Eksakta : Berkala Ilmiah Bidang MIPA
VOLUME 27 NO 05 2026, pp 1022-1038
ISSN : Print 1411-3724 — Online 2549-7464
DOI : https://doi.org/10.24036/eksakta/vol27-iss05/773

Eksakta
Berkala Ilmiah Bidang MIPA
http://www.eksakta.ppj.unp.ac.id/index.php/eksakta

Article

Comparative Analysis of Fuzzy Geographically Weighted Clustering with Flower Pollination and Chaotic Optimization for Flood Impacts

Article Info

Article history *:*
Received October 12, 2026
Revised October 24, 2026
Accepted October 26, 2026
Published October 30, 2026

Keywords *:*
FGWC,
FPA,
chaotic optimization,
flood impacts,
SDG 11

Ahmad Sayyif Zulfi Khoirudin[1], A'yunin Sofro[2*]

[1]Department of Mathematics, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Surabaya, Indonesia
[2]Department of Actuarial Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Surabaya, Surabaya, Indonesia

Abstract. Natural and non-natural disasters can cause socioeconomic losses and infrastructure damage that hinder sustainable development. Clustering the characteristics of flood impacts across regions can support mitigation planning. However, the classic Fuzzy Geographically Weighted Clustering (FGWC) method can be influenced by the initialization of cluster centers. This study aims to compare the performance of FGWC, FGWC-Flower Pollination Algorithm (FGWC-FPA), and Chaotic Flower Pollination Algorithm–FGWC (CFPA-FGWC) in clustering flood impacts across 38 regencies/cities in East Java using 10 indicators for the year 2025. The evaluation was conducted through 4,400 simulations using eight randomization schemes and six internal validation indices. The results show that CFPA-FGWC with the Circular chaotic map performed best on five of the six indices, while FGWC-FPA produced a slightly lower Classification Entropy value. Two clusters with distinct characteristics emerged. Cluster 1 was dominated by displacement and housing damage, while Cluster 2 exhibited higher frequencies of flooding, fatalities and missing persons, and bridge damage. These findings can serve as a basis for determining aid priorities, evacuation preparedness, and the strengthening of mitigation infrastructure. This study supports SDG 11 and SDG 13 by providing information for disaster risk reduction and enhancing regional resilience.

This is an open acces article under the CC-BY license.

This is an open access article distributed under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2026 by author.

Corresponding Author :
A'yunin Sofro
Department of Actuarial Science, Faculty of Mathematics and Natural Sciences,
Universitas Negeri Surabaya, Surabaya, Indonesia
Email : ayuninsofro@unesa.ac.id

1. Introduction

Floods have varying impacts across regions, ranging from loss of life and displacement to damage to settlements and infrastructure. These differences are influenced by regional characteristics and communities' exposure to disasters [1-4]. Previous research has shown that mapping patterns of vulnerability and the distribution of flood impacts can help identify groups and areas that require priority for mitigation efforts [5-9]. Analysis of impact distribution is also necessary to understand regional variations in risk [8]. These efforts support disaster risk reduction in line with SDG 11 and SDG 13.

Clustering analysis can group regions based on similarities in disaster impact characteristics. The Fuzzy Geographically Weighted Clustering (FGWC) approach considers fuzzy membership degrees and geographic components in cluster formation [10-12]. A region-based fuzzy approach has been applied to cluster food production in Indonesia [13]. In disaster studies, spatial analysis is also used to examine seismic characteristics and map seismic hazards [14-16]. Meanwhile, clustering methods and probabilistic approaches have been examined to support building damage analysis in the context of disaster mitigation [17].

Previous research on flood vulnerability clustering in East Java has utilized DBSCAN and K-Means++ [18]. However, the quality of the clustering results is also influenced by the selection of cluster centers and the optimization mechanism used [19]. The Flower Pollination Algorithm (FPA) offers a global search mechanism based on Lévy flights and a local search mechanism to refine candidate solutions [20-23]. The development of metaheuristic algorithms for clustering continues to be studied to improve search quality and convergence [19].

One development in FPA is the integration of chaotic maps to modify the search mechanism during optimization. This approach has been applied to function optimization and engineering problems through CFPA as well as chaos-based variations of FPA [24-25]. Mathematically, chaotic maps generate sequences through deterministic nonlinear mappings whose behavior can be sensitive to initial conditions. In optimization, these sequences can be used to vary the search steps or the process of updating candidate solutions. Their use has the potential to increase search diversity, but does not guarantee the avoidance of premature convergence for every problem.

The effect of chaotic maps on FGWC optimization needs to be evaluated empirically. Evaluating clustering results requires validation indices, as each index has different characteristics and sensitivities to data structure [26-29]. In fuzzy clustering, various indices have been developed to assess the quality of partitions based on membership characteristics and cluster structure [30-31]. Therefore, the use of multiple indices in this study is expected to provide a more comprehensive assessment of the optimization results.

This study compares the classical FGWC, the FGWC optimized using FPA (FGWC-FPA), and the Chaotic Flower Pollination Algorithm–FGWC (CFPA-FGWC) to classify the impacts of flooding in 38 regencies and cities in East Java based on data from 2025. Ten active indicators were selected to represent flood occurrences, impacts on people, and damage to settlements and infrastructure. This selection was based on the availability of impact indicators in BNPB data and their relevance to the social and physical dimensions of flood impacts. This scope aligns with previous studies that emphasize the importance of social and physical dimensions in flood vulnerability mapping [2], [3], [32]. Indicators for educational facilities, health facilities, and offices were not used because all observations had a value of zero, thus providing no variation to distinguish regional characteristics.

The contribution of this research lies in a comprehensive evaluation of FGWC optimization performance using eight randomization schemes: six chaotic maps Chebyshev, Circular, Dyadic, Kent, Logistic, and Sine and two random distributions, Normal and Uniform. The evaluation covered 44 configurations with 100 iterations per configuration, resulting in 4,400 simulation results. Cluster performance was assessed using six internal validation indices: Partition Coefficient (PC), Classification Entropy (CE), Partition Index (SC), Separation Index (SI), I-Index (IFV), and Kwon. With this design, this study not only compares clustering methods but also tests the consistency of their performance across variations in randomization mechanisms. The results are expected to provide an analytical basis for identifying the characteristics of flood impacts and support the prioritization of mitigation efforts at the regional level.

2. Experimental Section

This section presents the methodological framework used to compare the performance of Fuzzy Geographically Weighted Clustering (FGWC), FGWC combined with the Flower Pollination Algorithm (FGWC-FPA), and the Chaotic Flower Pollination Algorithm-FGWC (CFPA-FGWC) in clustering regencies and cities in East Java based on flood impact indicators for the year 2025. The research procedure includes data preparation and variable selection, data normalization, spatial component formation, the application of FGWC-based clustering methods, and evaluation using the internal cluster validity index. The entire procedure is designed to produce a systematic and reproducible comparison of clustering methods.

