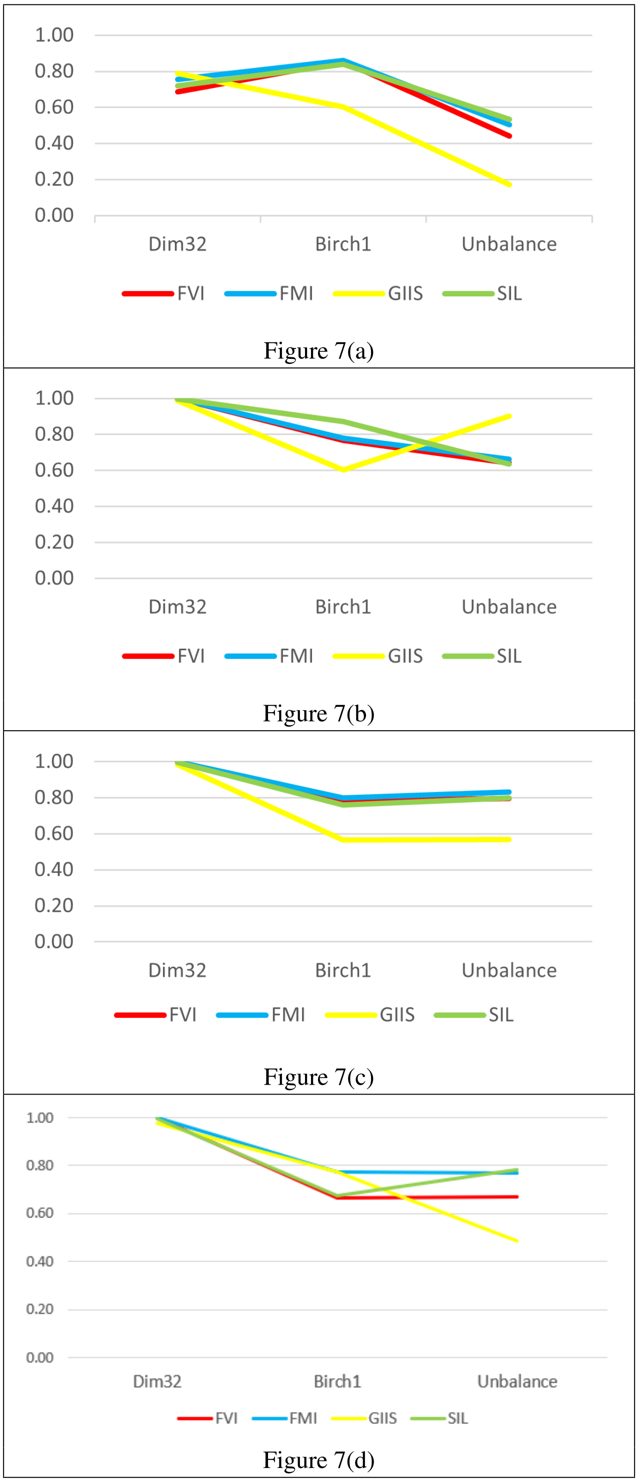


**FIGURE 7.** FVI scores by comparison with FMI, GIIS and SIL scores on multidimensional, structure, and unbalanced data properties for a weighting exponent of: (a) $m = 1.1$, (b) $m = 1.5$, (c) $m = 2.0$ and (d) $m = 2.5$.

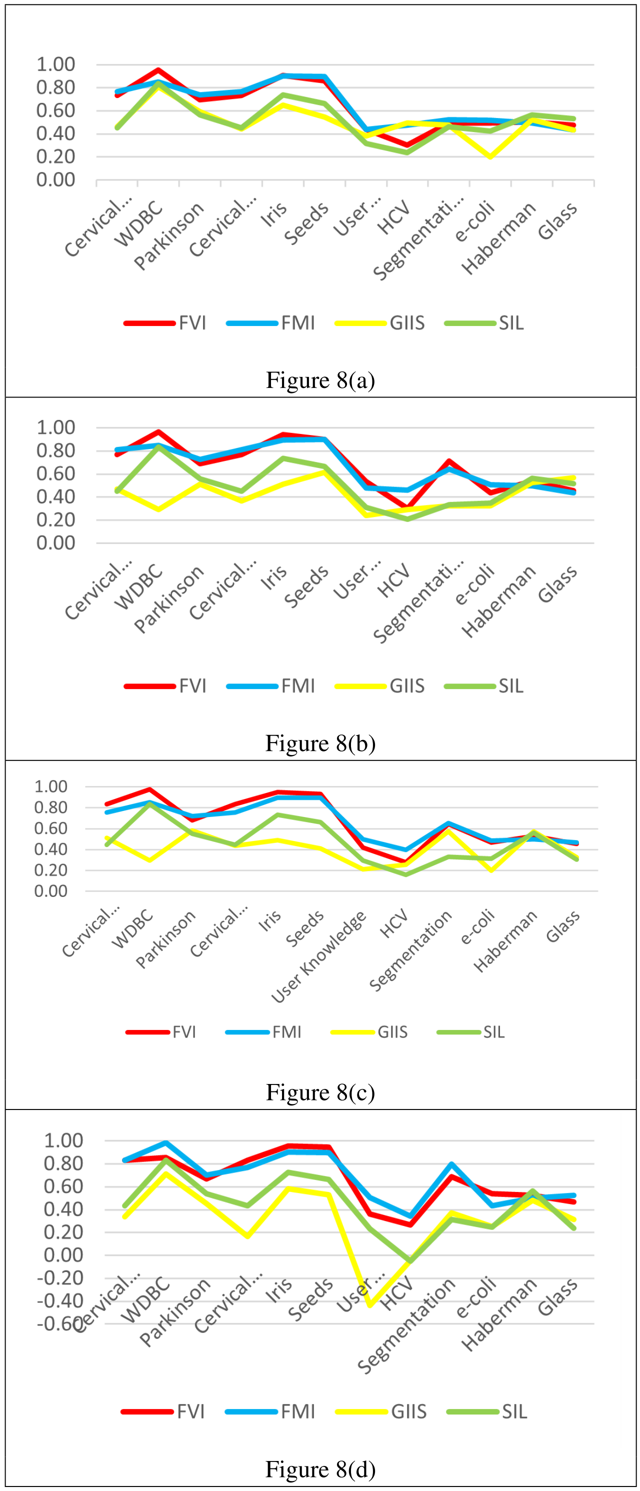


**FIGURE 8.** FVI scores by comparison with FMI, GIIS and SIL scores on real-world data properties for a weighting exponent of: (a) $m = 1.1$, (b) $m = 1.5$, (c) $m = 2.0$ and (d) $m = 2.5$.

be irrelevant for comparison because their scores were much lower than those obtained by the FVI and SIL.

