BAREKENG: Jurnal Ilmu Matematika dan Terapan
March 2023 Volume 17 Issue 1 Page 0125–0134
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss1pp0125-0134

IMPLEMENTATION OF THE FUZZY GUSTAFSON-KESSEL METHOD ON GROUPING DISTRICTS/CITIES IN KALIMANTAN ISLAND BASED ON POVERTY ISSUES FACTORS

Yunda Sasha Paradilla[1*], Memi Nor Hayati[2], Sifriyani[3]

[1,2,3] *Statistical Study Program, Faculty of Mathematics and Natural Sciences, Mulawarman University*
Jl. Barong Tongkok No.04, Samarinda, 75123, Indonesia

*Corresponding author's e-mail: * **yundaspd@gmail.com***

ABSTRACT

Article History:
Received: 28th August 2022
Revised: 21st November 2022
Accepted: 16th January 2023

Keywords:
Cluster Analysis;
FGK;
Poverty;
Validity Index.

Cluster analysis is an analysis that is useful in summarizing data by grouping objects based on certain similarity characteristics. One of the group analyses is Fuzzy Gustafson-Kessel (FGK) which is the development of the Fuzzy C-Means (FCM) method. The FGK method has a good way in adjusting the form of the cluster membership function correctly for data. This study aims to determine the optimal number of groups based on the Partition Coefficient (PC) and Classification Entropy (CE) validity indexes and to find out the results of grouping 56 districts/cities on the island of Kalimantan based on poverty issue factors in 2021. The optimal number of groups using the FGK method based on the validity indexes of PC and CE is two groups. The first group and the second group each consist of 28 districts/cities in Kalimantan Island.

How to cite this article:

Y. S. Paradilla, M. N. Hayati and Sifriyani, "IMPLEMENTATION OF THE FUZZY GUSTAFSON-KESSEL METHOD ON GROUPING DISTRICTS/CITIES IN KALIMANTAN ISLAND BASED ON POVERTY ISSUES FACTORS," *BAREKENG: J. Math. & App.*, vol. 17, iss. 1, pp. 0125-0134, March 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article • Open Access

1. INTRODUCTION

Cluster analysis is one of the analyzes used in data mining to group multidimensional objects, namely objects that can be described with several characteristics [1]. Cluster analysis was carried out using two grouping methods, namely the hierarchical method and the non-hierarchical method, where the group formation process needs to be carried out in such a way that each object is in exactly one group. But one day, it doesn't always happen because the object is between two or more other groups. So that the grouping is carried out using fuzzy clustering by considering the membership level of the fuzzy set as a weighting basis for the clustering [2].

There are several fuzzy clustering algorithms, one of which is Fuzzy C-Means (FCM). FCM is a data clustering where the existence of each point in a group is determined by the degree of membership [3]. An interesting extension of the FCM model is the Fuzzy Gustafson-Kessel (FGK) algorithm proposed by Gustafson and Kessel in 1979. This FGK method is a development of the FCM method [4]. The FGK method has a weakness when the fuzzy covariance matrix of the data is a singular matrix, so the matrix calculation cannot be applied [5]. In clustering with fuzzy clustering, optimal grouping results are needed by using the criterion of a measure of validity. One of the validity indexes that can be used is the Partition Coefficient (PC) index and the Classification Entropy (CE) index [6].

A previous study that compared the FCM and FGK methods using Quickbird satellite imagery data with the object of research in the Lubuk Batee Village area, Aceh Besar, by forming four grouping criteria. Based on this research, it can be seen that the FGK method is more accurate than the FCM method [7]. In addition, research on grouping analysis using the FCM and FGK methods on the LQ45 index financial report data, it was found that FGK had a better value than FCM [8]. Then research that discusses the application of fuzzy grouping methods, namely FCM and FGK, to group districts/cities based on the welfare data of the population in West Java Province in 2017, where it is known that the research results show FGK is better than FCM [9]. Furthermore, research that groups poverty in districts/cities in Papua based on information data on poverty in districts/cities in 2019 using the K-Means method, the results obtained using the silhouette index, the best results are 5 clusters [10].