2.1. Research Data and Variables

This study uses data on the impact of floods in 38 regencies and cities in East Java Province in 2025, obtained from the National Disaster Management Agency (BNPB). The data were used to identify the characteristics of flood impacts in each region, with each regency/city represented as a single observation, resulting in $n = 38$ observations.

The variables used consist of 10 active indicators, namely the number of incidents (X1), fatalities and missing persons (X2), injured or sick victims (X3), displaced and affected victims (X4), severely damaged homes (X5), moderately damaged homes (X6), slightly damaged homes (X7), flooded homes (X8), places of worship (X9), and bridges (X10). Three other indicators available in the initial data educational facilities, health facilities, and affected offices were not used because all values were recorded as zero in 38 regencies/cities, indicating that there was no physical damage to these facilities. Consequently, these three indicators showed no variation across regions and did not provide distinguishing information for the clustering process. Thus, each object j is represented by a vector

$$x_j = (x_{j1}, x_{j2}, \dots, x_{j10}), \qquad j = 1,2,3, \dots ,38$$

In addition to flood impact indicators, the geographic coordinates and population sizes of each regency/city were used to construct the spatial component. The geographic coordinates were obtained from a dataset of regency/city coordinates in Indonesia, while population data were obtained from the East Java Provincial Central Statistics Agency. Geographic coordinates are used to calculate distances between locations, while population data is used as a spatial weighting component in the FGWC, ensuring that the clustering accounts for spatial relationships between regions [10], [12].

2.2. Data Preprocessing

Preprocessing was performed to generate data that could be used consistently in the clustering process. Since the ten indicators had different units and value ranges, normalization was performed using the Min-Max method so that each indicator would be on a comparable scale [33]. For the r-th indicator, the observation value x_{jr} was transformed into

$$x_{jr}^{*} = \frac{x_{jr} - min_j(x_{jr})}{max_j(x_{jr}) - min_j(x_{jr})}$$

this yields $x_{jr}^{*} \in [0,1]$. The normalized data are then used as input in the FGWC, FGWC-FPA, and CFPA-FGWC processes.

Normalization is necessary because differences in units and ranges between indicators can affect the clustering process if the data are used on their original scales [33]. After normalization, the resulting data matrix is used throughout the clustering process so that each indicator has a uniform range.

2.3. Formation of Spatial Components

The spatial components are derived from the latitude and longitude coordinates of each regency/city. The geographic distance between locations is calculated using the Haversine formula to determine the Earth's surface distance between two locations. For two locations i and j, with latitudes ϕ_i,ϕ_j and longitudes λ_i, λ_j is defined

$$h_{ij} = \sin^2(\frac{\phi_i - \phi_j}{2}) + \cos(\phi_i)\cos(\phi_j)\sin^2(\frac{\lambda_i - \lambda_j}{2})$$

so that the geographic distance is

$$d_{ij} = 2R\ arcsin(\sqrt{h_{ij}})$$

where R represents the Earth's radius. This geographic distance matrix is then used to establish spatial relationships between regions.

In FGWC, spatial weights take into account both geographic distance and population size. The spatial weights used are

$$w_{ij} = \frac{(p_i p_j)^b}{d_{ij}^a}$$

where p_i and p_j represent the populations of objects i and j, respectively, while a and b are the parameters for distance and population effects [10]. In this study, $a = b = 1$ was used, in accordance with the FGWC implementation in the algorithm employed. This spatial matrix was then used to update the membership degrees in FGWC. The distance in the distance equation constitutes the spatial component, while the Euclidean distance is used to calculate the distance between data objects and cluster centers during the clustering process [10].

2.4. Fuzzy Geographically Weighted Clustering (FGWC)

FGWC is a fuzzy clustering method that incorporates geographic information into cluster formation by updating membership degrees based on spatial relationships between objects [10]. Suppose there are c clusters and n objects, then u_{ij} represents the membership degree of object j in cluster i, where $0 \le u_{ij} \le 1$ such that $\sum_{i=1}^{c} u_{ij} = 1$. The center of the i-th cluster is calculated based on the membership degrees as

$$v_i = \frac{\sum_{j=1}^{n} u_{ij}^m \boldsymbol{x_j}}{\sum_{j=1}^{n} u_{ij}^m}$$

where m is the fuzziness parameter. In this study, $m = 2$ is used. The objective function used is

$$J = \sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \|x_j - v_i\|^2$$

The spatial component is then used to update the membership degrees. With α as the previous membership weight and $\beta = 1 - \alpha$ as the spatial component weight, the membership update is expressed as

$$u'_{ki} = \alpha u_{ki} + \frac{\beta}{A_i} \sum_{j=1}^{n} w_{ij} u_{kj}$$

where $A_i = \sum_{k=1}^{c} \sum_{j=1}^{n} w_{ij}\, u_{kj}$.

This equation shows that an object's membership is influenced by its proximity to the cluster center and by membership information from other objects with which it has a spatial relationship. This formulation and implementation follow the FGWC mechanism used in previous studies as well as the implementation in the naspaclust package [10] .

In this study, we used $\alpha = 0.7, \beta = 0.3, a = b = 1$, a convergence tolerance of 10^{-5} and a maximum of 100 iterations. These parameters follow the FGWC implementation configuration in naspaclust. This study did not perform a sensitivity analysis on the parameters used. Therefore, the results are limited to a comparison of the methods' performance based on the specified parameter configurations.

2.5. Flower Pollination Algorithm (FPA)

To optimize the cluster center search process in FGWC, the Flower Pollination Algorithm (FPA) is used. FPA is a metaheuristic optimization algorithm inspired by the flower pollination process and employs both global and local pollination mechanisms [20]. FPA has also been developed to address optimization problems in other fields, demonstrating that the mechanism of this algorithm can be adapted to the characteristics of the problems at hand [21], [34].

In global updating, the i-th solution is updated based on the global optimal solution as

$$x_i^{t+1} = x_i^t + \gamma L(\lambda)(g^* - x_i^t)$$

where g^* is the optimal solution at the t-th iteration, γ is a scaling factor, and $L(\lambda)$ is a step based on the Lévy distribution [20]. Meanwhile, local crossover is performed using

$$x_i^{t+1} = x_i^t + \epsilon(x_j^t - x_k^t)$$

where x_j^t and x_k^t are two solutions selected from the population, and ϵ is the randomization value.

The global and local mechanisms are selected based on the replacement probability p. In this study, 10 flower agents were used with parameters $p = 0.8, \gamma = 1$, and $\lambda = 1.5$, while the optimization process was aimed at obtaining cluster centers that yield a lower FGWC objective function value. These values are treated as a consistent experimental setup for comparing methods, rather than as the results of a search for the optimal parameters for this flood dataset. Using the same setup helps ensure that differences in the results primarily reflect the optimization methods and their randomization schemes.