### C. DIMENSIONALITY AND CLUSTER OVERLAP

Figure 6 shows the experimental results for "dimensionality" and "cluster overlap" properties. This case study was based on datasets G2_2_30, G2_2_50, and G2_2_70, which combined "cluster overlap" and "high dimensionality" (1024 dimensions). Overall, the scores obtained by FVI (in red) were discernibly higher than the scores obtained by the baseline benchmark (SIL). It also demonstrated that the

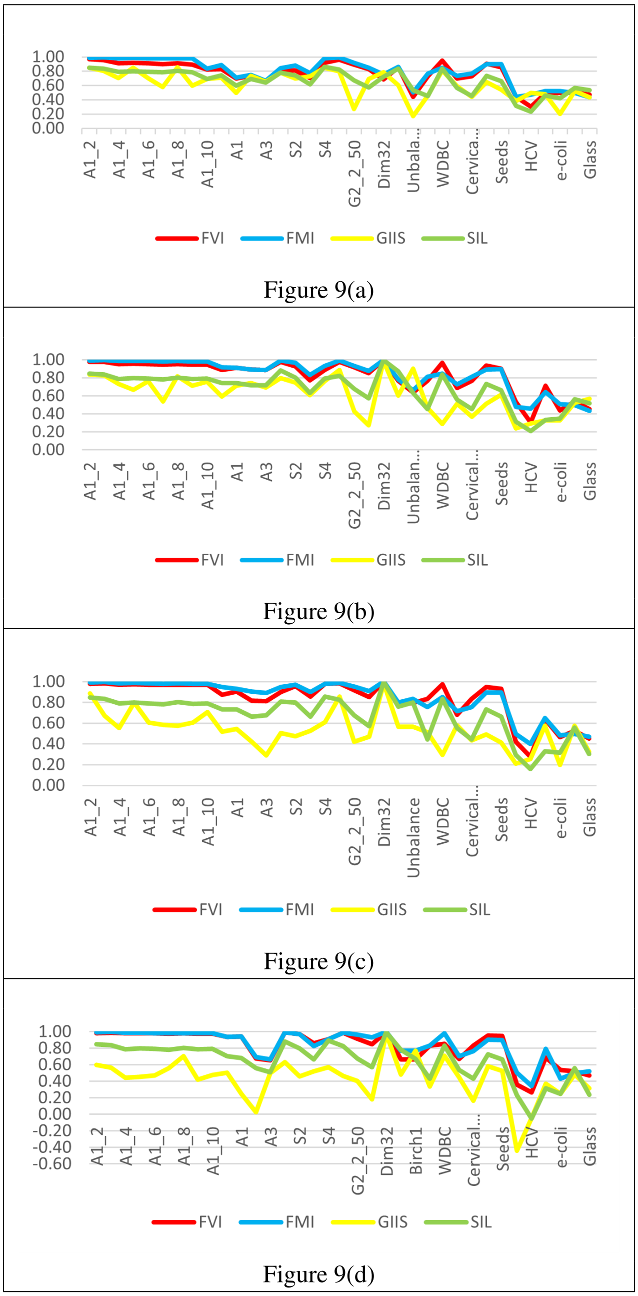


**FIGURE 9.** FVI scores by comparison with FMI, GIIS and SIL scores on all datasets for a weighting exponent of: (a) $m = 1.1$, (b) $m = 1.5$, (c) $m = 2.0$ and (d) $m = 2.5$.

scores obtained by FVI were also marginally lower than the scores obtained by FMI. However, the pattern of FVI scores exhibited a slight drop for $m = 2.0$ and 2.5, indicating that the FVI maintained similar performances as the previous case studies. In contrast, the scores obtained by GIIS (in yellow) were far below compared with those obtained from FVI and SIL. Thus, GIIS was irrelevant for comparison.

### D. MULTIDIMENSIONAL, STRUCTURE AND UNBALANCE DATA

Figure 7 shows the experimental results obtained for three datasets: DIM32, unbalanced, and Birch1, which represent multidimensional property, unbalanced property, and cluster structure property, respectively. Overall, the scores obtained by FVI were marginally lower than the scores obtained by both benchmarks, namely, SIL and FMI, particularly for $m = 1.1$ (Figure 7(a)) and $m = 2.5$ (Figure 7(d)). However, for $m = 2.0$, the scores obtained by FVI seemed to be almost similar to the scores obtained by both benchmarks (Figure 7(c)). From another perspective, the scores obtained by the GIIS were extremely inconsistent, especially for $m = 1.5$.

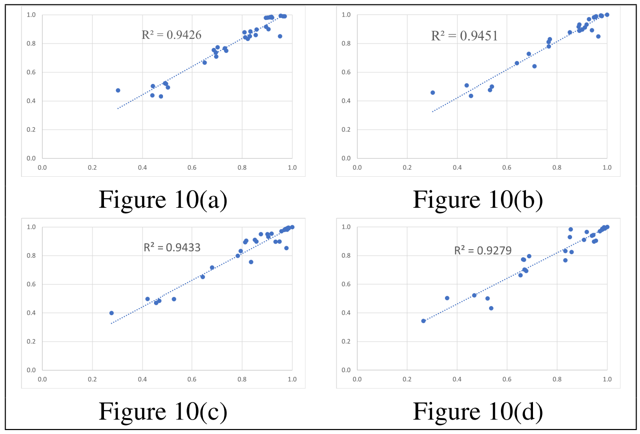


**FIGURE 10.** The correlation coefficient of FVI and FMI for an exponent of: (a) $m$=1.1, (b) $m$=1.5, (c) $m$=2.0 and (d) $m$=2.5.

### E. REAL-WORLD DATA

Figure 8 shows the experimental results obtained for 12 real-world datasets. Unlike synthetic datasets, the cluster properties such as cluster shape, overlap, and deviation of these datasets are unknown. However, the results are still important as they report the performances of FVI in handling such datasets, particularly for future datasets. Overall, FVI performed well in handling the unknown cluster properties represented by these datasets for each value of $m$ (as shown in Figures 8(a) and 8(b)). This is in contrast to the results obtained from synthetic data, where the FVI scores were consistently slightly lower than the FMI scores. Figure 8 shows that the scores obtained by FVI were sometimes higher than those obtained by FMI. However, the difference was not significant, and it was, in fact, marginal. The baseline benchmark represented by the SIL consistently produced lower scores than those obtained by FVI and FMI. Another baseline benchmark represented by the GIIS showed inconsistent patterns and lower scores than those recorded by the SIL and FVI.

### F. OVERALL PERFORMANCES

This section aimed to conclude overall performances of FVI throughout all datasets. Figure 9 shows the line graphs based on all 33 datasets combined. At first glance, it was evident that the scores obtained by FVI and FMI were almost similar throughout all datasets. Therefore, the goal set for the FVI was to achieve scores that were at least closer to