Poverty is a condition of a person's inability to meet basic needs such as food, clothing, housing, education, and health [11]. The number of poor people in Indonesia in September 2020 reached 27,55 million people, or equivalent to 10,19% of the total population in Indonesia. This number increased by 1,13 million people in March 2020. This was due to the Corona Virus Disease 2019 pandemic (COVID-19), which rose 1,3% in urban areas and 0,6% in rural areas [12].

The largest percentage of poor people are in Maluku and Papua. This region contributed the highest about 20,65% of the total poverty in September 2020. In contrast, the lowest number of poor people was on the island of Kalimantan. In March 2021, the province of South Kalimantan was the province with the lowest poverty on the island of Kalimantan with a percentage of 4,83%, followed by Central Kalimantan, then East Kalimantan, then West Kalimantan and North Kalimantan with the highest poverty rate on the island of Kalimantan that is equal to 7,41% [12].

Many factors cause poverty in Indonesia and each region. Various government programs or policies are always driven and carried out to overcome the problems of poverty. In order to obtain the characteristics of poverty in each region or district/city, it is necessary to group them using statistics related to cluster analysis. Based on this description, the authors are interested in conducting scientific research with the title "Implementation of the Fuzzy Gustafson-Kessel Method on Grouping Districts/Cities in Kalimantan Island Based on Poverty Issues Factors."

2. RESEARCH METHODS

This study used the Fuzzy Gustafson-Kessel (FGK) method to sample data on factors of poverty issues based on 13 variables in 56 districts/cities on the island of Kalimantan in 2021. The sampling technique used is purposive sampling, where the sample is taken by considering the availability of the latest data about the factors of poverty issues on the island of Kalimantan.

2.1 Fuzzy Gustafson-Kessel Clustering (FGK)

The FGK method is a development of the FCM method. The FGK algorithm changes the distance calculation function into an adaptive distance norm function which is always updated at each iteration by using a fuzzy covariance matrix and the distance used in the FGK algorithm is the mahalonobis distance [13].

The algorithm of the FGK method is as follows [4]:

1. Determine the number of clusters (c), rank (m), maximum iteration ($MaksIter$), the smallest expected error (ε) and the initial objective function ($P_0 = 0$).
2. Generate a random number α_{ik} with $i = 1,2,3,\dots,n$ which is the number of data and the number of clusters $k = 1,2,\dots,c$ as elements of the initial partition matrix $\mathbf{U}_0$:

$$\mathbf{U}_0 = \begin{bmatrix} \alpha_{1,1} & \alpha_{1,2} & \cdots & \alpha_{1,c} \\ \alpha_{2,1} & \alpha_{2,2} & \cdots & \alpha_{2,c} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{n,1} & \alpha_{n,2} & \cdots & \alpha_{n,c} \end{bmatrix}$$

   Where the initial partition matrix is chosen randomly with the number of each column element in one row is 1 (one).

$$\sum_{k=1}^{c} (\alpha_{ik}) = 1$$

   α_{ik} is the value of a matrix based on rows and columns or the degree of membership which refers to how likely it is that data can be a member of a cluster.
3. Calculate the center of the k: v_{kj}, with $k = 1,2,\dots,c$ is the number of clusters, $j = 1,2,\dots,p$ as the number of variables and x_{ij} is the i-th sample and the j-th variable.

$$v_{kj} = \frac{\sum_{i=1}^{n}((\alpha_{ik})^m \times x_{ij})}{\sum_{i=1}^{n}(\alpha_{ik})^m} \qquad (1)$$

