Available online at: http://journal.uny.ac.id/index.php/pythagoras
PYTHAGORAS: Jurnal Pendidikan Matematika, 16 (2), 2021, 217-232

Drought-prone areas mapping using fuzzy c-means method in Gunungkidul district

Kismiantini[1]*, Fajra Husniyah[2], Osval Antonio Montesinos-López[3]
[1] Statistics Study Program, Universitas Negeri Yogyakarta, Indonesia
[2] Mathematics Study Program, Universitas Negeri Yogyakarta, Indonesia
[3] Facultad de Telemática, Universidad de Colima, México
* Corresponding Author. E-mail: kismi@uny.ac.id

ARTICLE INFO

Article History:
Received: 10 Sep. 2021
Revised: 10 Sep. 2021
Accepted: 28 Dec. 2021

Keywords:
Gunungkidul district,
drought,
fuzzy c-means,
mapping.

ABSTRACT

Kabupaten Gunungkidul merupakan salah satu kabupaten di Daerah Istimewa Yogyakarta yang sering mengalami bencana kekeringan. Tujuan dari penelitian ini adalah memetakan daerah rawan kekeringan di Kabupaten Gunungkidul dengan menggunakan metode *fuzzy c-means*, untuk mempermudah pemerintah dalam mengalokasikan bantuan *dropping* air ke wilayah yang terdampak kekeringan. Variabel-variabel penelitian meliputi yaitu curah hujan, jenis tanah, infiltrasi, kemiringan lereng, dan penggunaan lahan. Variabel-variabel tersebut tergolong dalam data ordinal, sehingga perlu dilakukan transformasi data menggunakan metode suksesif interval agar dapat diolah menggunakan metode *fuzzy c-means*. Indeks validitas *cluster* yaitu *Xie and Beni index, partition coefficient,* dan *modification partition coefficient* digunakan untuk mencari k optimal. Hasil *fuzzy c-means clustering* diperoleh tiga cluster dengan tingkat kerentanan rendah ada di 7 kecamatan, tingkat kerentanan sedang ada di 8 kecamatan, dan tingkat kerentanan tinggi ada di 3 kecamatan. Curah hujan, penggunaan lahan, jenis tanah, infiltrasi, kemiringan lereng merupakan faktor bahaya kekeringan yang memiliki pengaruh paling besar hingga paling rendah pada penelitian ini.

Gunungkidul district is one of the districts in the Special Region of Yogyakarta that is frequently affected by drought disasters. The purpose of this study is to map drought-prone areas in Gunungkidul district using the fuzzy c-means method, making it easier for the government to allocate water-dropping assistance to drought-affected areas. The research variables include rainfall, soil type, infiltration, slope, and land use. The type of variables is in an ordinal scale, so they must be transformed using the successive interval method before being analyzed using the fuzzy c-means method. The cluster validity indexes of the Xie and Beni index, partition coefficient, and modification partition coefficient were used to find the optimal k. The results of fuzzy c-means clustering revealed three clusters with a low level of vulnerability consisting of 7 sub-districts, a moderate level of vulnerability consisting of 8 sub-districts, and a high level of vulnerability consisting of 3 sub-districts. Rainfall, land use, soil type, infiltration, and slope were the drought hazard factors with the greatest to least effect in this study.

How to Cite:
Kismiantini, Husniyah, F., & Montesinos-López, O. A. (2021). Drought-prone areas mapping using fuzzy c-means method in Gunungkidul district. *Pythagoras: Jurnal Pendidikan Matematika, 16*(2), 217-232. https://doi.org/10.21831/pythagoras.v16i2.43780