**TABLE 7.** FVI, FMI, SIL, and GIIS scores for $m = 1.1$ to $m = 2.5$.

| Dataset | VI | 1.1 | 1.2 | 1.3 | 1.4 | 1.5 | 1.6 | 1.7 | 1.8 | 1.9 | 2.0 | 2.1 | 2.2 | 2.3 | 2.4 | 2.5 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | FMI | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| | SIL | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 |
| | GIIS | 0.85 | 0.84 | 0.84 | 0.85 | 0.84 | 0.83 | 0.82 | 0.80 | 0.75 | 0.89 | 0.79 | 0.40 | 0.65 | 0.65 | 0.60 |
| | FVI | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| 2 | FMI | 0.99 | 0.99 | 0.99 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | SIL | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 |
| | GIIS | 0.81 | 0.83 | 0.85 | 0.62 | 0.83 | 0.60 | 0.56 | 0.74 | 0.78 | 0.67 | 0.66 | 0.62 | 0.69 | 0.73 | 0.57 |
| | FVI | 0.95 | 0.96 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| 3 | FMI | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| | SIL | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 |
| | GIIS | 0.71 | 0.73 | 0.77 | 0.78 | 0.73 | 0.79 | 0.45 | 0.74 | 0.66 | 0.55 | 0.58 | 0.59 | 0.40 | 0.48 | 0.44 |
| | FVI | 0.91 | 0.93 | 0.94 | 0.95 | 0.95 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| 4 | FMI | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| | SIL | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 |
| | GIIS | 0.85 | 0.82 | 0.82 | 0.82 | 0.67 | 0.71 | 0.75 | 0.56 | 0.64 | 0.80 | 0.64 | 0.55 | 0.75 | 0.60 | 0.45 |
| | FVI | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| 5 | FMI | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.98 |
| | SIL | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 |
| | GIIS | 0.71 | 0.74 | 0.74 | 0.75 | 0.77 | 0.77 | 0.72 | 0.64 | 0.60 | 0.60 | 0.55 | 0.71 | 0.64 | 0.47 | 0.47 |
| | FVI | 0.91 | 0.93 | 0.94 | 0.95 | 0.95 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| 6 | FMI | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| | SIL | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 | 0.78 |
| | GIIS | 0.58 | 0.71 | 0.80 | 0.78 | 0.54 | 0.68 | 0.61 | 0.67 | 0.70 | 0.59 | 0.44 | 0.65 | 0.51 | 0.47 | 0.56 |
| | FVI | 0.90 | 0.92 | 0.93 | 0.94 | 0.95 | 0.96 | 0.96 | 0.96 | 0.97 | 0.97 | 0.98 | 0.97 | 0.98 | 0.98 | 0.98 |
| 7 | FMI | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.98 | 0.98 | 0.94 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| | SIL | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 |
| | GIIS | 0.63 | 0.74 | 0.77 | 0.77 | 0.76 | 0.80 | 0.71 | 0.57 | 0.67 | 0.48 | 0.49 | 0.53 | 0.60 | 0.41 | 0.48 |
| | FVI | 0.89 | 0.91 | 0.92 | 0.93 | 0.94 | 0.95 | 0.95 | 0.96 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 |
| 8 | FMI | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.92 | 0.93 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 |
| | SIL | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.71 | 0.71 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 | 0.79 |
| | GIIS | 0.60 | 0.72 | 0.73 | 0.77 | 0.71 | 0.61 | 0.58 | 0.65 | 0.63 | 0.61 | 0.62 | 0.71 | 0.47 | 0.72 | 0.42 |
| | FVI | 0.89 | 0.92 | 0.93 | 0.94 | 0.95 | 0.82 | 0.82 | 0.96 | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 |
| 9 | FMI | 0.83 | 0.98 | 0.98 | 0.98 | 0.98 | 0.89 | 0.98 | 0.98 | 0.98 | 0.98 | 0.94 | 0.98 | 0.98 | 0.98 | 0.98 |
| | SIL | 0.69 | 0.79 | 0.79 | 0.79 | 0.79 | 0.68 | 0.79 | 0.79 | 0.79 | 0.79 | 0.71 | 0.79 | 0.79 | 0.79 | 0.79 |
| | GIIS | 0.69 | 0.73 | 0.70 | 0.77 | 0.76 | 0.63 | 0.65 | 0.70 | 0.79 | 0.71 | 0.65 | 0.58 | 0.48 | 0.47 | 0.48 |
| | FVI | 0.82 | 0.91 | 0.93 | 0.94 | 0.95 | 0.83 | 0.96 | 0.96 | 0.97 | 0.97 | 0.94 | 0.97 | 0.98 | 0.98 | 0.98 |
| 10 | FMI | 0.88 | 0.88 | 0.89 | 0.98 | 0.92 | 0.98 | 0.98 | 0.98 | 0.98 | 0.95 | 0.85 | 0.95 | 0.98 | 0.96 | 0.94 |
| | SIL | 0.74 | 0.88 | 0.74 | 0.79 | 0.74 | 0.78 | 0.78 | 0.78 | 0.78 | 0.73 | 0.73 | 0.74 | 0.78 | 0.73 | 0.71 |
| | GIIS | 0.71 | 0.88 | 0.79 | 0.65 | 0.59 | 0.77 | 0.67 | 0.62 | 0.74 | 0.52 | 0.57 | 0.58 | 0.61 | 0.29 | 0.51 |
| | FVI | 0.83 | 0.88 | 0.89 | 0.94 | 0.89 | 0.95 | 0.96 | 0.98 | 0.96 | 0.87 | 0.88 | 0.88 | 0.97 | 0.88 | 0.94 |
| 11 | FMI | 0.71 | 0.89 | 0.91 | 0.98 | 0.91 | 0.91 | 0.94 | 0.92 | 0.95 | 0.93 | 0.96 | 0.98 | 0.95 | 0.94 | 0.94 |
| | SIL | 0.60 | 0.74 | 0.74 | 0.78 | 0.74 | 0.71 | 0.75 | 0.71 | 0.75 | 0.73 | 0.75 | 0.79 | 0.73 | 0.70 | 0.68 |
| | GIIS | 0.50 | 0.71 | 0.77 | 0.73 | 0.71 | 0.67 | 0.51 | 0.58 | 0.59 | 0.55 | 0.62 | 0.56 | 0.46 | 0.21 | 0.25 |
| | FVI | 0.70 | 0.86 | 0.88 | 0.93 | 0.91 | 0.83 | 0.91 | 0.91 | 0.91 | 0.90 | 0.91 | 0.97 | 0.90 | 0.83 | 0.94 |
| 12 | FMI | 0.75 | 0.83 | 0.87 | 0.91 | 0.89 | 0.91 | 0.91 | 0.93 | 0.93 | 0.90 | 0.90 | 0.86 | 0.80 | 0.71 | 0.69 |
| | SIL | 0.69 | 0.70 | 0.72 | 0.74 | 0.72 | 0.73 | 0.73 | 0.74 | 0.71 | 0.66 | 0.65 | 0.56 | 0.65 | 0.53 | 0.56 |
| | GIIS | 0.73 | 0.71 | 0.60 | 0.70 | 0.74 | 0.72 | 0.69 | 0.46 | 0.23 | 0.42 | 0.47 | 0.57 | 0.60 | 0.49 | 0.03 |
| | FVI | 0.74 | 0.83 | 0.85 | 0.91 | 0.89 | 0.90 | 0.91 | 0.91 | 0.86 | 0.82 | 0.90 | 0.86 | 0.79 | 0.69 | 0.68 |
| 13 | FMI | 0.67 | 0.78 | 0.82 | 0.95 | 0.89 | 0.89 | 0.91 | 0.91 | 0.90 | 0.89 | 0.88 | 0.81 | 0.75 | 0.72 | 0.66 |
| | SIL | 0.64 | 0.68 | 0.69 | 0.77 | 0.72 | 0.70 | 0.72 | 0.70 | 0.68 | 0.68 | 0.64 | 0.66 | 0.61 | 0.59 | 0.51 |
| | GIIS | 0.66 | 0.61 | 0.55 | 0.78 | 0.69 | 0.53 | 0.50 | 0.46 | 0.50 | 0.29 | 0.38 | 0.59 | 0.65 | 0.39 | 0.50 |
| | FVI | 0.65 | 0.78 | 0.79 | 0.91 | 0.89 | 0.84 | 0.88 | 0.85 | 0.82 | 0.81 | 0.88 | 0.79 | 0.74 | 0.70 | 0.65 |
| 14 | FMI | 0.84 | 0.99 | 0.93 | 0.94 | 0.99 | 0.92 | 0.95 | 0.99 | 0.99 | 0.95 | 0.99 | 0.99 | 0.92 | 0.99 | 0.99 |
| | SIL | 0.78 | 0.88 | 0.82 | 0.75 | 0.88 | 0.84 | 0.83 | 0.88 | 0.88 | 0.81 | 0.88 | 0.88 | 0.81 | 0.88 | 0.88 |
| | GIIS | 0.78 | 0.82 | 0.81 | 0.71 | 0.80 | 0.68 | 0.67 | 0.31 | 0.61 | 0.51 | 0.65 | 0.46 | 0.46 | 0.49 | 0.63 |
| | FVI | 0.81 | 0.95 | 0.90 | 0.90 | 0.97 | 0.90 | 0.90 | 0.98 | 0.99 | 0.90 | 0.99 | 0.99 | 0.90 | 0.99 | 0.99 |