4. Calculate the grouping covariance matrix ($\mathbf{F}_k$) with the equation:

$$\mathbf{F}_k = \frac{1}{\sum_{i=1}^{n}(\alpha_{ik})^m} \times \mathbf{R}_k^T \times \mathbf{R}_k^* \qquad (2)$$

   where:

$$\mathbf{R}_k = \begin{bmatrix} (x_{11} - v_{k1}) & (x_{12} - v_{k2}) & \cdots & (x_{1p} - v_{kp}) \\ (x_{21} - v_{k1}) & (x_{22} - v_{k2}) & \cdots & (x_{2p} - v_{kp}) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{n1} - v_{k1}) & (x_{n2} - v_{k2}) & \cdots & (x_{np} - v_{kp}) \end{bmatrix}$$

$$\mathbf{R}_k^* = \begin{bmatrix} (\alpha_{1k})^m \times (x_{11} - v_{k1}) & (\alpha_{1k})^m \times (x_{12} - v_{k2}) & \cdots & (\alpha_{1k})^m \times (x_{1p} - v_{kp}) \\ (\alpha_{2k})^m \times (x_{21} - v_{k1}) & (\alpha_{2k})^m \times (x_{22} - v_{k2}) & \cdots & (\alpha_{2k})^m \times (x_{2p} - v_{kp}) \\ \vdots & \vdots & \ddots & \vdots \\ (\alpha_{nk})^m \times (x_{n1} - v_{k1}) & (\alpha_{nk})^m \times (x_{n2} - v_{k2}) & \cdots & (\alpha_{nk})^m \times (x_{np} - v_{kp}) \end{bmatrix}$$

5. Calculating distance

$$\mathbf{D}_k^2 = (\mathbf{R}_k \times \mathbf{A}_k)(\mathbf{R}_k^T) \qquad (3)$$

   Previously, we searched for the value of $\mathbf{A}_k$ which is a group distance matrix called the adaptive distance norm and n is the number of observational data.

$$\mathbf{A}_k = \left[\left(det(\mathbf{F}_k)^{\frac{1}{n}}\right)\right](\mathbf{F}_k{}^{-1}) \qquad (4)$$

The matrix formed from $\mathbf{D}_k^2$ is a square matrix that has the same number of rows as the number of columns shown as follows:

$$\mathbf{D}_k^2 = \begin{bmatrix} d_{11} & d_{12} & \cdots & d_{1n} \\ d_{21} & d_{22} & \cdots & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \cdots & d_{nn} \end{bmatrix}$$

Where is the main diagonal of the matrix $\mathbf{D}_k^2$ is the value that will be used for the mahalanobis distance symbolized as follows:

$$S_{ik}^2 = \begin{bmatrix} d_{11} \\ d_{22} \\ \vdots \\ d_{nn} \end{bmatrix} = \begin{bmatrix} s_{1k} \\ s_{2k} \\ \vdots \\ s_{nk} \end{bmatrix}$$

6. Calculating the objective function on the t-th iteration (P_t) with equation:

$$P_t = \sum_{i=1}^{n} \sum_{k=1}^{c} (\alpha_{ik})^m \times S_{ik}^2 \tag{5}$$

7. Calculates the new membership matrix change with elements α_{ik} with equation:

$$\alpha_{ik} = \left[\sum_{q=1}^{c} \left(\frac{S_{ik}}{S_{iq}} \right)^{\frac{2}{m-1}} \right]^{-1} \tag{6}$$

8. Repeat step 3 to step 7 until the condition $(|P_t - P_{t-1}|) < \varepsilon$ or (t > *MaksIter*) fulfilled.

2.2 Validity Index

The cluster validity index is an index to determine the number of optimal clusters formed **[4]**. Commonly used validity indexes:

1. *Partition Coefficient* (PC)

PC index is a method that measures the number of overlapping groups. The PC index equation is as follows:

$$PC\ (c) = \left(\frac{1}{n}\right) \sum_{i=1}^{n} \sum_{k=1}^{c} (\alpha_{ik})^2 \tag{7}$$

In the PC index the most optimal group is determined based on the largest PC value **[14]**.

2. *Classification Entropy* (CE)

The CE index is a method that measures the fuzziness of group partitions. The CE index equation is as follows:

$$CE\ (c) = \left(-\frac{1}{n}\right) \sum_{i=1}^{n} \sum_{k=1}^{c} \alpha_{ik} \ln(\alpha_{ik}) \tag{8}$$

The CE index evaluates the randomness of data in groups if all $\alpha_{i,k}$ values are in the range close to zero or one, so the lower the CE value, the better the grouping results **[14]**.