https://doi.org/10.21831/pythagoras.v16i2.43780

INTRODUCTION

The El Niño Southern Oscillation (ENSO) phenomenon is one of the causes of climate change in the world. Many studies have been conducted on the link of ENSO to climate anomalies throughout the world. Chiew et al. (1998) provided an overview of the relationship between ENSO and rainfall, drought, and streamflow in Australia. Baudoin et al. (2017) investigated how South Africa responded to drought over time due to the 2016 El Niño. Leitold et al. (2018) found that El Niño conditions accelerated canopy turnover in the central Amazon particularly during drought years. Ewbank et al. (2019) studied the resilience of rural communities in Nicaragua and Ethiopia to El Niño-related drought prior to and during the agricultural season. Costa et al. (2021) identified rainfall and drought extremes in the State of Alagoas, as well as their relationship with ENSO. The El Niño phenomenon, which occurred in 2015/2016 and was categorized as strong intensity, had a significant impact on meteorological droughts in Indonesia (Avia & Sofiati, 2018). The ENSO phenomenon is a sea-level condition in the Pacific Ocean that has experienced an increase or decrease in sea surface temperature resulting in a shift in seasons in Indonesia (Nabilah et al., 2017). According to duration and magnitude of drought strength, Muharsyah & Ratri (2015) reported that 22 of the 29 rainfall stations in Bali, Indonesia experienced the longest drought during the weak, moderate, and strong El Niño periods. Khasanah & Sastra (2017) stated also that the ENSO phenomenon affects the condition of the waters in West Sumatra as evidenced by the fact that when the El-Niño phenomenon occurs, the water conditions decline, and vice versa when the La-Nina phenomenon occurs, the water conditions increase. One of the areas experiencing the worst drought in Indonesia is Gunungkidul district in the Special Region of Yogyakarta (Apriani et al., 2014). Gunungkidul district's topography is dominated by karst hilly areas that make it difficult to store groundwater (Nugroho et al., 2020). Also, according to the number of dropped water distributions in 2019, Gunungkidul has the largest amount of 26868×1000 liters of water, followed by Bantul at 3218×1000 liters, Kulon Progo at 1037×1000 liters of water and Sleman at 1455×1000 liters of water(BPBDDIY, 2019).

The prolonged dry season can cause drought because water reserves will be depleted due to transpiration, evaporation and other human uses (Darojati et al., 2015). Drought can also be caused by soil characteristics in each area, such as soil type, infiltration level, slope, and land use. Soil type affects the ability of the soil to store water, so that it can indicate the level of dryness because it is related to the availability of water for plant roots (Halwatura et al., 2017). The size of the soil absorption capacity of rainwater falling on the surface can be seen from the infiltration capability; areas with low infiltration capabilities can allow drought disasters (Sigit, 2016b). Steeper slopes have less water content because water moves faster on a steep slope than on flat, sloping or wavy slopes (Banjarnahor et al., 2018). Uncontrolled change in land use, such as the loss of forest land cover for other types of land use, can result in droughts, floods, and landslides (Pawitan, 2014).

Drought-prone land has the potential to reduce agricultural production (Hidayati & Suryanto, 2015). Agricultural drought is a highly complex natural disaster that affects a large area and results in a significant decrease in agricultural production (Surmaini, 2016). One of the disaster countermeasures that can be done is mapping drought-prone areas. This mapping of drought-prone areas aims to assist the government in allocating and distributing water to drought-affected areas. Numerous studies have demonstrated the importance of drought-prone mapping. Al-Bakri et al. (2019) made drought vulnerability mapping to show the high rainfall zones in Jordan which could help on improving water management. Paparrizos et al. (2018) analyzed and mapped the present and future drought conditions on the Sperchios River Basin, Geropotamos River Basin, and Ardas River Basin in Greece where their findings could be used as a guide for predicting future climate conditions affecting water management. Moghbeli et al. (2020) investigated the application of the standardized precipitation index (SPI) for mapping drought severity as an effective water management tool in an arid climate region, southeastern Iran. Habibie et al. (2020) classified suitable lands for maize production in drought-prone areas of central East Java of Indonesia. Those studies were mapping of drought-prone areas with different methods, such as the combined drought index and mapping via geographic information system (GIS), the downscaling technique and the ordinary Kriging, the standardized precipitation index (SPI) at different timescales, and the analytic hierarchy process (AHP).