**TABLE 7.** ***(Continued.)***

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 15 | FMI | 0.88 | 0.88 | 0.87 | 0.99 | 0.97 | 0.93 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 |
| | SIL | 0.74 | 0.75 | 0.74 | 0.88 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 |
| | GIIS | 0.70 | 0.71 | 0.59 | 0.81 | 0.75 | 0.82 | 0.79 | 0.76 | 0.57 | 0.48 | 0.47 | 0.70 | 0.47 | 0.43 | 0.46 |
| | FVI | 0.81 | 0.85 | 0.87 | 0.97 | 0.93 | 0.94 | 0.94 | 0.95 | 0.95 | 0.96 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 |
| 16 | FMI | 0.77 | 0.79 | 0.79 | 0.97 | 0.83 | 0.86 | 0.87 | 0.86 | 0.89 | 0.90 | 0.91 | 0.91 | 0.82 | 0.90 | 0.83 |
| | SIL | 0.61 | 0.64 | 0.64 | 0.80 | 0.63 | 0.67 | 0.67 | 0.63 | 0.67 | 0.67 | 0.67 | 0.67 | 0.63 | 0.67 | 0.87 |
| | GIIS | 0.73 | 0.69 | 0.63 | 0.62 | 0.61 | 0.62 | 0.61 | 0.50 | 0.56 | 0.53 | 0.57 | 0.44 | 0.40 | 0.30 | 0.52 |
| | FVI | 0.70 | 0.75 | 0.78 | 0.92 | 0.77 | 0.86 | 0.86 | 0.79 | 0.86 | 0.86 | 0.86 | 0.86 | 0.80 | 0.86 | 0.86 |
| 17 | FMI | 0.98 | 0.98 | 0.93 | 0.80 | 0.93 | 0.85 | 0.98 | 0.96 | 0.73 | 0.98 | 0.97 | 0.97 | 0.84 | 0.97 | 0.91 |
| | SIL | 0.86 | 0.86 | 0.83 | 0.63 | 0.93 | 0.86 | 0.86 | 0.82 | 0.76 | 0.86 | 0.81 | 0.82 | 0.78 | 0.82 | 0.89 |
| | GIIS | 0.85 | 0.85 | 0.89 | 0.53 | 0.93 | 0.82 | 0.77 | 0.59 | 0.59 | 0.61 | 0.52 | 0.59 | 0.63 | 0.47 | 0.57 |
| | FVI | 0.92 | 0.94 | 0.89 | 0.78 | 0.93 | 0.82 | 0.97 | 0.89 | 0.81 | 0.98 | 0.97 | 0.90 | 0.88 | 0.90 | 0.91 |
| 18 | FMI | 0.99 | 0.99 | 0.99 | 0.82 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| | SIL | 0.82 | 0.82 | 0.82 | 0.76 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 |
| | GIIS | 0.80 | 0.84 | 0.88 | 0.78 | 0.89 | 0.57 | 0.56 | 0.43 | 0.70 | 0.86 | 0.23 | 0.48 | 0.64 | 0.78 | 0.46 |
| | FVI | 0.96 | 0.97 | 0.97 | 0.80 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 | 0.99 |
| 19 | FMI | 0.92 | 0.92 | 0.92 | 0.99 | 0.93 | 0.94 | 0.94 | 0.95 | 0.95 | 0.95 | 0.96 | 0.96 | 0.96 | 0.96 | 0.97 |
| | SIL | 0.68 | 0.68 | 0.68 | 0.82 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.68 | 0.96 | 0.68 | 0.68 | 0.68 |
| | GIIS | 0.27 | 0.67 | 0.56 | 0.35 | 0.43 | 0.41 | 0.41 | 0.43 | 0.32 | 0.42 | 0.51 | 0.61 | 0.96 | 0.47 | 0.40 |
| | FVI | 0.90 | 0.91 | 0.92 | 0.98 | 0.92 | 0.92 | 0.92 | 0.92 | 0.92 | 0.92 | 0.96 | 0.92 | 0.96 | 0.92 | 0.92 |
| 20 | FMI | 0.85 | 0.85 | 0.86 | 0.93 | 0.88 | 0.89 | 0.89 | 0.90 | 0.91 | 0.91 | 0.92 | 0.92 | 0.93 | 0.93 | 0.93 |
| | SIL | 0.57 | 0.57 | 0.57 | 0.68 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 | 0.57 |
| | GIIS | 0.69 | 0.55 | 0.40 | 0.39 | 0.27 | 0.53 | 0.71 | 0.57 | 0.48 | 0.47 | 0.44 | 0.49 | 0.51 | 0.23 | 0.18 |
| | FVI | 0.83 | 0.85 | 0.85 | 0.92 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 |
| 21 | FMI | 0.75 | 0.91 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | SIL | 0.72 | 0.90 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| | GIIS | 0.79 | 0.56 | 1.00 | 1.00 | 0.99 | 0.98 | 0.99 | 0.95 | 0.99 | 0.98 | 0.99 | 0.97 | 0.98 | 0.97 | 0.98 |
| | FVI | 0.69 | 0.89 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 | 1.00 |
| 22 | FMI | 0.86 | 0.84 | 0.77 | 0.93 | 0.78 | 0.77 | 0.77 | 0.83 | 0.67 | 0.80 | 0.77 | 0.83 | 0.83 | 0.67 | 0.77 |
