

Enhanced Possibilistic Fuzzy C-Means Clustering Stunting Prevalence in Indonesia

Alivia F. Zahro, Ria Dhea L.N. Karisma*, Usman Pagalay

Department of Mathematics, UIN Maulana Malik Ibrahim Malang, Indonesia
afaizah27@gmail.com, riadhea@uin-malang.ac.id, usman@mat.uin-malang.ac.id

**Correspondence: riadhea@uin-malang.ac.id*

ABSTRACT. Stunting in toddlers is a significant public health problem in Indonesia due to its potential to inhibit child development and cause long-term adverse effects. Clustering the prevalence of stunting provides valuable insights for designing effective prevention policies. This study employs the Possibilistic Fuzzy C-Means (PFCM) method, validated using the Modified Partition Coefficient (MPC) index, to cluster stunting prevalence in Indonesia. The PFCM method integrates Fuzzy C-Means (FCM) and Possibilistic C-Means (PCM), balancing membership degrees with probabilistic measures. The primary advantages of this method are its capability to handle data with uncertain membership degrees, robustness against noise, and flexibility in defining probabilistic membership values. The results obtained show that clusters with high stunting prevalence are dominated by nine provinces, namely Aceh, Jambi, Bengkulu, Bangka Islands, Central Kalimantan, Central Sulawesi, Gorontalo, West Papua, and Papua. The MPC validity score of 0.704 confirms the effectiveness of the PFCM method in categorizing stunting prevalence well, making it a robust tool to support policymaking in stunting prevention efforts.

1. INTRODUCTION

Clustering is an analytical technique that categorizes data into groups based on specific similarities and characteristics. Similarity is often measured by the proximity of objects, determined using a distance function [1]. This technique uncovers hidden patterns in data, facilitating more natural grouping [2]. In clustering, data can be grouped into two main categories, namely crisp clustering and fuzzy clustering [3].

Fuzzy clustering is a method used to identify the optimal cluster within a vector space by measuring the distance between vectors using the Euclidean norm [4]. Fuzzy clustering has the advantage of the concept of crisp clustering, which can overcome data that has membership degrees in several clusters [3]. Some algorithms in fuzzy clustering include Fuzzy C-Means (FCM), Fuzzy Possibilistic C-Means (FPCM), and Possibilistic Fuzzy C-Means (PFCM). Possibilistic Fuzzy C-Means (PFCM) is a clustering method that combines FCM and PCM. The method

Received: 14 Aug 2025.

Key words and phrases. Stunting Prevalence; Possibilistic Fuzzy C-Means; Clustering; Modified Partition Coefficient; Indonesia.

was first introduced by Pal et al. in 2005. PFCM is designed to overcome the weaknesses found in each method, namely sensitivity to outlier data in FCM, the problem of overlapping clusters in PCM, and limitations in setting possibilistic membership values in FPCM [5]. PFCM aims to balance the membership degree value with the typicality value [6].

Since clustering is an unsupervised learning process, validating clustering results is essential to assess the quality and reliability of the obtained clusters [7]. Modified Partition Coefficient (MPC) is a validity index that is an update of Partition Coefficient (PC) [8]. The method was introduced to address the limitation of monotonic changes in the Partition Coefficient (PC), as introduced by Dave in 1996 [9]. MPC values approaching 1 indicate higher accuracy and validity of clustering results [10].

Previous studies indicate that the PFCM algorithm outperforms others in various applications. For example, Simhachalam and Ganesan's research [11] implemented the PFCM algorithm for medical diagnosis, which proved to be more efficient than other algorithms. In addition, research by Kalist et al. [12] showed that PFCM excels in satellite image segmentation, and Ozdemir & Kaya's research [13] compared FCM, PCM, FPCM, and PFCM, where PFCM proved to have the best performance in clustering.

Several clustering methods are applied in various fields. However, research on the application of clustering in the healthcare sector, especially stunting, a national issue in Indonesia, remains limited. Conventional methods such as k-means and hierarchical clustering are limited, which is handling data uncertainty and overlapping regional characteristics due to the assumption that binary membership and the inability to handle data with partial degrees of membership [14]. Moreover, some studies only focus on improving the accuracy of clustering validation metrics, such as the Silhouette score and Davies-Bouldin Index (DBI), without exploring clustering models that can be effectively translated into data-based policy recommendations.