To know the different levels of drought-prone areas, a clustering method may be useful and easier to map them. Clustering is the process of identifying natural clusters, where similar objects are grouped together in the same cluster (Omran et al., 2007). One of the most popular clustering algorithms in use today is the K-means algorithm because it is relatively fast while remaining simple to understand and apply in practice (Raykov et al., 2016). In clustering, similarity between objects is determined by calculating the distance function. Nishom (2019) showed that the Euclidean distance had the highest accuracy in clustering the status of disparity in the needs of teachers in Tegal city when compared to the Manhattan distance and Minkowski distance. Hence, the choice of the distance function which is able to quantify the similarity among the data objects is important in clustering. Xulu et al. (2019) employed the K-means cluster analysis based on normalized difference water index (NDWI) to separate compartments into drought-affected and non-drought affected clusters in Zululand region, South Africa. Feng et al. (2019) used the K-means clustering algorithm to classify Bureau of Meteorology (BOM) climate stations in South-Eastern Australia into two coherent clusters. Gader et al. (2020) used the K-means cluster analysis to identify four regions with distinct rainfall regimes in the Mediterranean catchment of the Medjerda, Tunisia. Ramadhan et al.

(2020) used the K-means algorithm to obtain the clustering of drought-prone areas based on the analysis of hotspots data in Riau, Indonesia. The K-means algorithm is widely used in various fields but its drawbacks include the need to specify the number of clusters in advance, sensitive to outliers, inability to deal with non-convex clusters of varying size and density, sensitive to scale of the data set, and different initial centroids producing difference results (Govender & Sivakumar, 2020). Moradi & Dariane (2015) demonstrated that evolving neural network (ENN) conditioned on fuzzy c-means (FCM) outperformed than K-means clustering based ENN and regular ENN. Alam & Paul (2020) reported that the fuzzy c-means (FCM) algorithm outperforms the K-means clustering in terms of cluster homogeneity rainfall gauge stations in Bangladesh. Hence, in this study the FCM is used to cluster drought-prone areas.

The FCM algorithm is a popular fuzzy clustering method based on fuzzy theory (Simhachalam & Ganesan, 2015). The FCM clustering method is a data clustering technique that determines the presence of data in a cluster based on different degrees of membership between 0 and 1 (Rahakbauw et al., 2017). In the fuzzy clustering approach, FCM is the most-known method and performs well in cluster detection (Pimentel & de Souza, 2016). Another advantage of the FCM is that it can show the relationship between different cluster patterns (Sharma & Kamal Borana, 2014). Although the FCM method is the most widely used, it has a weakness in determining the optimal number of clusters (Yang & Nataliani, 2017). Many cluster validity indices for fuzzy clustering algorithms have been proposed in the literature, such as the Xie Beni index (XBI) (Xie & Beni, 1991), partition coefficient (PC) (Bezdek, 1973), and modified partition coefficient (MPC) (Dave, 1996). Pertiwi & Kurniawan (2017) used FCM with validation indices such as the Xie and Beni index, partition coefficient, modified partition coefficient, and others to map flood-prone areas in Indonesia.

In statistical analysis, differences in data types greatly affect the choice of models or statistical tests. The measurement scale is a rule that is required to quantify data derived from measurement variables (Febtriko, 2017). The variable used in fuzzy logic must have a continuous value (Rizal & Hakim, 2015), whereas the variables in this study include ordinal scale. Therefore, they must be transformed from ordinal into interval scale. A method which is common to transform the ordinal scale into interval scale is successive interval method (Maranell, 2017). The purpose of the successive interval method is to convert ordinal scale into interval scale by changing the cumulative proportion of each variable in the category to the default normal curve value (Ningsih & Dukalang, 2019). Many studies used the successive interval method to transform the data in ordinal scale into interval scale (Herawaty (2014), Sofiani et al. (2017), Mardiana et al. (2020)).

The purpose of this study is to use the fuzzy c-means method to map the results of clustering the level of drought vulnerability in Gunungkidul district. The variables used in this study include rainfall, soil type, infiltration, slope, and land use; where all variables were on an ordinal scale. Hence, the successive interval method was employed to them prior to data analysis. The novel aspect of this paper is that it uses the FCM clustering method to group drought-prone areas in Gunungkidul district based on the transformed variables via the successive interval method.