2.3 Research Variables

The variables used in this study consist of 13 variables, namely X_1 to X_{13} in the form of data on factors of poverty issues in 56 districts/cities on the island of Kalimantan in 2021. Obtained from the National Socio-Economic Survey (SUSENAS) March 2021 by the Central Statistics Agency (BPS) in every regency/city in Indonesia, including 5 Provinces on the island of Kalimantan through publication in the form of the book “Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2021” **[15]**.

2.4 Stages of Analysis

The stages of analysis carried out are as follows:

1. *Input* data to be grouped
2. Describing research data
3. Determine c, m, *MaksIter*, ε, and P_0
4. Generating random numbers α_{ik}
5. Calculate cluster center (v_{kj})
6. Calculating the group covariance matrix ($\mathbf{F}_k$)
7. Calculating distance $\mathbf{D}_k^2$
8. Calculating objective function (P_t)
9. Calculating membership matrix changes
10. Repeat steps 4 through 8 until ($|P_t - P_{t-1}|$) $< \varepsilon$ or (t > *MaksIter*) fulfilled
11. Calculating validity index

3. RESULTS AND DISCUSSION

Before entering into the analysis stage, first the data input will be grouped, namely: x_{ij}, with $i = 1,2, \dots ,56$; and $j = 1,2, \dots ,13$.

3.1 Descriptive Statistical Analysis

Descriptive statistical analysis on poverty issue factors data based on 13 variables in 56 districts/cities on the island of Kalimantan in 2021 using the minimum value, maximum value, average and standard deviation. The results of descriptive statistical analysis can be seen in **Table 1**. as follows:

Table 1. Descriptive Statistical Analysis

Variable (X_j)	Minimum	Maximum	Average	Standard Deviation
X_1	6,1900	49,7200	25,5200	11,3362
X_2	36,5200	74,3700	54,4400	9,6656
X_3	3,4700	56,9200	20,0300	10,3303
⋮	⋮	⋮	⋮	⋮
X_{13}	1,4500	71,8300	23,2000	15,6867

3.2 Fuzzy Gustafson-Kessel (FGK) Clustering Method

The grouping process using the Fuzzy Gustafson-Kessel (FGK) method is carried out with the following steps:

3.2.1 Determination of Parameter Values of FGK Method

This study uses clusters (c) = 2, 3, 4, 5 dan 6, $m = 3{,}75$, $MaxIter = 1000$, $\varepsilon = 10^{-3}$ and $P_0 = 0$. As an example of the calculation in this study, the FGK grouping will be shown using $c = 2$.

3.2.2 Generating Random Numbers

After determining the parameter values, the next step in the grouping process using the FGK method is to generate random numbers α_{ik}, as elements of the initial membership matrix $\mathbf{U}_0$ with the help of software R. The initial membership values with $c = 2$ are presented in **Table 2** as following:

Table 2. Initial Membership Value with $c = 2$ in FGK Method

No.	District/City	Initial Membership Value Cluster 1 (α_{i1})	 Cluster 2 (α_{i2})
1.	Sambas	0,3525	0,6475
2.	Bengkayang	0,0723	0,9277
3.	Landak	0,4420	0,5580
⋮	⋮	⋮	⋮
56.	Tarakan City	0,1509	0,8491

The value α_{ik} is determined randomly with the condition that the number of element values in each row must be 1, the initial membership matrix of $\mathbf{U}_0$ is 56 × 2 as follows:

$$\mathbf{U}_0 = \begin{bmatrix} 0,3525 & 0,6475 \\ 0,0723 & 0,9277 \\ \vdots & \vdots \\ 0,1509 & 0,8491 \end{bmatrix}$$

3.2.3 Calculating Cluster Center

After generating the random number α_{ik}, the next step is to calculate the center of the initial group using **Equation (1)** with the help of Microsoft Excel. So that the center of the initial group can be seen in **Table 3** as follows:

Table 3. Initial Cluster Center Element with $c = 2$ in FGK Method

Variable	Cluster Center	
	1	2
X_1 (Percentage of the Poor Age 15 and Over Graduated below Elementary School)	25,0003	26,2425
X_2 (Percentage of the Poor Age 15 and Over Graduated from Elementary School/Junior High School)	55,3567	54,9399
X_3 (Percentage of the Poor Age 15 and Over Graduated High School or College Education)	19,6416	18,8023
⋮	⋮	⋮
X_{13} (Percentage of Poor Households Receiving Poor Rice/Prosperous Rice)	20,4244	27,0254