In conventional methods such as FCM, a high dependence on the fuzzier parameter (m), directly influences the membership distribution. Meanwhile, PCM has weaknesses in handling outliers, as its membership values often cannot be effectively controlled [15]. To address this limitation, PFCM integrates both probabilistic and possibilistic elements. However, optimally tuning its parameters, specifically a , b , m , and η to achieve a balance of fuzzy and possibilistic memberships is still challenging. The constants a and b define the relative importance of fuzzy membership and typicality values in the objective function. If a high value of b is used compared to a , the centroid value will be more influenced by the typicality value than the membership value, which can help reduce outliers [6]. Therefore, to address this issue, this study employs the Modified Partition Coefficient (MPC). MPC serves as a validation index and facilitates adjusting parameters in the PFCM. Thus, its approach is expected to yield more balanced membership contributions and typicality values, enhance resilience to outliers, and improve clustering accuracy, particularly in cases of stunting prevalence in Indonesia.

In the context of stunting in Indonesia, clustering can help categorize regions based on prevalence rates to support more specific policies. According to the WHO, stunting is a condition in which the growth and development of children under five is impaired due to chronic malnutrition over a long period of time, as indicated by length or height below the appropriate standard [16]. In Indonesia, the prevalence of stunting reached 21.6% in 2022 from 24.4% in 2021 [17]. However, according to WHO standards, this decrease in prevalence is still categorized as high (>20%). Stunting significantly affects children's physical, cognitive, and social development, necessitating data-driven interventions for more effective prevention. In this study, the PFCM method was chosen due to its ability to handle data with uncertain membership degrees and sensitivity to noisy data. The clustering results are intended to serve as a basis for designing more targeted policies tailored to each region's needs.

2. RESEARCH METHODS

2.1. Data Analysis. The steps of analysis in this study using R software were carried out through several steps [18]. First, the data is entered in the form of a $p \times q$ matrix. Next, the data standardization process was carried out. After that, the initial parameters were determined, including the number of clusters (c), FCM weighting rank (m), PCM weighting rank (η), maximum iteration, minimum error (ε), initial objective function ($J_0 = 0$), initial iteration ($t = 1$), coefficient γ (K), weighting coefficient for relative typicality matrix U (a), and weighting coefficient for absolute typicality matrix T (b).

The PFCM clustering process requires the results of FCM clustering, starting with generating the partition matrix elements U using a random number u_{ik} . Then, the number of each row (Q_i) is calculated by equation (1) where $1 \leq i \leq p$. The partition matrix element U is normalized using equation (2). Next, the cluster center V is calculated by equation (3) and the objective function $J_m(t)$ is calculated by equation (5). The process continues by fixing the elements of the partition matrix U using equation (7). Then the stop condition is checked with the criteria If $(|J_m(t) - J_m(t-1)|) < \varepsilon$ or $t > \text{MaxIter}$, if it has not been achieved, then the iteration continues by fixing the cluster center V .

$$Q_i = \sum_{k=1}^c u_{ik}, \quad (1)$$

$$\mu_{ik} = \frac{u_{ik}}{Q_i}, \quad (2)$$

$$v_{kj} = \frac{\sum_{i=1}^p (\mu_{ik})^m x_{ij}}{\sum_{i=1}^p (\mu_{ik})^m}, \quad (3)$$

$$d_{ik} = d(x_i, v_k) = \left[\sum_{j=1}^q (x_{ij} - v_{kj})^2 \right]^{1/2}, \quad (4)$$

$$J_{m(t)} = \sum_{i=1}^p \sum_{k=1}^c (\mu_{ik})^m (d_{ik})^2, \quad (5)$$

$$d_{iz} = d(x_i, v_z) = \left[\sum_{j=1}^q (x_{ij} - v_{zj})^2 \right]^{1/2}, \quad (6)$$

$$\mu_{ik} = \left[\sum_{z=1}^c \left(\frac{d_{ik}}{d_{iz}} \right)^{2/(m-1)} \right]^{-1}. \quad (7)$$

After obtaining the results of the partition matrix U and cluster center V in FCM, the PFCM clustering process can be started by calculating the typicality matrix element T (t_{ik}) where the partition matrix elements U and cluster center V of the FCM clustering results. The element T is calculated by equation (8). The cluster center is fixed by equation (9) and the objective function $J_{m,\eta}(t)$ is calculated by equation (11). In calculating the objective function $J_{m,\eta}(t)$, the coefficient γ_k is required to control the spread of the cluster, ensuring greater stability.

$$t_{ik} = \frac{1}{1 + \left(\frac{b(d_{ik})^2}{\gamma_k} \right)^{1/(\eta-1)}}, \quad (8)$$

$$v_{kj} = \frac{\sum_{i=1}^p (a(\mu_{ik})^m + b(t_{ik})^\eta) x_{ij}}{\sum_{i=1}^p (a(\mu_{ik})^m + b(t_{ik})^\eta)}, \quad (9)$$

$$\gamma_k = K \frac{\sum_{i=1}^p (\mu_{ik})^\eta (d_{ik})^2}{\sum_{i=1}^p (\mu_{ik})^\eta}, \quad (10)$$

$$J_{m,\eta}(t) = \sum_{i=1}^p \sum_{k=1}^c (a(\mu_{ik})^m + b(t_{ik})^\eta) (d_{ik})^2 + \sum_{k=1}^c \gamma_k \sum_{i=1}^p (1 - t_{ik})^\eta. \quad (11)$$