2.6. Chaotic Flower Pollination Algorithm (CFPA)

CFPA is an extension of FPA that uses chaotic maps to generate random values during the optimization process. The use of chaotic maps in FPA has been developed to improve the exploration mechanism and solution-finding characteristics of optimization algorithms [24-25]. In this study, six chaotic maps were used: Chebyshev, Circular, Dyadic, Kent, Logistic, and Sine.

In addition to randomization based on chaotic maps, the naspaclust implementation provides Normal and Uniform distributions as non-chaotic randomization mechanisms. The Normal distribution is defined as $\epsilon_t \sim N(0,1)$, while the Uniform distribution is defined as $\epsilon_t \sim U(-1,1)$. Both distributions are used as comparison mechanisms against the six chaotic maps; thus, the Normal and Uniform distributions are not categorized as chaotic maps but rather as randomization schemes in the optimization process.

The six chaotic maps used in this study are defined as follows.

a) Chebyshev map: $\epsilon_{t+1} = \cos(t \cos^{-1}(\epsilon_t))$
b) Circular map: $\epsilon_{t+1} = \epsilon_t + 0.2 - (0.5 - 2\pi)\sin(2\pi\epsilon_t)$
c) Dyadic map: $\epsilon_{t+1} = (2\epsilon_t)\ mod\ 1$
d) Kent map: $\epsilon_{t+1} = \begin{cases} \frac{\epsilon_t}{m_c}, 0 < \epsilon_t \leq m_c \\ \frac{1-\epsilon_t}{1-m_c}, m_c < \epsilon_t < 1 \end{cases}$
e) Logistic map: $\epsilon_{t+1} = r\epsilon_t(1 - \epsilon_t)$
f) Sine map: $\epsilon_{t+1} = \frac{m_c}{4}\sin(\pi\epsilon_t)$

where $m_c = 0.7$ and $r = 4$.

The mapping parameters and method for generating random values follow the CFPA implementation in the naspaclust package. The difference between FGWC-FPA and CFPA-FGWC lies in the randomization mechanism during optimization, while the spatial structure of FGWC and the objective function used are retained. Previous research on FPA and CFPA optimization provides a foundation for comparing these randomization mechanisms [20], [24]. To clarify the integration of CFPA optimization and FGWC spatial membership updates, the algorithm's computational flow is presented in Figure 1.

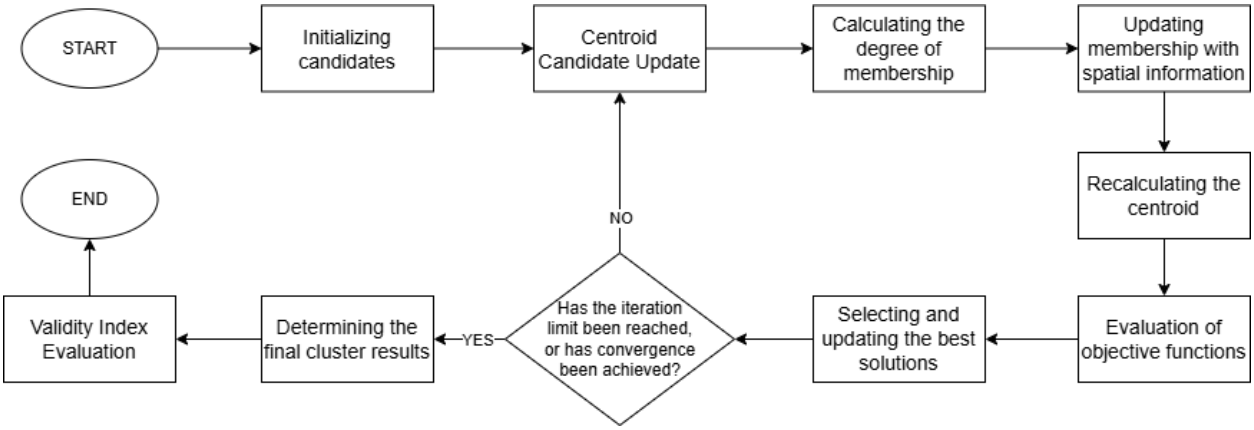


Figure 1. CFPA Integration Workflow and FGWC Spatial Membership Update

Based on Figure 1, optimization is performed by updating cluster center candidates, calculating membership degrees, and adjusting membership while taking spatial information into account. The cluster centers are then updated, and each candidate is evaluated using the objective function. This process is repeated until the termination condition is met, after which the final cluster results are determined and the internal validity index is evaluated.

2.7. Internal Validation

The performance of the clustering results was evaluated using six internal validation indices: Partition Coefficient (PC), Classification Entropy (CE), Partition Index (SC), Separation Index (SI), I-Index (IFV), and Kwon. Internal validation was used because the clustering results lacked external class labels to serve as ground truth. The use of multiple indices is necessary because each index evaluates different partition characteristics, such as membership strength, compactness, inter-cluster separation, and uncertainty. Studies on cluster validation show that evaluation results can vary depending on the characteristics and mathematical properties of the indices used [26], [28], [29], [35]. In fuzzy clustering, the evaluation must also consider the degree of membership of each object in a cluster as

well as the characteristics of the resulting fuzzy partition [27], [30], [36], [37]. The Partition Coefficient (PC) measures the degree of membership of an object in a cluster and is calculated as

$$PC = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2$$

where u_{ij} represents the degree of membership of the j-th object in the i-th cluster, c is the number of clusters, and n is the number of objects. A higher PC value indicates a more distinct fuzzy partition [27], [30]. Classification Entropy (CE) measures the uncertainty in fuzzy membership and is calculated as

$$CE = -\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\, ln(u_{ij})$$

A lower CE value indicates a lower level of membership uncertainty [27], [28], [29]. The Partition Index (SC) measures the compactness of objects relative to the cluster centers and the separation between cluster centers as

$$SC = \sum_{i=1}^{c} \frac{\sum_{j=1}^{n} u_{ij}^m \left\|x_j - v_i\right\|^2}{\sum_{k=1}^{c} N_k \|v_i - v_k\|^2}$$

where N_k denotes the number of objects included in the k-th cluster based on hard assignment, and m is the fuzzification parameter, which in this study is set to $m = 2$. A lower SC value indicates a better partition [27], [29], [35]. The Separation Index (SI) evaluates the compactness of objects by considering the minimum distance between cluster centers and is calculated as

$$SI = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2 \left\|x_j - v_i\right\|^2}{n \min\limits_{i\neq k} \|v_i - v_k\|^2}$$

A lower SI value indicates better clustering results [27], [29], [35], [38]. The I-Index (IFV) takes into account the dispersion of objects relative to the cluster centers, the separation between centers, and the strength of membership. The components of average dispersion and maximum distance between centers are defined as

$$\sigma_D = \frac{1}{nc}\sum_{i=1}^{c}\sum_{j=1}^{n} \left\|x_j - v_i\right\|^2$$

$$SD_{max} = \max_{i\neq k} \|v_i - v_k\|^2$$

Based on these two components, the IFV is calculated as

$$IFV = \frac{SD_{max}}{\sigma_D}\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2 \left[\log_2 c - \frac{1}{n}\sum_{k=1}^{n} \log_2 u_{ik}\right]^2$$

A higher IFV value indicates better partitioning [27], [29], [30]. The Kwon index considers the compactness of objects relative to the cluster center and the dispersion among cluster centers. The index is expressed as

$$V_K = \frac{\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2 \left\|x_j - v_i\right\|^2 + \frac{1}{c}\sum_{i=1}^{c}\left[\sum_{l=1}^{p}(v_{il} - \bar{x}_l)\right]^2}{\min\limits_{i\neq k} \|v_i - v_k\|^2}$$

where $\bar{x}_l = \frac{1}{n}\sum_{j=1}^{n} x_{jl}$. A lower Kwon value indicates a better partition [31].