The matrix form of the center of group $\mathbf{V}$ formed in the first iteration $(t = 1)$, where v_{kj} is a row vector with 13 elements as follows:

$$\mathbf{V} = \begin{bmatrix} 25,0003 & 55,3567 & \cdots & 20,4244 \\ 26,2425 & 54,9399 & \cdots & 27,0254 \end{bmatrix}$$

3.2.4 Calculating Group Covariance Matrices

After calculating the initial group center, the next step is to calculate the group covariance matrix in the first iteration $(t = 1)$ with the number of groups as much as $c = 2$ using **Equation (2)**. The value of the covariance matrix of the groups formed in the first iteration with FGK for group 1 and group 2 measuring 13 × 13 is as follows:

$$\boldsymbol{F}_1 = \begin{bmatrix} 147,9979 & -64,0699 & \cdots & 21,0043 \\ -64,0699 & 107,1799 & \cdots & 21,8042 \\ \vdots & \vdots & \cdots & \vdots \\ 21,0043 & 21,8042 & \cdots & 157,2831 \end{bmatrix}$$

$$\boldsymbol{F}_2 = \begin{bmatrix} 133,3552 & -69,3813 & \cdots & -21,7215 \\ -69,3813 & 80,9000 & \cdots & -27,1330 \\ \vdots & \vdots & \cdots & \vdots \\ -21,7215 & -27,1330 & \cdots & 400,1525 \end{bmatrix}$$

3.2.5 Calculating Distance

After calculating the group covariance matrix, the next step is to calculate the mahalonobis distance using **Equation (3)**. Where the first calculation of the value of the distance matrix of the group $\mathbf{A}_k$ with a value of $n = 56$ using **Equation (4)**. The value of the distance matrix for group $\mathbf{A}_k$ formed at $t = 1$ with FGK for group 1 and group 2 is 13 × 13 as follows:

$$\mathbf{A}_1 = \begin{bmatrix} 1.633.214.846.479.460 & 1.632.480.665.760.660 & \cdots & 1.417.217.495.817 \\ 1.632.480.665.760.660 & 1.631.786.567.926.820 & \cdots & 1.404.703.475.763 \\ \vdots & \vdots & \ddots & \vdots \\ 1.417.217.495.817 & 1.404.703.475.763 & \cdots & 28.277.624.480 \end{bmatrix}$$

$$\mathbf{A}_2 = \begin{bmatrix} 4.453.944.720.660.670 & 4.455.808.335.996.100 & \cdots & -3.525.273.658.344 \\ 4.455.808.335.996.100 & 4.457.895.429.185.800 & \cdots & -3.498.275.802.749 \\ \vdots & \vdots & \ddots & \vdots \\ -3.525.273.658.344 & -3.498.275.802.749 & \cdots & 71.770.206.287 \end{bmatrix}$$

After calculating the distance matrix of the $\mathbf{A}_k$ group, it can be continued to calculate the distance of the mahalonobis $\mathbf{D}_k^2$ in the k-th group. The distance matrix value of the first group ($\mathbf{D}_1^2$) and the second group ($\mathbf{D}_2^2$) is 56 × 56. Then the main diagonal elements formed in the distance matrix $\mathbf{D}_1^2$ and $\mathbf{D}_2^2$ are the results of the mahalonobis distance using the FGK method at $t = 1$ for group 1 (S_{i1}^2) and group 2 (S_{i2}^2). The results of the calculation of the mahalonobis distance for each data can be seen in **Table 4** as follows:

Tabel 4. Distance Value with $c = 2$ in FGK Method

No.	District/City	Distance value	
		Cluster 1 (S_{i1}^2)	Cluster 2 (S_{i2}^2)
1.	Sambas	39.554.058.556.916	136.165.064.723.348
2.	Bengkayang	35.127.489.879.202	78.214.865.243.053
3.	Landak	24.174.810.579.990	117.582.158.420.883
⋮	⋮	⋮	⋮
56.	Tarakan City	41.690.953.938.701	77.542.564.633.453