Then, the partition matrix element U is fixed by equation (7), and the absolute typicality matrix T is updated using equation (8). With the enhancement of the absolute distinctiveness matrix T , the clustering results become more robust against outliers. Therefore, coefficients a and b are needed to handle them. Coefficient b must be larger than coefficient a to better handle outliers. The stop condition is checked by the criteria $(|J_{m,\eta}(t) - J_{m,\eta}(t-1)|) < \varepsilon$ or $t > \text{MaxIter}$. If the stopping criterion is not fulfilled, then the iteration continues from the step of updating the cluster center V based on equation (9). After the clustering results are obtained, the next step is to calculate the validity of the clustering results using equation (13) [8]. The greater the MPC value, which is close to 1, it shows better cluster validity [9]. Then the final stage is to interpret the clustering results.

2.2. Z-Score. Data standardization is used when variables have different measurement scales. One method for standardization is the Z-Score method. Z-scores convert variable values into a standardized form, where each variable follows a distribution with a mean of 0 and a standard deviation of 1, allowing differences in scale to be overcome. The Z-score can be calculated using the following equation (12).

$$Z_{ij} = \frac{X_{ij} - \mu_j}{\sigma_j} \quad (12)$$

2.3. Modified Partition Coefficient Validity. MPC is a validity index that can reduce monotonous changes in the PC proposed by Dave [9]. MPC calculation is done by calculating between the value of the membership degree and the cluster center value. The MPC value has a range between 0 to 1 [18]. MPC equation can be written as the following equation (13).

$$\text{MPC} = 1 - \frac{c}{c-1} \left(1 - \frac{1}{p} \sum_{k=1}^c \sum_{i=1}^p (\mu_{ik})^2 \right) \quad (13)$$

From equation (13), the greater the MPC value, the better the cluster validity [9]. As for the MPC validity criteria can be seen in Table 1 [19].

TABLE 1. MPC Validity Criteria

MPC Value	Criteria
0.8–1.0	Excellent
0.5–0.7	Good
0.3–0.4	Fair
0.0–0.2	Poor

2.4. Data Source. This study uses data from the Ministry of Health Republic Indonesia (Kementerian Kesehatan Indonesia), including health profile data from 399 districts/cities in 2022. The data was then analyzed using the Possibilistic Fuzzy C-Means method (PFCM) and validation of clustering results using the Modified Partition Coefficient. The data variables used in the study are shown in Table 2.

TABLE 2. Data Variables

No.	Variables	Scale
1	The Number of Pregnant Women Receiving Blood Enhancement Tablets (X_1)	person
2	The Number of Infants Receiving Early Breastfeeding Initiation (X_2)	person
3	The Number of Infants Receiving Exclusive Breastfeeding for 6 Months (X_3)	person
4	The Number of Infants Receiving Vitamin A (X_4)	person
5	The Number of Infants Receiving Complete Basic Immunization (X_5)	person
6	The Number of Low-Weight Infants (X_6)	person
7	The Number of Families with Access to Proper Sanitation Facilities (X_7)	household

The stages of the research are as follows

Algorithm 1 Stages of the Research

- 1: Collect research data.
- 2: Input data for clustering.
- 3: Perform data standardization.
- 4: Set initial parameters: $c, m, \eta, \text{MaxIter}, \varepsilon, J_0 = 0, t = 1, K, a, b$.
- 5: Generate partition matrix U using random numbers u_{ik} .
- 6: Calculate Q_i using (1).
- 7: Normalize U using (2).
- 8: Calculate cluster center V using (3).
- 9: Calculate objective function $J_m(t)$ using (5).
- 10: Update U using (7).
- 11: **if** $|J_m(t) - J_m(t - 1)| < \varepsilon$ **or** $t > \text{MaxIter}$ **then**
- 12: **Stop iteration**
- 13: **else**
- 14: If the stopping criterion is not satisfied, go back to Step 8 to update the cluster centers V and continue the iterations.
- 15: **end if**
- 16: Calculate typicality matrix T using (8), where γ_k and d_{ik} are obtained from (4) and (10).
- 17: Update V using (9).
- 18: Calculate $J_{m,\eta}(t)$ using (11).
- 19: Update U using (7).
- 20: Update T using (8).
- 21: **if** $|J_{m,\eta}(t) - J_{m,\eta}(t - 1)| < \varepsilon$ **or** $t > \text{MaxIter}$ **then**
- 22: **Stop iteration**
- 23: **else**
- 24: Repeat the update steps until the stopping criterion is satisfied.
- 25: **end if**
- 26: Calculate MPC validity using (13).