| | SIL | 0.84 | 0.57 | 0.80 | 0.80 | 0.87 | 0.80 | 0.80 | 0.81 | 0.63 | 0.76 | 0.79 | 0.81 | 0.80 | 0.62 | 0.78 |
| | GIIS | 0.60 | 0.23 | 0.89 | 0.85 | 0.60 | 0.76 | 0.65 | 0.67 | 0.61 | 0.57 | 0.51 | 0.50 | 0.73 | 0.50 | 0.48 |
| | FVI | 0.85 | 0.81 | 0.76 | 0.92 | 0.77 | 0.66 | 0.66 | 0.79 | 0.63 | 0.78 | 0.64 | 0.79 | 0.79 | 0.65 | 0.67 |
| 23 | FMI | 0.50 | 0.78 | 0.86 | 0.87 | 0.66 | 0.81 | 0.82 | 0.83 | 0.84 | 0.83 | 0.79 | 0.78 | 0.77 | 0.76 | 0.77 |
| | SIL | 0.53 | 0.87 | 0.86 | 0.87 | 0.63 | 0.81 | 0.82 | 0.83 | 0.84 | 0.80 | 0.79 | 0.78 | 0.77 | 0.76 | 0.67 |
| | GIIS | 0.17 | 0.60 | 0.58 | 0.59 | 0.90 | 0.58 | 0.59 | 0.60 | 0.60 | 0.57 | 0.56 | 0.55 | 0.54 | 0.53 | 0.77 |
| | FVI | 0.44 | 0.75 | 0.75 | 0.76 | 0.64 | 0.84 | 0.85 | 0.86 | 0.87 | 0.79 | 0.82 | 0.81 | 0.80 | 0.79 | 0.66 |
| 24 | FMI | 0.77 | 0.77 | 0.77 | 0.81 | 0.81 | 0.81 | 0.77 | 0.77 | 0.77 | 0.76 | 0.78 | 0.78 | 0.77 | 0.77 | 0.83 |
| | SIL | 0.45 | 0.34 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.44 | 0.44 | 0.43 | 0.43 | 0.43 | 0.43 |
| | GIIS | 0.46 | 0.50 | 0.45 | 0.41 | 0.47 | 0.40 | 0.46 | 0.50 | 0.45 | 0.51 | 0.38 | 0.41 | 0.38 | 0.41 | 0.34 |
| | FVI | 0.73 | 0.71 | 0.80 | 0.77 | 0.77 | 0.77 | 0.83 | 0.83 | 0.83 | 0.84 | 0.84 | 0.84 | 0.83 | 0.83 | 0.83 |
| 25 | FMI | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.98 |
| | SIL | 0.84 | 0.85 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.84 | 0.83 | 0.83 | 0.83 | 0.83 | 0.83 | 0.83 |
| | GIIS | 0.81 | 0.85 | 0.86 | 0.80 | 0.29 | 0.79 | 0.76 | 0.72 | 0.68 | 0.30 | 0.73 | 0.88 | 0.89 | 0.85 | 0.71 |
| | FVI | 0.95 | 0.85 | 0.96 | 0.96 | 0.97 | 0.97 | 0.97 | 0.97 | 0.97 | 0.98 | 0.98 | 0.98 | 0.98 | 0.98 | 0.85 |
| 26 | FMI | 0.74 | 0.74 | 0.74 | 0.74 | 0.73 | 0.73 | 0.73 | 0.73 | 0.72 | 0.72 | 0.71 | 0.71 | 0.70 | 0.70 | 0.70 |
| | SIL | 0.57 | 0.57 | 0.57 | 0.57 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.55 | 0.54 | 0.54 | 0.54 | 0.54 | 0.54 |
| | GIIS | 0.59 | 0.66 | 0.42 | 0.55 | 0.51 | 0.59 | 0.57 | 0.59 | 0.57 | 0.58 | 0.38 | 0.42 | 0.25 | 0.56 | 0.45 |
| | FVI | 0.70 | 0.70 | 0.70 | 0.70 | 0.69 | 0.69 | 0.69 | 0.69 | 0.68 | 0.68 | 0.67 | 0.67 | 0.67 | 0.67 | 0.67 |
| 27 | FMI | 0.77 | 0.77 | 0.80 | 0.81 | 0.81 | 0.81 | 0.83 | 0.78 | 0.77 | 0.76 | 0.78 | 0.78 | 0.77 | 0.77 | 0.77 |
| | SIL | 0.45 | 0.34 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.45 | 0.44 | 0.44 | 0.43 | 0.43 | 0.43 | 0.43 |
| | GIIS | 0.44 | 0.46 | 0.32 | 0.36 | 0.37 | 0.43 | 0.49 | 0.49 | 0.44 | 0.44 | 0.35 | 0.27 | 0.34 | 0.40 | 0.16 |
| | FVI | 0.73 | 0.71 | 0.77 | 0.76 | 0.78 | 0.76 | 0.79 | 0.75 | 0.74 | 0.74 | 0.74 | 0.73 | 0.72 | 0.72 | 0.83 |
| 28 | FMI | 0.85 | 0.86 | 0.86 | 0.85 | 0.85 | 0.85 | 0.85 | 0.86 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.88 |
| | SIL | 0.78 | 0.79 | 0.79 | 0.79 | 0.79 | 0.80 | 0.80 | 0.80 | 0.80 | 0.80 | 0.81 | 0.81 | 0.81 | 0.81 | 0.80 |
| | GIIS | 0.48 | 0.43 | 0.36 | 0.47 | 0.57 | 0.51 | 0.49 | 0.47 | 0.45 | 0.45 | 0.40 | 0.36 | 0.34 | 0.28 | 0.46 |
| | FVI | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.87 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.88 | 0.95 |