The use of multiple indices allows clustering results to be evaluated based on a number of different partition characteristics. Research on validation indices shows that not all indices exhibit the same evaluation behavior; therefore, the choice of evaluation metric must be tailored to the partition characteristics and the objectives of the analysis [35], [36], [37], [39]. Research on the development of

fuzzy indices also indicates that partition quality evaluation continues to evolve to address diverse data characteristics and cluster structures [40-42]. Therefore, in this study, better partition quality is indicated by higher PC and IFV values and lower CE, SC, SI, and Kwon values.

Each algorithm configuration was simulated 100 times for the number of clusters $c = 2, 3, 4,$ and 5. The simulations were repeated to evaluate the stability and consistency of each configuration's performance. All implementations were carried out using the R programming language with the naspaclust package, which provides the FGWC, FGWC-FPA, and CFPA-FGWC methods, as well as the validation index functions used in this study.

3. Results and Discussion

The research results are presented based on an evaluation of the performance of FGWC, FGWC-FPA, and CFPA-FGWC under various clustering configurations. The evaluation was conducted using six internal validation indices: Partition Coefficient (PC), Classification Entropy (CE), Partition Index (SC), Separation Index (SI), I-Index (IFV), and Kwon. The evaluation results were used to determine the final configuration for each method and subsequently to analyze regional characteristics based on the clustering results

3.1. Experimental Results and Internal Validation

The experiment was conducted 100 times for each configuration so that the evaluation results would not depend solely on a single initialization process. In total, the experiment covered 44 configurations and produced 4,400 simulation results, consisting of 400 FGWC simulation results, 800 FGWC-FPA simulation results, and 3,200 CFPA-FGWC simulation results. For FGWC-FPA, two randomization schemes were used Normal and Uniform while CFPA-FGWC utilized these two schemes along with six chaotic maps: Chebyshev, Circular, Dyadic, Kent, Logistic, and Sine. The average values from the 100 repetitions were used to determine the best configuration for each validation index, as presented in Table 1.

Table 1. Best Configuration Based on Each Validation Index

Method	Index	Configuration	Top grades
FGWC	PC	$c = 2$	0.6223
	CE	$c = 2$	0.5598
	SC	$c = 5$	0.7808
	SI	$c = 4$	8549.09
	IFV	$c = 5$	4.1566
	Kwon	$c = 4$	385133.14
FGWC-FPA	PC	$c = 2$, Normal	0.6656
	CE	$c = 2$, Normal	0.5098
	SC	$c = 5$, Uniform	0.3754
	SI	$c = 2$, Normal	1.0290
	IFV	$c = 5$, Uniform	25.8791
	Kwon	$c = 2$, Normal	41.6860
CFPA-FGWC	PC	$c = 2$, Circular	0.6746
	CE	$c = 2$, Circular	0.4988
	SC	$c = 5$, Sine	0.3462
	SI	$c = 2$, Circular	0.9625
	IFV	$c = 5$, Sine	27.6763
	Kwon	$c = 2$, Circular	39.1497

Based on Table 1, the configuration $c = 2$ achieved optimal values on four of the six indices for both FGWC-FPA and CFPA-FGWC. In FGWC-FPA, this configuration used Normal randomization, while in CFPA-FGWC it used a Circular chaotic map. In contrast, SC and IFV in both methods achieved optimal values at $c = 5$. These differences indicate that each index has distinct evaluation characteristics. SC considers the compactness of objects within clusters and the separation between clusters, whereas IFV evaluates cluster structure based on the distribution of objects and fuzzy membership degrees. Increasing the number of clusters can affect group compactness, the distance between cluster centers, and the distribution of membership degrees; thus, SC and IFV may yield optimal values at different numbers of clusters than other indices. This is consistent with Wang et al. [27] as well as comparative studies and reviews of clustering validity indices [28], [29], [35], which show that validity measures can assess partition structures from different perspectives. Therefore, there is no single configuration that outperforms others based on all validity indices; thus, determining the final configuration requires considering several validation measures simultaneously.

In classical FGWC, the optimal values are distributed across $c = 2, c = 4,$ and $c = 5$. Therefore, these three configurations need to be further compared based on member structure and the characteristics of the resulting clusters. The results of this analysis are used to determine the final FGWC configuration.

3.2. Selection of the Final Clustering Configuration

The final configuration was selected based on the results of 100 iterations, the patterns of the six validation indices, and the resulting cluster structure. Based on this comprehensive evaluation, the final number of clusters was set to $c = 2$ for all three methods, with FGWC using the classical configuration, FGWC-FPA using Normal randomization, and CFPA-FGWC using the Circular chaotic map. Using the same number of clusters allows for a fair comparison of the three methods' performance across equivalent group sizes.

In the classical FGWC, the selection of $c = 2$ required further examination because the optimal index values were distributed across $c = 2, c = 4,$ and $c = 5$. At $c = 2$, two clusters were formed, consisting of 8 and 30 regions. Meanwhile, $c = 4$ produces clusters of 7, 22, 1, and 8 regions, while $c = 5$ produces clusters of 1, 20, 10, and 7 regions with one cluster label left unassigned. This structure indicates that increasing the number of clusters results in a more fragmented partition and the formation of single-member clusters. Since the choice of the number of clusters must take into account both the index value and the interpretability of the partition structure [19], [27], [35], $c = 2$ was selected because it produces a simpler and more interpretable structure, with all clusters having an adequate number of members.

For FGWC-FPA, the $c = 2$ configuration with Normal randomization was selected because it achieved optimal values on four of the six indices: PC, CE, SI, and Kwon. For CFPA-FGWC, the $c = 2$ configuration with a Circular chaotic map also achieved optimal values on the same four indices. Thus, these two configurations provide more consistent validation support compared to the other configurations.

Table 2. Final Configuration of the Three Clustering Methods

Method	Number of Clusters	Randomization
FGWC	2	-
FGWC-FPA	2	Normal
CFPA-FGWC	2	Circular

The configuration in Table 2 was then used to obtain the final clustering results, including cluster labels, fuzzy membership, centroids, and the characteristics of each cluster. Thus, the results of 100

simulations were used as the basis for determining the clustering configuration, while the selected configuration was used to obtain the clustering results, which were then analyzed.