3. RESULTS AND DISCUSSION

3.1. Data Standardization. Data standardization was conducted in this study because there were variables with different measurement scales, namely on the access to proper sanitation facilities variable. Data standardization in this study is performed using the Z-Score method, as shown in equation (12). The results of this data standardization process are presented in Table 3.

3.2. Possibilistic Fuzzy C-Means Clustering Result. PFCM clustering begins with FCM clustering to obtain the partition matrix U and cluster center matrix V that will be used in

TABLE 3. Data Standardization

No.	Districts/Cities	X_1	X_2	X_3	\dots	X_7
1	South Aceh	-0.433	-0.485	-0.427	\dots	-0.576
2	Southeast Aceh	-0.452	-0.455	-0.463	\dots	-0.667
3	East Aceh	-0.076	0.017	-0.267	\dots	-0.277
\vdots	\vdots	\vdots	\vdots	\vdots	\ddots	\vdots
397	Intan Jaya	-0.750	-0.730	-0.598	\dots	-0.793
398	Deiyai	-0.750	-0.730	-0.598	\dots	-0.794

the PFCM stage. Before the clustering stage, a series of pre-research trials were conducted to determine the optimal parameters as follows:

- Number of clusters (c), $c = 3$
- Rank of FCM weights (m), $m = 2$
- Rank of PCM weights (η), $\eta = 2$
- Maximum iterations = 1000
- Minimum error = 10^{-1}
- Initial objective function (J_0) = 0
- Initial iteration ($t = 1$)
- Coefficient $\gamma(K) = 1$
- Weighting coefficient for relative typicality matrix U (a), $a = 1$
- Weight coefficient for absolute typicality matrix T (b), $b = 9$

Based on the research by [20] the optimal weighting rank m is found to be $m = 2$ and $\eta = 2$. Additionally, based on the research by Pal et al. [5] the optimal value of K is determined to be $K = 1$, so the coefficient K used is 1. The number of clusters to be formed is three, based on the classification of stunting categories in Indonesia is low, medium, and high. So the parameter c is set to $c = 3$.

The clustering results using the FCM method produce an initial partition matrix normalized, with the number of elements in each row equal to 1. Furthermore, an iterative process is carried out to calculate the cluster center and objective function until it gets convergent results. The FCM iterative process reached convergence after 17 iterations, with the last objective function value of 579.899. The final partition matrix \mathbf{U} is normalized, such that the sum of the elements in each row equals one.

$$U = \begin{bmatrix} 0.055 & 0.006 & 0.939 \\ 0.008 & 0.001 & 0.991 \\ 0.207 & 0.016 & 0.777 \\ \vdots & \vdots & \vdots \\ 0.050 & 0.008 & 0.942 \\ 0.050 & 0.008 & 0.942 \\ 0.035 & 0.005 & 0.961 \end{bmatrix}$$

$$V = \begin{bmatrix} 0.502 & 0.542 & 0.403 & \cdots & 0.825 \\ 3.042 & 3.015 & 1.982 & \cdots & 2.854 \\ -0.423 & -0.428 & -0.331 & \cdots & -0.501 \end{bmatrix}$$

The final partition matrix U and cluster center V show convergent results and are ready to be used for PFCM clustering of stunting prevalence.

PFCM clustering starts by calculating the absolute typicality matrix by calling the results of the partition matrix and cluster center in the FCM process. The aim is to improve clustering results through a possibilistic approach.

$$T = \begin{bmatrix} 0.048 & 0.027 & 0.158 \\ 0.043 & 0.026 & 0.550 \\ 0.083 & 0.032 & 0.068 \\ \vdots & \vdots & \vdots \\ 0.029 & 0.023 & 0.108 \\ 0.029 & 0.023 & 0.109 \\ 0.041 & 0.026 & 0.205 \end{bmatrix}$$

The absolute typicality matrix T in the PFCM algorithm is not subject to the same constraints as traditional fuzzy membership values. Consequently, typicality values do not require normalization across clusters.