3.2.6 Calculating Objective Functions

After calculating the distance, then the initial membership value in Table 2 and the distance value in **Table 4** are used to calculate the value of the objective function in the first iteration ($t = 1$) with the number of groups of $c = 2$ using **Equation (5)**. The objective function value at $t = 1$ is 1.802.428.978.784.920. The initial objective function value (P_0) is 0 so $|P_1 - P_0|$ = 1.802.428.978.784.920 > $\varepsilon = 10^{-3}$ because the change in the objective function is still greater than the value of the process continues to the next iteration.

3.2.7 Calculating Membership Matrix Changes

After calculating the value of the objective function in the first iteration ($t = 1$), the next step is to calculate the change in membership value with the number of groups as much as $c = 2$ using **Equation (6).**

After calculating the change in the membership matrix, the next step is to recalculate the group center, the objective function with the updated group center, and calculate the change in the matrix again with $c = 2$. The iteration stops when $|P_t - P_{t-1}| < 0{,}001$ or $t > 1000$. In this study, the stop step is at the 43rd iteration with the help of software R. The final results obtained are the membership values of 56 districts/cities which are shown in **Table 5** as follows:

Table 5. Membership Value at Last Iteration with $c = 2$ In FGK Method

No.	District/City	Membership Value		Followed Cluster
		Cluster 1	Cluster 2	
1.	Sambas	0,4838	0,5162	2
2.	Bengkayang	0,4682	0,5318	2
3.	Landak	0,4616	0,5384	2
⋮	⋮	⋮	⋮	⋮
56.	Tarakan City	0,5143	0,4857	1

3.2.8 Calculating Validity Index

The results of the calculation of the validity index values for c = 2, 3, 4, 5 and 6 can be seen in **Table 6** as follows:

Table 6. Validity Index Value for Overall Parameter c in FGK Method

Clusters (c)	2	3	4	5	6
PC	**0,5363**	0,3927	0,3077	0,2692	0,2437
CE	**0,6500**	1,0105	1,2850	1,4726	1,6248

Based on **Table 6** the PC and CE values obtained from many groups 2, 3, 4, 5 and 6, the largest PC value is 0,5363 in clusters 2 while the smallest CE value is 0,6500 in clusters 2. So that the optimal group is clusters 2.

3.3 Interpretation of Grouping Results

After grouping using the Fuzzy Gustafson-Kessel (FGK) method, group validation was carried out by looking at the PC and CE values, where the most optimal grouping results were the groups that produced the largest PC values and the smallest CE values.

The results of the best grouping of districts/cities on the island of Kalimantan based on the factors of poverty issues are the results of grouping with the optimal number of groups as much as 2. The following group members are based on the results of calculations with c = 2 in the FGK method, namely:

Cluster 1: Sintang, Kubu Raya, Kotawaringin Timur, Kapuas, Barito Selatan, Barito Utara, Lamandau, Seruyan, Katingan, Pulang Pisau, Gunung Mas, Murung Raya, Palangka Raya City, Banjar, Tapin, Hulu Sungai Tengah, Tanah Bumbu, Balangan, Banjarmasin City, Banjar Baru City, Kutai Kartanegara, Kutai Timur, Mahakam Hulu, Samarinda City, Bontang City, Bulungan, Nunukan dan Tarakan City.

Cluster 2: Sambas, Bengkayang, Landak, Pontianak, Sanggau, Ketapang, Kapuas Hulu, Sekadau, Melawi, Kayong Utara, Pontianak City, Singkawang City, Kotawaringin Barat, Sukamara, Barito Timur, Tanah Laut, Kota Baru, Barito Kuala, Hulu Sungai Selatan, Hulu Sungai Utara, Tabalong, Paser, Kutai Barat, Berau, Penajam Paser Utara, Balikpapan City, Malinau dan Tana Tidung.