3.3. Performance Comparison of Three Methods

A performance comparison of the three methods was conducted using the validation results for the final configuration of each method. All six indices were used simultaneously to compare the quality of the partitions generated by FGWC, FGWC-FPA, and CFPA-FGWC. The validation results for the final configurations are presented in Table 3.

Table 3. Results of the Final Configuration Validation for the Three Methods

Method	PC	CE	SC	SI	IFV	Kwon
FGWC	0.6329	0.5482	4.7115	1.3945	0.7448	56.0514
FGWC-FPA	0.6560	**0.5210**	3.7698	1.1041	0.9814	44.4429
CFPA-FGWC	**0.6577**	0.5214	**3.3564**	**0.9815**	**1.0984**	**39.4252**

Based on Table 3, CFPA-FGWC yielded higher PC and IFV values than FGWC and FGWC-FPA. This method also produced lower SC, SI, and Kwon values. Meanwhile, FGWC-FPA produced a CE value of 0.5210, which was slightly lower than that of CFPA-FGWC at 0.5214. Thus, in clustering using the selected configuration, CFPA-FGWC showed better values on five of the six indices, while FGWC-FPA had a slightly lower CE value.

These differences indicate that the application of optimization can affect the quality of the resulting partitions. These findings are relevant to the study by Nasution et al. [10] on the application of FGWC optimized using CFPA in flood vulnerability analysis, as well as the studies by Daqaq et al. [24] and Ozsoydan et al. [25] on the development of FPA using chaotic maps for optimization. However, the advantages of CFPA-FGWC do not apply to all indices; therefore, performance evaluations must still consider several validity measures.

In addition to cluster quality, performance improvements must also be considered in conjunction with the required computational costs. Therefore, we compared the average execution time and the number of iterations for the three methods, as shown in Table 4.

Table 4. Comparison of Computation Time and Iterations

Method	Configuration	Average Time (s)	Average Iteration
FGWC	$c = 2$	0.0409	19.51
FGWC-FPA	$c = 2$, Normal	0.2060	22.43
CFPA-FGWC	$c = 2$, Circular	0.3326	24.62

Based on Table 4, FGWC has the lowest average execution time, at 0.0409 seconds with an average of 19.51 iterations. FGWC-FPA takes 0.2060 seconds with an average of 22.43 iterations, while CFPA-FGWC takes 0.3326 seconds with an average of 24.62 iterations. This indicates that these optimization methods require more computation time and a greater number of iterations compared to the classical FGWC. Nevertheless, the execution times of all three methods remain under one second in the tested configurations. Considering these validation results and computational costs, CFPA-FGWC provides a viable alternative for analyzing flood impacts across 38 regencies and cities in East Java.

Based on this validation pattern, CFPA-FGWC with a circular chaotic map was used as the method for further analysis of the clustering results. The clustering results from this configuration were then used to analyze regional distribution, fuzzy membership, and cluster characteristics.

3.4. Clustering Results and Spatial Distribution

After the final configuration was established, the CFPA-FGWC results were used to identify groups of regencies and cities based on the similarity of flood impact patterns, taking spatial components into account. The clustering results yielded two clusters: 30 regions in Cluster 1 and 8 regions in Cluster 2. The list of regions in each cluster is presented in Table 5.

Table 5. List of Regencies/Cities Based on the Results of CFPA-FGWC Clustering

Cluster	Regency/City
Cluster 1	Pacitan, Ponorogo, Blitar, Kediri, Lumajang, Banyuwangi, Bondowoso, Situbondo, Pasuruan, Sidoarjo, Mojokerto, Jombang, Nganjuk, Madiun, Magetan, Tuban, Lamongan, Gresik, Bangkalan, Sampang, Pamekasan, Kota Kediri, Kota Blitar, Kota Malang, Kota Probolinggo, Kota Pasuruan, Kota Mojokerto, Kota Madiun, Kota Surabaya, Kota Batu
Cluster 2	Trenggalek, Tulungagung, Malang, Jember, Probolinggo, Ngawi, Bojonegoro, Sumenep

Based on Table 5, Cluster 1 consists of 30 regions, while Cluster 2 consists of 8 regions. The members of Cluster 2 are not concentrated in a single contiguous administrative region but are spread across several parts of East Java. This indicates that the clustering results do not solely form groups based on geographic proximity but rather result from a combination of similarities in flood impact indicators and spatial relationships between regions. This interpretation is consistent with research that combines clustering and geospatial analysis to understand flood patterns, as well as the fuzzy spatial clustering approach to flood vulnerability mapping [6], [10]. The spatial distribution of the CFPA-FGWC results is further illustrated in Figure 2 to show the geographic locations of each cluster member.

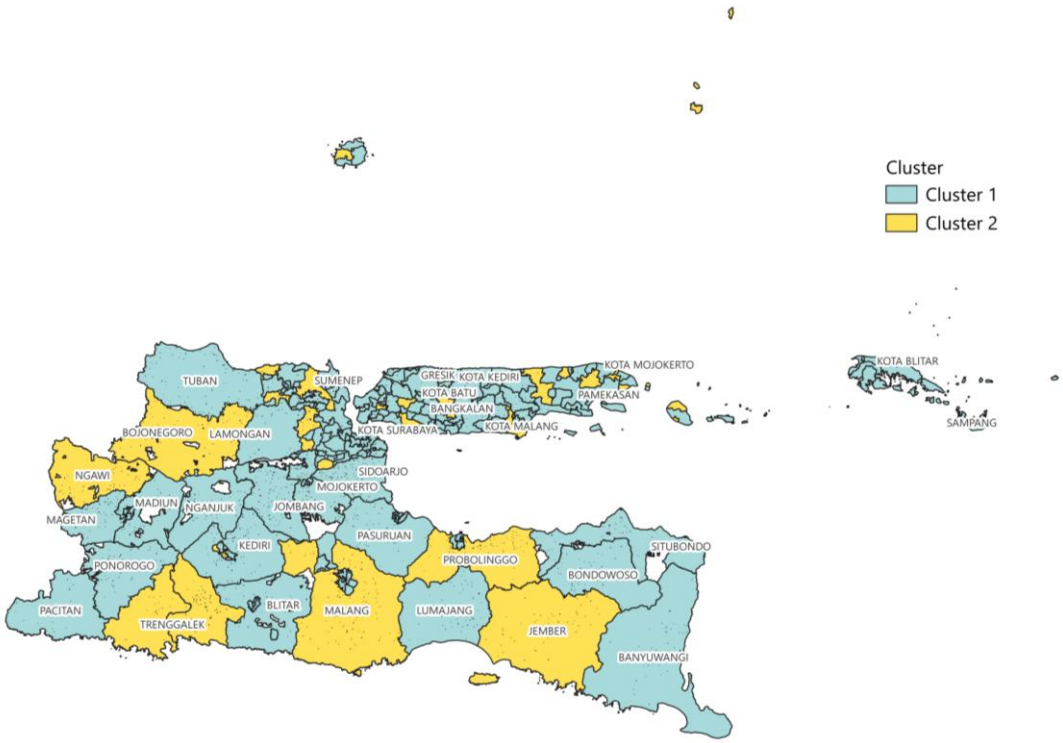


Figure 2. Map of Clustering Results Using the CFPA-FGWC Method

In Figure 2, Cluster 2 appears to be scattered across several parts of East Java, while Cluster 1 covers most of the remaining area. The two clusters do not form two completely separate geographic zones. This indicates that the CFPA-FGWC clustering is not determined solely by geographic proximity, but also takes into account similarities in flood impact indicators and spatial components.