Subsequently, the cluster center V is updated so that each cluster center is determined by both the probabilistic membership degree and the possibilistic typicality degree.

$$V = \begin{bmatrix} 0.426 & 0.445 & 0.326 & \cdots & 0.762 \\ 2.477 & 2.450 & 1.505 & \cdots & 2.373 \\ -0.446 & -0.445 & -0.356 & \cdots & -0.515 \end{bmatrix}$$

The result of the cluster center matrix is used to calculate the objective function. Then, it is followed by fixing the relative typicality matrix U and the absolute typicality matrix T .

$$U = \begin{bmatrix} 0.121 & 0.871 & 0.008 \\ 0.135 & 0.864 & 0.001 \\ 0.083 & 0.889 & 0.028 \\ \vdots & \vdots & \vdots \\ 0.166 & 0.825 & 0.009 \\ 0.166 & 0.825 & 0.009 \\ 0.139 & 0.856 & 0.006 \end{bmatrix}$$

$$T = \begin{bmatrix} 0.299 & 0.091 & 0.917 \\ 0.270 & 0.090 & 0.988 \\ 0.441 & 0.111 & 0.800 \\ \vdots & \vdots & \vdots \\ 0.193 & 0.076 & 0.886 \\ 0.194 & 0.076 & 0.887 \\ 0.262 & 0.089 & 0.938 \end{bmatrix}$$

It aims to obtain results that are more robust to outliers by considering the cluster centers obtained previously. The stage is repeated until it reaches convergence. The PFCM iterative process reached convergence after 12 iterations, with a final objective function value of 3.866943×10^3 .

$$|J_{12} - J_{11}| < \varepsilon$$

$$|3.866943 \times 10^3 - 3.867011 \times 10^3| < 10^{-1}$$

$$6.8 \times 10^{-2} < 10^{-1}$$

The calculation proves that the 12th iteration has reached convergence then the results of the cluster center matrix V, relative typicality matrix U, absolute typicality matrix T can be used for clustering result analysis.

$$V = \begin{bmatrix} 0.540 & 0.579 & 0.442 & \dots & 0.862 \\ 3.033 & 2.997 & 2.055 & \dots & 2.834 \\ -0.407 & -0.414 & -0.315 & \dots & -0.485 \end{bmatrix}$$

$$U = \begin{bmatrix} 0.051 & 0.006 & 0.942 \\ 0.009 & 0.001 & 0.990 \\ 0.185 & 0.015 & 0.800 \\ \vdots & \vdots & \vdots \\ 0.052 & 0.009 & 0.939 \\ 0.052 & 0.009 & 0.939 \\ 0.035 & 0.005 & 0.961 \end{bmatrix}$$

$$T = \begin{bmatrix} 0.006 & 0.001 & 0.105 \\ 0.001 & 0.000 & 0.101 \\ 0.030 & 0.003 & 0.116 \\ \vdots & \vdots & \vdots \\ 0.007 & 0.001 & 0.106 \\ 0.007 & 0.001 & 0.106 \\ 0.004 & 0.001 & 0.104 \end{bmatrix}$$

3.3. Modified Partition Coefficient Validity Result. Evaluation related to the accuracy of Possibilistic Fuzzy C-Means clustering results on the prevalence of stunting toddlers is determined using the Modified Partition Coefficient (MPC). The MPC value is calculated using equation (13).

$$\text{MPC} = 1 - \frac{3}{3-1} \left(1 - \frac{1}{399} \sum_{k=1}^3 \sum_{i=1}^{399} (\mu_{ik})^2 \right) = 0.704$$

The MPC validity results show a significant accuracy value of 0.704. This value shows that the accuracy is close to 1 and the clustering results used have a good level of precision in grouping data.

3.4. Interpretation of Clustering Results. The result of the relative typicality matrix U is finally represented as the result of the clustering process using the PFCM method as the degree of membership in each district/city in Indonesia in each cluster as shown in Table 4.

TABLE 4. Degree of Membership

No.	Districts/Cities	Cluster 1	Cluster 2	Cluster 3	Clustering Result
1	South Aceh	0.051	0.006	0.942	Cluster 3
2	Southeast Aceh	0.009	0.001	0.990	Cluster 3
3	East Aceh	0.185	0.015	0.800	Cluster 3
⋮	⋮	⋮	⋮	⋮	⋮
149	Kuningan	0.795	0.084	0.121	Cluster 1
150	Cirebon	0.051	0.926	0.023	Cluster 2
⋮	⋮	⋮	⋮	⋮	⋮
397	Intan Jaya	0.052	0.009	0.939	Cluster 3
398	Deiyai	0.052	0.009	0.939	Cluster 3
399	Jayapura City	0.035	0.005	0.961	Cluster 3

The determination of clusters is based on the highest degree of membership of each district/city, by looking at which cluster the district has the most dominant membership value.