**TABLE 7.** *(Continued.)*

| | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 29 | FMI | 0.85 | 0.85 | 0.85 | 0.86 | 0.85 | 0.85 | 0.85 | 0.85 | 0.86 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 |
| | SIL | 0.80 | 0.80 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.81 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 | 0.82 |
| | GIIS | 0.46 | 0.46 | 0.45 | 0.44 | 0.44 | 0.43 | 0.48 | 0.46 | 0.42 | 0.42 | 0.38 | 0.39 | 0.31 | 0.42 | 0.21 |
| | FVI | 0.89 | 0.88 | 0.88 | 0.89 | 0.89 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.89 | 0.88 | 0.89 | 0.89 | 0.95 |
| 30 | FMI | 0.92 | 0.93 | 0.92 | 0.91 | 0.91 | 0.91 | 0.91 | 0.91 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.90 | 0.92 |
| | SIL | 0.84 | 0.84 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.85 | 0.86 | 0.86 | 0.86 | 0.86 | 0.86 | 0.86 |
| | GIIS | 0.30 | 0.35 | 0.27 | 0.21 | 0.20 | 0.26 | 0.32 | 0.31 | 0.30 | 0.29 | 0.17 | 0.22 | 0.08 | 0.26 | 0.15 |
| | FVI | 0.95 | 0.95 | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.94 | 0.93 | 0.94 | 0.94 | 0.94 | 0.36 |
| 31 | FMI | 0.47 | 0.47 | 0.64 | 0.66 | 0.46 | 0.45 | 0.44 | 0.43 | 0.40 | 0.40 | 0.40 | 0.26 | 0.26 | 0.46 | 0.34 |
| | SIL | 0.24 | 0.23 | 0.47 | 0.22 | 0.21 | 0.20 | 0.19 | 0.18 | 0.17 | 0.16 | 0.16 | 0.15 | -0.06 | -0.05 | -0.05 |
| | GIIS | 0.49 | 0.14 | 0.30 | 0.27 | 0.29 | 0.31 | 0.62 | 0.35 | 0.45 | 0.25 | 0.60 | 0.16 | 0.51 | 0.01 | -0.05 |
| | FVI | 0.30 | 0.30 | 0.54 | 0.56 | 0.30 | 0.30 | 0.30 | 0.30 | 0.28 | 0.28 | 0.25 | 0.38 | 0.38 | 0.37 | 0.27 |
| 32 | FMI | 0.52 | 0.63 | 0.67 | 0.70 | 0.64 | 0.57 | 0.57 | 0.60 | 0.65 | 0.65 | 0.67 | 0.67 | 0.54 | 0.80 | 0.80 |
| | SIL | 0.46 | 0.42 | 0.39 | 0.39 | 0.33 | 0.39 | 0.39 | 0.43 | 0.34 | 0.33 | 0.33 | 0.32 | 0.30 | 0.31 | 0.31 |
| | GIIS | 0.48 | 0.46 | 0.21 | 0.38 | 0.33 | 0.16 | 0.12 | 0.37 | 0.32 | 0.58 | 0.30 | 0.49 | 0.49 | -0.21 | 0.37 |
| | FVI | 0.49 | 0.54 | 0.60 | 0.58 | 0.71 | 0.57 | 0.57 | 0.60 | 0.67 | 0.64 | 0.67 | 0.66 | 0.66 | 0.53 | 0.69 |
| 33 | FMI | 0.52 | 0.53 | 0.56 | 0.60 | 0.51 | 0.50 | 0.50 | 0.49 | 0.50 | 0.49 | 0.52 | 0.52 | 0.49 | 0.57 | 0.43 |
| | SIL | 0.42 | 0.40 | 0.36 | 0.60 | 0.35 | 0.34 | 0.34 | 0.33 | 0.32 | 0.31 | 0.34 | 0.35 | 0.28 | 0.31 | 0.25 |
| | GIIS | 0.20 | 0.46 | 0.57 | 0.60 | 0.33 | 0.17 | 0.17 | 0.40 | 0.32 | 0.20 | 0.13 | -0.05 | 0.10 | -0.09 | 0.25 |
| | FVI | 0.50 | 0.52 | 0.51 | 0.60 | 0.44 | 0.48 | 0.48 | 0.47 | 0.47 | 0.47 | 0.44 | 0.46 | 0.44 | 0.42 | 0.54 |
| 34 | FMI | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.49 | 0.50 | 0.50 | 0.50 |
| | SIL | 0.57 | 0.52 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 | 0.56 |
| | GIIS | 0.52 | 0.50 | 0.58 | 0.55 | 0.52 | 0.59 | 0.55 | 0.52 | 0.50 | 0.57 | 0.55 | 0.54 | 0.52 | 0.50 | 0.48 |
| | FVI | 0.50 | 0.51 | 0.55 | 0.54 | 0.54 | 0.53 | 0.53 | 0.53 | 0.53 | 0.52 | 0.45 | 0.45 | 0.55 | 0.55 | 0.52 |
| 35 | FMI | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.43 | 0.47 | 0.49 | 0.49 | 0.52 | 0.50 | 0.52 |
| | SIL | 0.53 | 0.58 | 0.57 | 0.57 | 0.51 | 0.51 | 0.50 | 0.50 | 0.50 | 0.30 | 0.26 | 0.26 | 0.26 | 0.28 | 0.23 |
| | GIIS | 0.43 | 0.53 | 0.50 | 0.59 | 0.56 | 0.54 | 0.52 | 0.50 | 0.58 | 0.32 | 0.22 | 0.20 | 0.64 | 0.25 | 0.31 |
| | FVI | 0.47 | 0.45 | 0.50 | 0.46 | 0.45 | 0.45 | 0.40 | 0.39 | 0.39 | 0.45 | 0.45 | 0.44 | 0.46 | 0.49 | 0.46 |

the scores obtained by FMI, which served as the ground truth benchmark. FVI was able to achieve this goal fairly. Furthermore, the scores obtained by FVI were expected to be higher than those obtained by the SIL. In this context, the FVI was able to achieve this goal. In most cases, the FVI demonstrated significantly higher scores for each case. Finally, the FVI scores were definitely better than the scores obtained by the GIIS, which were not comparable to the FVI scores.

For further analysis, the correlation coefficient was used to measure the strength of a linear relationship between the scores obtained by FVI and FMI. The results are shown in Figure 10, where the $r^2$ values were recorded above 0.9 for $m = 1.1, 1.5, 2.0$, and 2.5. These results clearly indicated that the scores obtained by FVI were 90% similar to the scores obtained by FMI. Thus, this provides strong evidence that the new internal validity index, FVI, can be used to measure cluster quality, similar to external validity indices such as FMI.

## VII. CONCLUSION

Overall, the results showed that the FVI produced similar scores compared to the external validity index. Therefore, the following conclusions can be drawn:

a) The FVI could also serve as a ground truth for measuring cluster quality such as FMI. By considering various weighting exponents required by FCM (from a minimum of 1.1 to a maximum of 2.5), the performance of FVI was very promising. The experimental results clearly demonstrated that the FVI could adequately handle the exponents from 1.1 to 2.0, and there was only a marginal difference when the exponents were set to be greater than 2.0 toward 2.5.
b) The FVI was comparable to the FMI, which was considered a ground-truth external validity index for benchmarking cluster quality. Based on the statistical analysis, it was proven that the correlation coefficient between the FVI and FMI was mostly greater than 0.9.
c) The FVI could also serve as a ground truth from an internal validation perspective. One advantage of the FVI is its ability to validate clusters with unknown labels, which aligns with real-world data and applications.

## VIII. FUTURE WORK

Furthermore, the FVI could be further investigated and validated in two areas.

a) First, it could be examined if the scores obtained by the FVI can come closer or be identical to the scores obtained by any external validity indices, especially when the weighting exponent $m$ is chosen by users.

**TABLE 8. Correlation Coefficient Between FVI and FMI Reported for Each Cluster Category.**

| $m$ | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| 1.1 | 0.96 | 0.96 | 1.00 | 0.99 | 0.91 | 0.94 |
| 1.2 | 0.93 | 1.00 | 1.00 | 1.00 | 0.90 | 0.94 |
| 1.3 | 0.93 | 0.99 | 0.99 | 0.80 | 0.88 | 0.93 |
| 1.4 | 0.71 | 0.98 | 1.00 | 0.94 | 0.90 | 0.93 |
| 1.5 | 0.92 | 0.99 | 1.00 | 1.00 | 0.86 | 0.95 |
| 1.6 | 0.85 | 0.93 | 0.99 | 0.85 | 0.89 | 0.93 |
| 1.7 | 0.80 | 0.89 | 1.00 | 0.86 | 0.93 | 0.93 |
| 1.8 | 0.92 | 0.95 | 1.00 | 0.88 | 0.95 | 0.95 |
| 1.9 | 0.99 | 0.90 | 1.00 | 0.98 | 0.97 | 0.94 |
| 2.0 | 0.95 | 0.92 | 1.00 | 0.99 | 0.94 | 0.94 |
| 2.1 | 0.90 | 0.98 | 0.94 | 0.80 | 0.95 | 0.93 |
| 2.2 | 0.92 | 0.82 | 0.99 | 0.90 | 0.90 | 0.93 |
| 2.3 | 0.99 | 0.90 | 0.99 | 0.90 | 0.91 | 0.93 |
| 2.4 | 0.92 | 0.90 | 0.99 | 0.98 | 0.96 | 0.91 |
| 2.5 | 1.00 | 0.98 | 0.99 | 1.00 | 0.87 | 0.93 |

* A-Number of Clusters
* B-Cluster Overlap
* C-Dimensionality and Overlap
* D-Multidimensionality, Structure and Unbalance
* E- Real-World Dataset
* F- Overall Dataset

b) Second, the underlying theory behind the adjusted fuzzy matrix should also be further explored as it plays a significant role in developing the FVI.

## APPENDIX

## FVI, FMI, SIL & GIIS SCORES FOR M = 1.1 TO M = 2.5

See Table 7.