The distribution of Cluster 2 can be further analyzed by considering the geographical conditions of its member regions. Ngawi and Bojonegoro are part of the Bengawan Solo river basin system; therefore, rainfall conditions in the upstream areas and changes in river discharge are relevant factors in flood analysis. Meanwhile, Trenggalek, Tulungagung, Malang, and Jember have landscapes that include the southern coastal region as well as hills and mountains. These conditions can serve as a context for assessing the potential for hydrometeorological hazards. However, a direct relationship between geographic conditions and flood impact patterns in each region cannot yet be concluded without more detailed supporting data.

3.5. Fuzzy Membership Results from CFPA-FGWC

Unlike hard clustering, which assigns a single label to each region, CFPA-FGWC generates fuzzy membership degrees for each cluster. Membership degrees are denoted by u_{1j} for Cluster 1 and u_{2j} for Cluster 2, where $u_{1j} + u_{2j} = 1$ for each region. These values indicate the relative proximity of a region to the two formed clusters. The dominant cluster is determined based on the larger membership value.

Table 6. Fuzzy Membership Results from CFPA-FGWC

Regency/City	u_{1j}	u_{2j}	Cluster	Regency/City	u_{1j}	u_{2j}	Cluster
Pacitan	0.8461	0.1539	1	Magetan	0.8205	0.1795	1
Ponorogo	0.8103	0.1897	1	Ngawi	0.3019	0.6981	2
Trenggalek	0.3111	0.6889	2	Bojonegoro	0.3332	0.6668	2
Tulungagung	0.3428	0.6572	2	Tuban	0.8653	0.1347	1
Blitar	0.8525	0.1475	1	Lamongan	0.8565	0.1435	1
Kediri	0.8267	0.1733	1	Gresik	0.6327	0.3673	1
Malang	0.4306	0.5694	2	Bangkalan	0.8590	0.1410	1
Lumajang	0.8419	0.1581	1	Sampang	0.8621	0.1379	1
Jember	0.4053	0.5947	2	Pamekasan	0.6925	0.3075	1
Banyuwangi	0.8064	0.1936	1	Sumenep	0.2985	0.7015	2
Bondowoso	0.5469	0.4531	1	Kota Kediri	0.8114	0.1886	1
Situbondo	0.5300	0.4700	1	Kota Blitar	0.8080	0.1920	1
Probolinggo	0.3305	0.6695	2	Kota Malang	0.6961	0.3039	1
Pasuruan	0.6101	0.3899	1	Kota Probolinggo	0.8379	0.1621	1
Sidoarjo	0.8769	0.1231	1	Kota Pasuruan	0.7743	0.2257	1
Mojokerto	0.8868	0.1132	1	Kota Mojokerto	0.8271	0.1729	1
Jombang	0.8745	0.1255	1	Kota Madiun	0.8170	0.1830	1
Nganjuk	0.8310	0.1690	1	Kota Surabaya	0.8182	0.1818	1
Madiun	0.7712	0.2288	1	Kota Batu	0.8021	0.1979	1

Based on Table 6, most regions have a higher membership degree in Cluster 1. The eight regions included in Cluster 2 have u_{2j} values greater than u_{1j}, with u_{2j} values ranging from 0.5694 to 0.7015. Meanwhile, Bondowoso and Situbondo have relatively balanced membership values between the two clusters, at 0.5469–0.4531 and 0.5300–0.4700, respectively.

These differences in membership degrees provide additional information regarding the resulting cluster labels. Areas with large differences in membership values indicate a clearer dominance toward

one of the clusters, while more balanced values indicate a relatively closer position to both cluster structures. Thus, fuzzy membership provides an indication of the relative proximity of each area to the clusters, which is not fully represented by the dominant cluster label. The use of membership degrees to evaluate partition structures is also consistent with studies on validity indices in fuzzy clustering [27], [31], [35].

Bondowoso has a membership degree of 0.5469 in Cluster 1 and 0.4531 in Cluster 2, while Situbondo has membership degrees of 0.5300 and 0.4700, respectively. These values indicate that both regions have relatively balanced membership degrees compared to other regions that are dominant in one cluster. Nevertheless, Bondowoso and Situbondo are still classified as belonging to Cluster 1 because they have a higher membership degree in that cluster. The relatively balanced membership degrees indicate that the characteristics of both regions are similar to the profiles of both Cluster 1 and Cluster 2, so both can be viewed as occupying a transitional position between the characteristics of the two clusters. This condition must be taken into account in disaster mitigation planning by considering the characteristics of flood impacts in each region, in line with the importance of considering social, economic, and environmental factors in flood resilience policies. However, in this study, fuzzy membership degrees were used to describe relative proximity to clusters, not as a probability of flooding or a direct measure of disaster risk levels.

3.4. Characteristics of Flood Impact Clusters

After the regional structure and fuzzy membership values were obtained, the CFPA-FGWC results were analyzed based on the average of 10 indicators for each cluster. A summary of the characteristics of each cluster on the original data scale is presented in Table 7.

Table 7. Average Profile of Flood Impact Indicators Based on CFPA-FGWC Clusters

Cluster	X1	X2	X3	X4	X5	X6	X7	X8	X9	X10
1	4.07	0.00	0.10	12,780.37	4.67	1.37	6.30	2,374.10	0.17	0.47
2	7.75	1.25	0.00	9,581.00	1.50	0.88	3.50	1,577.88	0.13	1.75

Based on Table 7, Cluster 1, consisting of 30 regencies/cities, showed higher average values for the number of displaced and affected people as well as for the housing damage indicator. The average number of displaced and affected people (X4) reached 12,780.37, compared to 9,581 in Cluster 2. The average number of severely damaged (X5), moderately damaged (X6), slightly damaged (X7), and flooded homes (X8) in Cluster 1 was 4.67, 1.37, 6.30, and 2,374.10 units, respectively. This pattern indicates that Cluster 1 was more significantly affected by the flood's impact on the population and damage to dwellings.