TABLE 5. Mean Data of Each Cluster

Cluster	PW-TTD	IMD	EBF-6mo	AdeqSan
1	14.216	12.622	5.986	311.306
2	43.409	36.742	17.704	680.235
3	3.908	3.070	1.674	59.543

Notes:

PW-TTD = Number of Pregnant Women Receiving TTD

IMD = Number of Infants Receiving IMD

EBF-6mo = Number of Infants Exclusive Breastfeeding for 6 months

AdeqSan = Number of Families with Access to Adequate Sanitation Facilities

The final result of the cluster center matrix V in the PFCM process can be used to identify the characteristics of clusters that have high, medium, and low stunting causal factors. This identification can be adjusted to the average value of the original data from each variable presented in Table 5, thus providing a more accurate picture of the pattern and distribution of data in each cluster. From the analysis in Table 5, it can be seen that cluster 2 has the highest average value. This shows that cluster 2 is categorized as an area with a low prevalence of stunting. While cluster 3 has the lowest average value so it is categorized as an area with a high prevalence of stunting. Then cluster 1 is classified as a cluster with stunting-causing factors that are in the medium category.

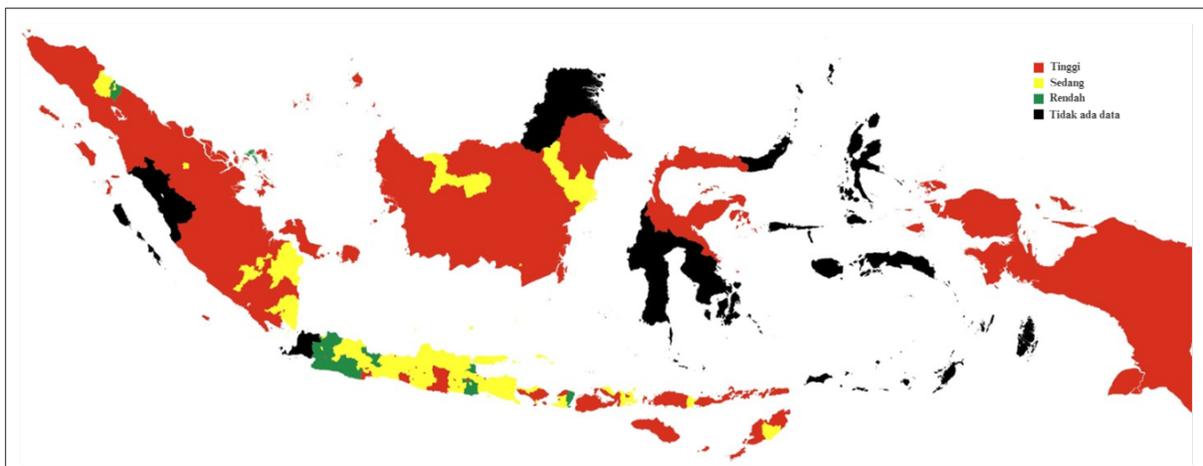


FIGURE 1. Distribution map of stunting prevalence in Indonesia shows that there are still many areas with high stunting rates. This is indicated by the dominance of red color spread in various regions, indicating districts/cities with a high prevalence of stunting

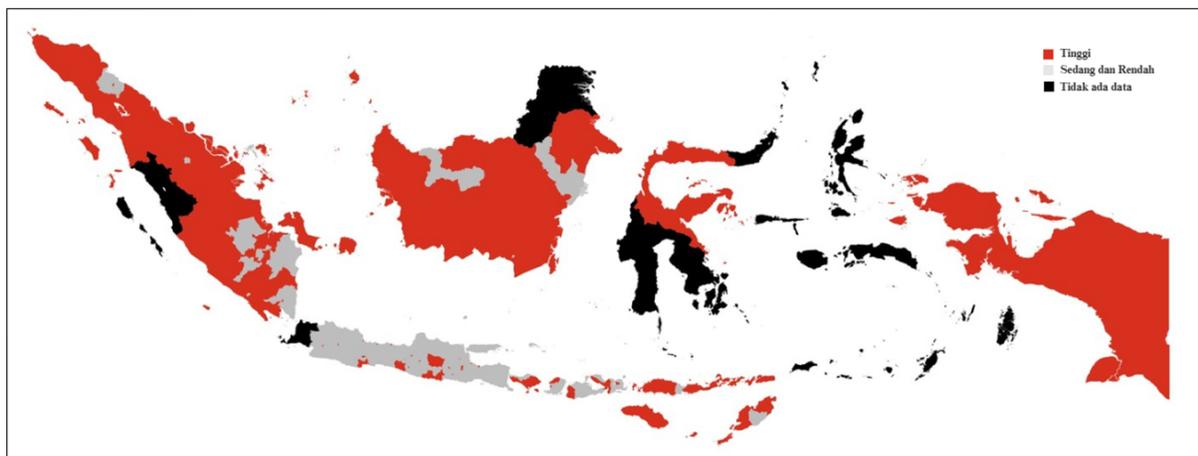


FIGURE 2. Distribution map shows the areas included in the high category cluster. The dominance of red color on the map is the provinces of Aceh, Jambi, Bengkulu, Bangka Belitung Islands, Central Kalimantan, Central Sulawesi, Gorontalo, West Papua, and Papua

Figure 1 shows that there are 22 districts/cities in the low category with details of 14 districts and 8 cities. There are 87 districts/municipalities in the medium category with details of 75 districts and 12 cities. As for the high category, there are 290 districts/cities with details of 178 districts and 112 cities. This indicates that stunting prevalence in Indonesia remains high, as the number of districts/cities in Cluster 3 exceeds those in Clusters 2 and 1.