## CORRELATION COEFFICIENT BETWEEN FVI AND FMI REPORTED FOR EACH CLUSTERING PROPERTY

See Table 8.

## ACKNOWLEDGMENT

The authors extend their gratitude to those who have contributed toward the completion of this study.

## REFERENCES

[1] G. Gan, C. Ma, and J. Wu, *Data Clustering: Theory, Algorithms, and Applications*. Philadelphia, PA, USA: SIAM, 2020.
[2] J. C. Bezdek, *Pattern Recognition With Fuzzy Objective Function Algorithms*. Springer, 2013.
[3] J. C. Dunn, "A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters," *J. Cybern.*, vol. 3, no. 3, pp. 32–57, 1973.
[4] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proc. 5th Berkeley Symp. Math. Statist. Probab.*, vol. 1, 1967, pp. 281–297.
[5] J. Kogan, *Introduction to Clustering Large and High-Dimensional Data*. Cambridge, U.K.: Cambridge Univ. Press, 2007.
[6] C. C. Aggarwal, "An introduction to cluster analysis," in *Data Clustering*. Boca Raton, FL, USA: CRC Press, 2018, pp. 1–28.
[7] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *J. Amer. Stat. Assoc.*, vol. 66, no. 336, p. 846, Dec. 1971.
[8] E. B. Fowlkes and C. L. Mallows, "A method for comparing two hierarchical clusterings," *J. Amer. Stat. Assoc.*, vol. 78, no. 383, p. 553, Sep. 1983.
[9] B. Larsen and C. Aone, "Fast and effective text mining using linear-time document clustering," in *Proc. 5th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 1999, pp. 16–22.
[10] S. Günnemann, I. Färber, E. Müller, I. Assent, and T. Seidl, "External evaluation measures for subspace clustering," in *Proc. 20th ACM Int. Conf. Inf. Knowl. Manage.*, Oct. 2011, pp. 1363–1372.
[11] B. E. Dom, "An information-theoretic external cluster-validity measure," 2012, *arXiv:1301.0565*.
[12] Y. Lei, J. C. Bezdek, S. Romano, N. X. Vinh, J. Chan, and J. Bailey, "Ground truth bias in external cluster validity indices," *Pattern Recognit.*, vol. 65, pp. 58–70, May 2017.
[13] P. J. Rousseeuw, "Silhouettes: A graphical aid to the interpretation and validation of cluster analysis," *J. Comput. Appl. Math.*, vol. 20, pp. 53–65, Nov. 1987.
[14] Y. Fukuyama, "A new method of choosing the number of clusters for fuzzy C-means method," in *Proc. 5th Fuzzy Syst. Symp.*, 1989, pp. 247–250.
[15] S. H. Kwon, "Cluster validity index for fuzzy clustering," *Electron. Lett.*, vol. 34, no. 22, pp. 2176–2177, 1998.
[16] X. L. Xie and G. Beni, "A validity measure for fuzzy clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 8, pp. 841–847, Aug. 1991.
[17] J. C. Bezdek and S. K. Pal, *Fuzzy Models for Pattern Recognition: Methods That Search for Structures in Data*. IEEE Press, 1992.
[18] T. Calinski and J. Harabasz, "A dendrite method for cluster analysis," *Commun. Statist. Simul. Comput.*, vol. 3, no. 1, pp. 1–27, 1974.
[19] A. Strehl and J. Ghosh, "Cluster ensembles—A knowledge reuse framework for combining multiple partitions," *J. Mach. Learn. Res.*, vol. 3, pp. 583–617, Jun. 2002.
[20] R. J. G. B. Campello and E. R. Hruschka, "A fuzzy extension of the silhouette width criterion for cluster analysis," *Fuzzy Sets Syst.*, vol. 157, no. 21, pp. 2858–2875, Nov. 2006.
[21] M. Rawashdeh and A. Ralescu, "Crisp and fuzzy cluster validity: Generalized intra-inter silhouette index," in *Proc. Annu. Meeting North Amer. Fuzzy Inf. Process. Soc. (NAFIPS)*, Aug. 2012, pp. 1–6.
[22] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-1, no. 2, pp. 224–227, Apr. 1979.
[23] L. Hubert and P. Arabie, "Comparing partitions," *J. Classification*, vol. 2, no. 1, pp. 193–218, Dec. 1985.
[24] W. J. Krzanowski and Y. T. Lai, "A criterion for determining the number of groups in a data set using sum-of-squares clustering," *Biometrics*, vol. 44, no. 1, pp. 23–34, Mar. 1988.
[25] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, Sep. 1993.
[26] J. C. Bezdek and N. R. Pal, "Some new indexes of cluster validity," *IEEE Trans. Syst., Man Cybern. B, Cybern.*, vol. 28, no. 3, pp. 301–315, Jun. 1998.
[27] B. Rezaee, "A cluster validity index for fuzzy clustering," *Fuzzy Sets Syst.*, vol. 161, no. 23, pp. 3014–3025, Dec. 2010.
[28] H.-L. Shieh and P.-L. Chang, "A new robust validity index for fuzzy clustering algorithm," in *Proc. IEEE Int. Conf. Ind. Eng. Eng. Manage.*, Dec. 2010, pp. 767–771.
[29] E. Pomponi and A. Vinogradov, "A real-time approach to acoustic emission clustering," *Mech. Syst. Signal Process.*, vol. 40, no. 2, pp. 791–804, Nov. 2013.
[30] M. A. Rahman and M. Z. Islam, "CRUDAW: A novel fuzzy technique for clustering records following user defined attribute weights," in *Proc. 10th Australas. Data Mining Conf.*, 2012, pp. 27–42.
[31] S. Banerjee, Y. Badr, and E. T. Al-Shammari, "Analyzing tweet cluster using standard fuzzy C means clustering," in *Social Networks: A Framework of Computational Intelligence*. Berlin, Germany: Springer, 2014, pp. 377–406.
[32] P. Civicioglu, U. H. Atasever, C. Ozkan, E. Besdok, A. E. Karkinli, and A. Kesikoglu, "Performance comparison of evolutionary algorithms for image clustering," *Int. Arch. Photogramm., Remote Sens. Spatial Inf. Sci.*, vol. XL-7, pp. 71–74, Sep. 2014.
[33] S. Banerjee, A. Choudhary, and S. Pal, "Empirical evaluation of K-means, bisecting K-means, fuzzy C-means and genetic K-means clustering algorithms," in *Proc. IEEE Int. WIE Conf. Electr. Comput. Eng. (WIECON-ECE)*, Dec. 2015, pp. 168–172.

[34] E. A. Fernandez, J. J. G. Torres, A. M. C. Soto, and N. G. González, "Demodulation of m-ary non-symmetrical constellations using clustering techniques in optical communication systems," in *Proc. IEEE Latin Amer. Conf. Comput. Intell. (LA-CCI)*, Nov. 2016, pp. 1–6.

[35] N. H. Bakar, Z. M. Kasirun, N. Salleh, and H. A. Jalab, "Extracting features from online software reviews to aid requirements reuse," *Appl. Soft Comput.*, vol. 49, pp. 1297–1315, Dec. 2016.

[36] D. Nurjanah, K. Dewanto, and F. D. Sari, "Homogeneous group formation in collaborative learning using fuzzy C-means," in *Proc. IEEE 6th Int. Conf. Teaching, Assessment, Learn. Eng. (TALE)*, Dec. 2017, pp. 74–79.