Cluster 2, consisting of 8 regencies/cities, exhibits different characteristics across several indicators. The average number of incidents (X1) reached 7.75, higher than the 4.07 incidents recorded in Cluster 1. Cluster 2 also has an average of 1.25 fatalities and missing persons (X2) and 1.75 affected bridges (X10), higher than Cluster 1's figures of 0 fatalities and 0.47 bridges, respectively. Thus, Cluster 2 stands out in terms of incident frequency, fatalities and missing persons, and impacts on bridges.

Overall, the two clusters exhibit distinct, multidimensional flood impact profiles. Cluster 1 is characterized more by the large number of displaced and affected residents as well as damage to homes, whereas Cluster 2 stands out in terms of the number of incidents, fatalities and missing persons, and damage to bridges. These differences indicate that the CFPA-FGWC classifies regions based on a combination of several flood impact indicators, rather than on any single indicator.

The importance of considering multiple indicators in the analysis of flood-affected areas is also reflected in the study by Nasution et al. [10], which used a fuzzy spatial clustering approach to examine

flood vulnerability in Jakarta. Other studies emphasize that flood vulnerability must be assessed through both social and physical dimensions, including population characteristics and exposure to flood impacts [2], [3], [32]. In line with these previous studies, the results of this study show that differences in flood impacts across regions are reflected in a combination of indicators: frequency of occurrence, number of affected residents, damage to homes, fatalities and missing persons, and impacts on infrastructure.

The differences in the characteristics of the two clusters can serve as an initial consideration in determining the direction of flood mitigation efforts. Cluster 1 has higher average numbers of displaced and affected residents (X4) as well as greater damage to homes (X5–X8). Therefore, mitigation efforts in this cluster can be directed toward providing evacuation shelters, distributing emergency supplies, improving the flood resilience of homes, and upgrading settlements based on the area's vulnerability level.

Conversely, Cluster 2 has higher average numbers of flood events (X1), fatalities and missing persons (X2), and affected bridges (X10). Mitigation efforts in this cluster can focus on inspecting bridge conditions, maintaining transportation infrastructure, strengthening early warning systems, and conducting community evacuation drills. These recommendations serve as initial guidelines based on the cluster profiles; therefore, their implementation must be tailored to the risk assessments and specific conditions of each region.

4. Conclusion

This study compares the performance of Fuzzy Geographically Weighted Clustering (FGWC), FGWC combined with the Flower Pollination Algorithm (FGWC-FPA), and CFPA-FGWC in clustering flood impacts across 38 regencies and cities in East Java in 2025. The evaluation results, using six internal validity indices, show that CFPA-FGWC yields the best results for five indices namely, the Partition Coefficient (PC), Partition Index (SC), Separation Index (SI), I-Index (IFV), and Kwon while FGWC-FPA produces a lower Classification Entropy (CE) value. These results indicate that chaos-based optimization has the potential to improve the quality of clustering based on most of the indices used, although this advantage does not apply to all validity measures.

The optimal CFPA-FGWC clustering yielded two clusters with distinct flood impact characteristics. The first cluster consists of 30 regencies/cities and is characterized by a large number of displaced and affected residents as well as damage to homes, while the second cluster consists of 8 regencies/cities characterized by a high number of flood events, fatalities and missing persons, and damage to bridges. These differences indicate that flood impacts vary across regions; thus, the clustering results can serve as supporting information for mitigation planning tailored to regional characteristics. However, the findings of this study are limited to data from 2025 and the available indicators. Therefore, future research should test the consistency of the clustering results across different time periods and incorporate additional indicators to strengthen the interpretation of flood impact characteristics.

References

[1] Wen, J., Wan, C., Ye, Q., Yan, J., & Li, W. (2023). Disaster Risk Reduction, Climate Change Adaptation and Their Linkages with Sustainable Development over the Past 30 Years: A Review: Wen et al. Disaster Risk Reduction, Climate Change Adaptation and Sustainable Development. *International Journal of Disaster Risk Science*, *14*(1), 1-13.

[2] Tate, E., Rufat, S., Rahman, M. A., & Hoover, S. (2025). Profiles of social vulnerability for flood risk reduction. *International Journal of Disaster Risk Reduction*, *118*, 105250.

[3] Dulawan, J. M. T., Imamura, Y., Konishi, T., Amaguchi, H., & Ohara, M. (2024). A systematic framework for assessing social vulnerability to flood for integrated flood risk management: A case study in Metro Manila, Philippines. *International Journal of Disaster Risk*

Reduction, *112*, 104778.
[4] Ranabahu, S. S., Lowry, J. H., & Imran, M. (2025). An indicator-based GIS approach for assessing a community's vulnerability to mobility-related disruptions caused by floods. *International Journal of Disaster Risk Reduction*, *120*, 105342.
[5] Yang, J., Ahn, D., Bahk, J., Park, S., Rizqihandari, N., & Cha, M. (2024). Assessing climate risks from satellite imagery with machine learning: A case study of flood risks in Jakarta. *Climate Risk Management*, *46*, 100651.
[6] Duykers, J. L., Ardana, K., Yulistiani, R., Irwansyah, E., & Fitrianah, D. (2023). Identifying factors for supporting early warning flood using clustering approach and geo-spatial analysis. *Procedia Computer Science*, *227*, 540-547.
[7] Askari, K., Zheng, W., Shi, S., Chu, J., & Wang, F. (2025). A novel entropy-based machine learning frame work for flood risk mapping in Pakistan. *Journal of Hydrology: Regional Studies*, *62*, 102911.
[8] Guan, X., Xu, T., Meng, Y., & Lv, H. (2025). Characterization of spatial and temporal distribution of flood impacts in China. *Natural Hazards*, *121*(18), 21745-21769.
[9] Chakraborty, L., Thistlethwaite, J., Scott, D., Henstra, D., Minano, A., & Rus, H. (2023). Assessing social vulnerability and identifying spatial hotspots of flood risk to inform socially just flood management policy. *Risk Analysis*, *43*(5), 1058-1078.
[10] Nasution, B. I., Saputra, F. M., Kurniawan, R., Ridwan, A. N., Fudholi, A., & Sumargo, B. (2022). Urban vulnerability to floods investigation in jakarta, Indonesia: A hybrid optimized fuzzy spatial clustering and news media analysis approach. *International Journal of Disaster Risk Reduction*, *83*, 103407.
[11] Caraka, R. E., Kurniawan, R., Nasution, B. I., Jamilatuzzahro, J., Gio, P. U., Basyuni, M., & Pardamean, B. (2021). Micro, small, and medium enterprises' business vulnerability cluster in Indonesia: An analysis using optimized fuzzy geodemographic clustering. *Sustainability*, *13*(14), 7807.
[12] Nasution, B. I., Kurniawan, R., Siagian, T. H., & Fudholi, A. (2020). Revisiting social vulnerability analysis in Indonesia: An optimized spatial fuzzy clustering approach. *International Journal of Disaster Risk Reduction*, *51*, 101801.
[13] Yozza, H., & Mukhlis, A. (2024). Clustering Provinces in Indonesia Based on The Main Food Crop Production Using The Spatial Fuzzy C-Means. *EKSAKTA: Berkala Ilmiah Bidang MIPA*, *25*(04), 498-507.
[14] Yanti, T. D. H., Styawan, Y., & Wulandari, R. (2026). Spatial-Temporal Analysis of b-value Before and After the 2018 Lombok Earthquake Using OK1993 Based on Voronoi in the Bali-NTB Region. *EKSAKTA: Berkala Ilmiah Bidang MIPA*, *27*(03), 263-274.
[15] Ramadhan, R. R., Wulandari, R., Styawan, Y., Nugraha, P., & Syafitri, M. N. (2026). Integrating Time-Dependent and Time-Independent Earthquake Recurrence Models for Fault-Based Seismic Hazard Mapping in Indonesia. *EKSAKTA: Berkala Ilmiah Bidang MIPA*, *27*(04), 771-789.
[16] Surbakti, E. N. B., Saputra, A., Wulandari, R., & Styawan, Y. (2026). Peak Ground Acceleration Mapping of Indonesia Using Complementary Area Source and Smoothed Seismicity Approaches. *EKSAKTA: Berkala Ilmiah Bidang MIPA*, *27*(03), 241-256.
[17] Sari, D. P., & Rosadi, D. (2022). Disaster Mitigation Efforts Using K-Medoids Algorithm and Bayesian Network. *EKSAKTA: Berkala Ilmiah Bidang MIPA*, *23*(03), 231-241.
[18] Nugrahanto, Z. Z., & Sofro, A. Y. (2026). Implementation of DBSCAN and K-MEANS++ Methods for Flood Vulnerability Cluster Mapping in East Java Province, 2024. *CAUCHY: Jurnal Matematika Murni dan Aplikasi*, *11*(1), 32-44.
[19] Kaur, A., Kumar, Y., & Sidhu, J. (2024). Exploring meta-heuristics for partitional clustering: methods, metrics, datasets, and challenges: A. Kaur et al. *Artificial Intelligence Review*, *57*(10),