Figure 2 represents the distribution of districts/cities in cluster 3 in Indonesia based on the results of clustering with factors that cause stunting in the high category. Based on these results, cluster 3 is dominated by 9 provinces, namely Aceh Province, Jambi Province, Bengkulu Province, Bangka Islands Province, Central Kalimantan Province, Central Sulawesi Province, Gorontalo Province, West Papua Province, and Papua Province. The high prevalence of stunting in some of these provinces is caused by various factors specific to each region. In Aceh Province, Gayo Lues District to be precise, the low level of clean and healthy living behavior (PHBS), poverty, suboptimal uptake of POSYANDU, and the inactivity of the Stunting Reduction Acceleration Team (TPPS) are the main factors. According to the Ministry of Health, POSYANDU is an Integrated Service Post which is a community-based health effort that aims to provide integrated health and child development service. In Jambi Province, to be precise in Jambi City, anemia in pregnant women due to low compliance with the consumption of blood supplement tablets is one of the important problems. In Central Sulawesi, specifically in Palu City, low community participation in Posyandu activities worsens immunization efforts and health monitoring of children under five. Meanwhile, in Arfak Mountains District, West Papua, the lack of access to health services, clean water facilities, and supplementary food hampers nutrition improvement efforts. In Papua Province, specifically in Jayapura City, the main challenges are lack of nutrition awareness, availability of nutritious food, and access to clean water, although

Jayapura City has shown a decrease in stunting rates. These challenges highlight the importance of holistic and sustainable interventions to reduce stunting in Indonesia.

4. CONCLUSION

Based on the problem formulation and discussion in this study, it can be concluded that the Possibilistic Fuzzy C-Means (PFCM) clustering method on the prevalence of stunting toddlers with optimal parameters are the weighting rank for FCM (m)=2, the weighting rank for PCM (η)=2, the coefficient $K=1$, the weighting coefficient for the relative typicality matrix U (a)=1, and the weighting coefficient for the absolute typicality matrix T (b)=9. The clustering results show the grouping of stunting prevalence categories in 399 districts/cities in Indonesia, namely 22 districts/cities with low categories, 87 districts/cities with medium categories, and 290 districts/cities with high categories. The accuracy of the clustering results using the Modified Partition Coefficient (MPC) method reached 0.704. It indicates clustering model is close to the optimal value and is included in the good criteria. Hence, the result is applicable in further policy for handling stunting in Indonesia.

Based on the results of this study, clusters with high stunting prevalence are dominated by nine provinces, namely Aceh, Jambi, Bengkulu, Bangka Belitung Islands, Central Kalimantan, Central Sulawesi, Gorontalo, West Papua, and Papua. These areas require greater government attention to reduce stunting prevalence and achieve the annual reduction target. The findings of this study can serve as a basis for formulating more targeted and effective policies to enhance local government performance in addressing stunting. Future research can explore different clustering and validity methods to compare the accuracy of different clustering methods. Additionally, different factors contributing to stunting can be explored to identify more accurate determinants.

Competing interests: The authors declare that there is no conflict of interest regarding the publication of this paper.

REFERENCES

- [1] J. Han, M. Kamber, J. Pei, Data Mining: Concepts and Techniques, 3rd ed, Elsevier, Amsterdam, 2012.
- [2] A.K. Jain, M.N. Murty, P.J. Flynn, Data Clustering, ACM Comput. Surv. 31 (1999), 264–323. <https://doi.org/10.1145/331499.331504>.
- [3] B.S. Everitt, S. Landau, M. Leese, D. Stahl, Cluster Analysis, 5th ed., John Wiley & Sons, 2011.
- [4] D.L. Rahakbauw, V.Y.I. Ilwaru, M.H. Hahury, Implementasi Fuzzy C-Means Clustering Dalam Penentuan Beasiswa, BAREKENG: J. Ilmu Mat. Dan Ter. 11 (2017), 1–12. <https://doi.org/10.30598/barekengvol11iss1pp1-12>.
- [5] N. Pal, K. Pal, J. Keller, J. Bezdek, A Possibilistic Fuzzy C-Means Clustering Algorithm, IEEE Trans. Fuzzy Syst. 13 (2005), 517–530. <https://doi.org/10.1109/TFUZZ.2004.840099>.
- [6] J. Chen, H. Zhang, D. Pi, M. Kantardzic, Q. Yin, et al., A Weight Possibilistic Fuzzy C-Means Clustering Algorithm, Sci. Program. 2021 (2021), 9965813. <https://doi.org/10.1155/2021/9965813>.