[37] S. Renjith, A. Sreekumar, and M. Jathavedan, "Evaluation of partitioning clustering algorithms for processing social media data in tourism domain," in *Proc. IEEE Recent Adv. Intell. Comput. Syst. (RAICS)*, Dec. 2018, pp. 127–131.

[38] K. V. Rajkumar, A. Yesubabu, and K. Subrahmanyam, "Fuzzy clustering and fuzzy $c$-means partition cluster analysis and validation studies on a subset of citescore dataset," *Int. J. Electr. Comput. Eng. (IJECE)*, vol. 9, no. 4, p. 2760, Aug. 2019.

[39] P. Mishra and N. Bhoi, "Kalman filtering based fuzzy C-means clustering and artificial neural network for classification of microarray data," in *Proc. Int. Conf. Robot. Ind. Appl.*, Jan. 2019, pp. 305–315.

[40] J. Eliyanto and S. Surono, "Distance functions study in fuzzy C-means core and reduct clustering," *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika*, vol. 7, no. 1, p. 118, Apr. 2021.

[41] E. M. S. Rochman and H. Suprajitno, "Comparison of clustering in tuberculosis using fuzzy $c$-means and $k$-means methods," *Commun. Math. Biol. Neurosci.*, vol. 2022, pp. 1–21, Mar. 2022.

[42] A. A. Altameem and A. M. Hafez, "Behavior analysis using enhanced fuzzy clustering and deep learning," *Electronics*, vol. 11, no. 19, p. 3172, Oct. 2022. [Online]. Available: https://www.mdpi.com/2079-9292/11/19/3172

[43] W. Zhao, J. Ma, Q. Liu, J. Song, M. Tysklind, C. Liu, D. Wang, Y. Qu, Y. Wu, and F. Wu, "Comparison and application of SOFM, fuzzy $c$-means and $k$-means clustering algorithms for natural soil environment regionalization in China," *Environ. Res.*, vol. 216, Jan. 2023, Art. no. 114519.

[44] S. Zhu, R. Kong, X. Luo, Z. Xu, and F. Zhu, "An improved potential landslide hazard points evaluating method considering the heterogeneity of environmental features," *Int. J. Environ. Sci. Technol.*, vol. 20, no. 7, pp. 7951–7968, Jul. 2023.

[45] N. Bhandari and P. Pahwa, "Evaluating partitioning based clustering methods for extended non-negative matrix factorization (NMF)," *Intell. Autom. Soft Comput.*, vol. 35, no. 2, pp. 2043–2055, 2023.

[46] M. K. Ng and L. Jing, "A new fuzzy $k$-modes clustering algorithm for categorical data," *Int. J. Granular Comput., Rough Sets Intell. Syst.*, vol. 1, no. 1, pp. 105–119, 2009.

[47] A. Seman, Z. A. Bakar, and M. N. Isa, "An efficient clustering algorithm for partitioning Y-short tandem repeats data," *BMC Res. Notes*, vol. 5, no. 1, 2012, Art. no. 2101791285670500.

[48] D. Steinley, "Local optima in K-means clustering: What you don't know may hurt you," *Psychol. Methods*, vol. 8, no. 3, pp. 294–304, 2003.

[49] P. Fränti and S. Sieranoja, "K-means properties on six clustering benchmark datasets," *Int. J. Speech Technol.*, vol. 48, no. 12, pp. 4743–4759, Dec. 2018.

[50] I. Kärkkäinen and P. Fränti, "Dynamic local search algorithm for the clustering problem," Dept. Comput. Sci., Univ. Joensuu, Joensuu, Finland, Tech. Rep. A-2006-6, 2002.

[51] P. Fränti and O. Virmajoki, "Iterative shrinking method for clustering problems," *Pattern Recognit.*, vol. 39, no. 5, pp. 761–775, May 2006.

[52] P. Fränti, R. Mariescu-Istodor, and C. Zhong, "XNN graph," in *Proc. Structural, Syntactic, Stat. Pattern Recognit. (SSPR)*, Mérida, Mexico, Dec. 2016, pp. 207–217.

[53] P. Franti, O. Virmajoki, and V. Hautamaki, "Fast agglomerative clustering using a $k$-nearest neighbor graph," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1875–1881, Nov. 2006.

[54] T. Zhang, R. Ramakrishnan, and M. Livny, "BIRCH: An efficient data clustering method for very large databases," *ACM SIGMOD Rec.*, vol. 25, no. 2, pp. 103–114, Jun. 1996.

[55] M. Rezaei and P. Fränti, "Set matching measures for external cluster validity," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 8, pp. 2173–2186, Aug. 2016.

[56] C. J. Merz and P. M. Murphy, "UCI repository of machine learning databases," Dept. Inf. Comput. Sci., Univ. California, Irvine, CA, USA, 1996. [Online]. Available: http://www.ics.uci.edu/mlearn/MLRepository.html

[57] P. Christen, D. J. Hand, and N. Kirielle, "A review of the F-measure: Its history, properties, criticism, and alternatives," *ACM Comput. Surv.*, vol. 56, no. 3, pp. 1–24, Mar. 2024.

[58] D. J. Hand, P. Christen, and N. Kirielle, "F*: An interpretable transformation of the F-measure," *Mach. Learn.*, vol. 110, no. 3, pp. 451–456, Mar. 2021.

[59] K. N. Ismail, A. Seman, and K. A. F. A. Samah, "A comparison between external and internal cluster validity indices," in *Proc. IEEE 11th Int. Conf. Syst. Eng. Technol. (ICSET)*, Nov. 2021, pp. 229–233.

**KHAIRUL NURMAZIANNA ISMAIL** was born in 1984. She received the Diploma and B.Sc. (Hons.) degrees in computer science and the M.Sc. degree in information technology and quantitative science from the Universiti Teknologi MARA (UiTM), Malaysia, where she is currently pursuing the Ph.D. degree in computer science. With 11 years of working experience, including time spent in the industry, she is a Lecturer with the College of Computing, Informatics and Mathematics, UiTM, Melaka Branch, Jasin Campus. Her primary research interest includes machine learning.

**ALI SEMAN** was born in 1969. He received the Ph.D. degree in computer science from the Universiti Teknologi MARA (UiTM), Malaysia, in 2013. He has 28 years of working experience including in industry. He is currently a Senior Lecturer with the School of Computing, College of Computing, Informatics and Mathematics, UiTM. He is the author of several papers in international journals and conferences. His research interests include machine learning, optimization, algorithms, and bioinformatics applications. He also collaborates as a reviewer of several international journals.

**KHYRINA AIRIN FARIZA ABU SAMAH** (Member, IEEE) was born in Melaka, Malaysia, in 1976. She received the degree in information technology and management from Open University Malaysia, and the master's degree in computer science and the Ph.D. degree in information technology from the Universiti Teknikal Malaysia Melaka, Malaysia. She has 13 years of working experience in the semiconductor industry before joining Universiti Teknologi MARA (UiTM). She has been nominated as an Adjunct Professor with Universitas Pendidikan Indonesia and collaborates in research. She is currently a Senior Lecturer with the College of Computing, Informatics and Mathematics, UiTM, Melaka Branch, Jasin Campus. She has received several awards in publications, leads research projects, and is active in research. She has published articles under scientific standards. Her research interests include artificial intelligence, algorithm analysis, optimization, evacuation algorithms, and big data.