4. CONCLUSIONS

Based on the results of data analysis and discussion, the following conclusions can be drawn:

1. Based on the results of the validity index, namely Partition Coefficient (PC) and Classification Entropy (CE), the optimal number of groups to classify districts/cities in Kalimantan Island based on the factors of poverty issues are 2 cluster swith the largest PC value being 0,5363 and the smallest CE is 0,6500.
2. The results of grouping using the Fuzzy Gustafson-Kessel (FGK) method in districts/cities on the island of Kalimantan based on poverty issue factors with the largest PC value is 0,5363 and the smallest CE value is 0,6500 in many groups 2, namely, there are 28 districts/cities that are included in cluster 1 members and 28 districts/cities that are included in cluster 2 members.

REFERENCES

[1] Nurjanah, F. Andi dan F. Indriani, "Implementasi Metode Fuzzy C-Means Pada Sistem Clustering Data Varietas Padi," *Kumpulan Jurnal Ilmu Komputer (KLIK),* vol. 01, no. 01, pp. 23-22, 2014.

[2] J.-S. R. Jang, C.-T. Sun and E. Mizutani, Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence, New Jersey: Prentice-Hall Inc, 1997.

[3] S. Kusumadewi dan H. Purnomo, Aplikasi Logika Fuzzy: Untuk Pendukung Keputusan Edisi 2, Yogyakarta: Graha Ilmu, 2010.

[4] A. Balasko, J. Abonyi and B. Feil, Fuzzy Clustering and Data Analysis Toolbox: For Use with Matlab, Veszprem: Department of Process Engineering University of Veszprem, 2007.

[5] R. Babuka, P. v. d. Veen and U. Kaymak, "Improved Covariance Estimation for Gustafson-Kessel Clustering," *In Proceedings of the 2002 IEEE International Conference on Fuzzy Systems,* pp. 1081-1085, 2002.

[6] S. M. H. Jansen, Customer Segmentation and Customer Profiling for a Mobile Telecommunications Company Based on Usage Behavior: A Vodafone Case Study, Maastricht: University of Maastricht (UM), 2007.

[7] A. M. M, H. Sofyan and M. Subianto, "Perbandingan Metode Fuzzy C-Means (FCM) dan Fuzzy Gustafson-Kessel (FGK) Menggunakan Data Citra Satelit Quickbird (Studi Kasus: Desa Lubuk Batee, Aceh Besar)," *Jurnal Transenden,* vol. 00, pp. 01-05, 2013.

[8] L. Rahmatika, S. Suparti and D. Safitri, "Analisis Kelompok dengan Algoritma Fuzzy C-Means dan Gustafson Kessel Clustering Pada Indeks LQ45," *Jurnal Gaussian,* vol. 4, no. 3, pp. 543-551, 2015.

[9] I. Izzati, Penerapan Algoritma Fuzzy Clustering Pada Pengelompokan Kabupaten/Kota di Provinsi Jawa Barat Berdasarkan Indikator Kesejahteraan Rakyat., Jatinangor: Unpad Repository, 2020.

[10] I. Minitansia, Pengelompokan Kemiskinan Kabupaten/Kota di Papua dengan Menggunakan Metode K-Means, Surabaya: UIN Sunan Ampel Surabaya, 2021.

[11] K. PPN/Bappenas, Analisis Wilayah dengan Kemiskinan Tinggi, Jakarta: Kedeputian Bidang Kependudukan dan Ketenagakerjaan Kementrian PPN/Bappenas, 2018.

[12] B. P. S. BPS, Jumlah Penduduk Miskin Menurut Provinsi, 2007-2021, Jakarta: Badan Pusat Statsitik Republik Indonesia, 2021.

[13] D. E. Gustafson and W. C. Kessel, "Fuzzy Clustering with a Fuzzy Covariance Matrix," *IEEE,* vol. 17, pp. 761-766, 1979.

[14] H. Sakai, M. K. Chakraborty, D. Slezak, A. E. Hassanien dan W. Zhu, Transactions on Rough Sets XIV, London: Springer, 2011.

[15] A. Arie, T. Mumtaz and N. Dwima, Data dan Informasi Kemiskinan Kabupaten/Kota Tahun 2021, Jakarta: Badan Pusat Statistik, 2021.