287.
[20] Mergos, P. E., & Yang, X. S. (2021). Flower pollination algorithm parameters tuning. *Soft computing*, *25*(22), 14429-14447.
[21] Ong, K. M., Ong, P., & Sia, C. K. (2022). A new flower pollination algorithm with improved convergence and its application to engineering optimization. *Decision Analytics Journal*, *5*, 100144.
[22] Song, H., Bei, J., Zhang, H., Wang, J., & Zhang, P. (2024). Hybrid algorithm of differential evolution and flower pollination for global optimization problems. *Expert Systems with Applications*, *237*, 121402.
[23] Jia, Y., Wang, S., Liang, L., Wei, Y., & Wu, Y. (2023). A flower pollination optimization algorithm based on cosine cross-generation differential evolution. *Sensors*, *23*(2), 606.
[24] Daqaq, F., Ouassaid, M., Kamel, S., Ellaia, R., & El-Naggar, M. F. (2022). A novel chaotic flower pollination algorithm for function optimization and constrained optimal power flow considering renewable energy sources. *Frontiers in Energy Research*, *10*, 941705.
[25] Ozsoydan, F. B., & Baykasoglu, A. (2021). Chaos and intensification enhanced flower pollination algorithm to solve mechanical design and unconstrained function optimization problems. *Expert systems with applications*, *184*, 115496.
[26] Simpson, C., Muñoz, M. A., & Campello, R. J. (2026). Instance space of clustering validation measures. *Data Mining and Knowledge Discovery*, *40*(3), 25.
[27] Wang, H. Y., Wang, J. S., & Wang, G. (2022). A survey of fuzzy clustering validity evaluation methods. *Information Sciences*, *618*, 270-297.
[28] Ikotun, A. M., Habyarimana, F., & Ezugwu, A. E. (2025). Cluster validity indices for automatic clustering: A comprehensive review. *Heliyon*, *11*(2).
[29] Hassan, B. A., Tayfor, N. B., Hassan, A. A., Ahmed, A. M., Rashid, T. A., & Abdalla, N. N. (2024). From A-to-Z review of clustering validation indices. *Neurocomputing*, *601*, 128198.
[30] Ismail, K. N., Seman, A., & Samah, K. A. F. A. (2024). A new internal validity index for fuzzy c-means algorithm. *IEEE Access*, *12*, 85995-86012.
[31] Kwon, S. H., Kim, J., & Son, S. H. (2021). Improved cluster validity index for fuzzy clustering. *Electronics Letters*, *57*(21), 792-794.
[32] Coleman, N., Clarke, A., Esparza, M., & Mostafavi, A. (2025). Analyzing common social and physical features of flash-flood vulnerability in urban areas. *International Journal of Disaster Risk Reduction*, *122*, 105437.
[33] Wongoutong, C. (2024). The impact of neglecting feature scaling in k-means clustering. *PloS one*, *19*(12), e0310839.
[34] Ong, P., Jian, J., Yin, J., & Ma, G. (2023). Characteristic wavelength optimization for partial least squares regression using improved flower pollination algorithm. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, *302*, 123095.
[35] Todeschini, R., Ballabio, D., Termopoli, V., & Consonni, V. (2024). Extended multivariate comparison of 68 cluster validity indices. A review. *Chemometrics and intelligent laboratory systems*, *251*, 105117.
[36] Sirmen, R. T., & Üstündağ, B. B. (2022). Internal validity index for fuzzy clustering based on relative uncertainty. *Computers, Materials and Continua*, *72*(2), 2909-2926.
[37] Wang, H. Y., Wang, J. S., & Wang, G. (2023). Multi-fuzzy clustering validity index ensemble: A Dempster-Shafer theory-based parallel and series fusion. *Egyptian Informatics Journal*, *24*(4), 100417.
[38] Kalpanarani, K., & Grace, G. H. (2025). A Statistical Test Based Separability Measure for Internal Cluster Validation: K. Kalpanarani, GH Grace. *Neural Processing Letters*, *57*(6), 95.
[39] Yan, B., & Xie, Z. (2026). A robust fuzzy cluster validity index based on local distances. *Expert*

Systems with Applications, *295*, 128883.
[40] Liu, F., Wang, J., & Liu, Y. (2024). IMI2: a fuzzy clustering validity index for multiple imbalanced clusters. *Expert Systems with Applications*, *238*, 122231.
[41] Oskouei, A. G., Samadi, N., Khezri, S., Moghaddam, A. N., Babaei, H., Hamini, K., ... & Arasteh, B. (2025). Feature-weighted fuzzy clustering methods: An experimental review. *Neurocomputing*, *619*, 129176.
[42] Tang, Y., Li, B., Pedrycz, W., & Wang, X. (2025). A clustering validity index with multi-granularity fusion for multiple fuzzy clustering algorithms. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *47*(10), 8379-8396.