- [7] Y. Chen, S. Zhou, Revisiting Possibilistic Fuzzy C-Means Clustering Using the Majorization-Minimization Method, *Entropy* 26 (2024), 670. <https://doi.org/10.3390/e26080670>.
- [8] G.N.S. Putri, D. Ispriyanti, T. Widiharih, Implementasi Algoritma Fuzzy C-Means dan Fuzzy Possibilistics C-Means Untuk Klasterisasi Data Tweets Pada Akun Twitter Tokopedia, *J. Gaussian* 11 (2022), 86–98. <https://doi.org/10.14710/j.gauss.v11i1.33996>.
- [9] R.N. Dave, Validating Fuzzy Partitions Obtained Through C-Shells Clustering, *Pattern Recognit. Lett.* 17 (1996), 613–623. [https://doi.org/10.1016/0167-8655\(96\)00026-8](https://doi.org/10.1016/0167-8655(96)00026-8).
- [10] A. Suleman, A New Perspective of Modified Partition Coefficient, *Pattern Recognit. Lett.* 56 (2015), 1–6. <https://doi.org/10.1016/j.patrec.2015.01.008>.
- [11] B. Simhachalam, G. Ganesan, Possibilistic Fuzzy C-Means Clustering on Medical Diagnostic Systems, in: 2014 International Conference on Contemporary Computing and Informatics (IC3I), IEEE, 2014: pp. 1125–1129. <https://doi.org/10.1109/IC3I.2014.7019729>.
- [12] V. Kalist, P. Ganesan, B. Sathish, J.M.M. Jenitha, K. Basha.shaik, Possibilistic-Fuzzy C-Means Clustering Approach for the Segmentation of Satellite Images in HSL Color Space, *Procedia Comput. Sci.* 57 (2015), 49–56. <https://doi.org/10.1016/j.procs.2015.07.364>.
- [13] M. Lestari, D. Kartini, I. Budiman, M.R. Faisal, M. Muliadi, Comparison of Industrial Business Grouping Using Fuzzy C-Means and Fuzzy Possibilistic C-Means Methods, *Telematika* 16 (2023), 91–102. <https://doi.org/10.35671/telematika.v16i2.2548>.
- [14] M.Z. Rodriguez, C.H. Comin, D. Casanova, O.M. Bruno, D.R. Amancio, et al., Clustering Algorithms: A Comparative Approach, *PLOS ONE* 14 (2019), e0210236. <https://doi.org/10.1371/journal.pone.0210236>.
- [15] Y. Gu, T. Ni, Y. Jiang, Deep Possibilistic C-Means Clustering Algorithm on Medical Datasets, *Comput. Math. Methods Med.* 2022 (2022), 3469979. <https://doi.org/10.1155/2022/3469979>.
- [16] World Health Organization, Stunting in a Nutshell, (2015). <https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell>.
- [17] Kementerian Kesehatan Republik Indonesia, Hasil Survei Status Gizi Indonesia (SSGI) 2022, (2022). <https://repository.badankebijakan.kemkes.go.id/id/eprint/4855/3/Buku%20Saku%20SSGI%202022%20rev%20270123%200K.pdf>.
- [18] K. Nida, M.N. Hayati, R. Goejantoro, Implementasi Metode Fuzzy Possibilistic C-Means Pada Pengelompokan Provinsi Di Indonesia Berdasarkan Data Jumlah Kejadian Dan Dampak Bencana Banjir, *J. Math. Educ. Sci.* 7 (2024), 33–42. <https://doi.org/10.32665/james.v7i1.1919>.
- [19] F.N.M. Falah, Y.R.W. Utami, W.L.Y. Saptomo, Penerapan Fuzzy C-Means Dalam Penentuan Kelas Siswa Di Sekolah Menengah Kejuruan, *J. Teknol. Inf. Komun. (TIKomsin)* 12 (2024), 41. <https://doi.org/10.30646/tikomsin.v12i1.743>.
- [20] A. Keller, F. Klawonn, Fuzzy Clustering With Weighting of Data Variables, *Int. J. Uncertain. Fuzziness Knowl.-Based Syst.* 08 (2000), 735–746. <https://doi.org/10.1142/S0218488500000538>.