DATA AND METHOD

Data

Data in this study were collected based on secondary data from eighteen sub-districts in Gunungkidul district. The research variables included rainfall (categories were coded as 4: < 1000 mm/year, 3: 1000-1500 mm/year, 2: 1500-2000 mm/year, and 1: 2000-2500 mm/year), soil type (categories were coded as 3: mediteran, 2: latosol, and 1: grumusol), soil infiltration rate (categories were coded as 2: slow and 1: medium), land use (categories were coded as 6: settlement, 5: rainfed rice fields, 4: irrigation paddy, 3: garden, 2: production forest, 1: conservation forest and protected forest) (DPTRDIY, 2016), and slope (categories were coded as 3: steep (> 25°), 2: moderate (15°-25°), and 3: gentle (<15°) (BPS Gunungkidul, 2016). All variables were on an ordinal scale, they must be transformed into interval scales using the successive interval method prior to analysis with FCM algorithm. Details on how to carry out the successive interval method can be found in Maranell (2017). The results of clustering for mapping drought-prone areas were visualized using a Geographic Information System (GIS) (Chang et al., 2019). Figure 1 depicts the research flow chart.

Method

Fuzzy C-Means Algorithm

Fuzzy c-means (FCM) is a data clustering technique that uses the degree of membership to determine the existence of each data in a cluster. The FCM was first introduced by Jim Bezdek in 1981. The basic concept of the FCM is to find the center of each cluster, which will represent the average location of each cluster. The cluster center in this initial stage is not accurate. Each data point has a degree of membership in each cluster. In order for the cluster center to be accurate, the cluster center and the degree of membership at each data point are repaired repeatedly. This iteration is based on the minimization of the objective function, which describes the distance from the data point to the center of the cluster and is weighted by the degree of membership of the data point. The output of fuzzy c-means is not a fuzzy inference system, but rather a series of cluster centers and degrees of membership for each data point (Kusumadewi & Purnomo, 2010). The following are the steps for the FCM algorithm (Bezdek, 1981).

a. The input data to be clustered is in the form of a matrix of order $n \times p$, where n = the number of observations; and p = the number of variables; X_{ij} = the ith observation (i=1,2,...,n) from the j attribute (j=1,2,...,p).

b. Determine the number of clusters ($k \geq 2$); the weighting exponent ($m \in [1, \infty)$, for example $m = 2$); maximum iteration (*MaxIter)*; initial iteration ($t = 1$); smallest error($\varepsilon =$ a very small positive value).

c. Determine the partition matrix $\boldsymbol{U}$ (degree of membership) at random.

$$\boldsymbol{U} = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_1) & \cdots & \mu_{1c}(x_1) \\ \mu_{21}(x_2) & \mu_{22}(x_2) & \cdots & \mu_{2c}(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n1}(x_n) & \mu_{n2}(x_n) & \cdots & \mu_{nc}(x_n) \end{bmatrix} \quad (1)$$

where $\boldsymbol{U}$ is a partition matrix $\boldsymbol{U}$ or a degree of membership; $\boldsymbol{U}$ =$\sum_{k=1}^{c} \mu_{ik} = 1$, i=1,2,..., n; k=1,2,..., c.

d. Calculate the cluster centers (V) for each cluster using the equation:

$$V_{kj} = \frac{\sum_{k=1}^{c} {\mu_{ik}}^m X_{ij}}{\sum_{k=1}^{c} {\mu_{ik}}^m} \quad (2)$$

where V_{kj} is a cluster center, k=1,2,..., c; j=1,2,..., p; $\sum_{k=1}^{c} \mu_{ik}$ is an element of the partition matrix $\boldsymbol{U}$ (i=1,2,..., n; k=1,2,..., c); X_{ij} is sample data ($i = 1,2,\ldots,n$; $j = 1,2,\ldots,$p); m is the weighting exponent. Then it can be obtained a matrix $\boldsymbol{V}$ as follows:

$$\boldsymbol{V} = \begin{bmatrix} V_{11} & V_{12} & \cdots & V_{1p} \\ V_{21} & V_{22} & \cdots & V_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ V_{c1} & V_{c2} & \cdots & V_{cp} \end{bmatrix} \quad (3)$$

e. Calculate the distance between the object and the center of the cluster (d),

$$d_{ik} = \left[\sum_{j=1}^{p} \left(X_{ij} - V_{kj}\right)^2\right]^{\frac{1}{2}} \quad (4)$$

where d_{ik} is the distance between the object and the center of the cluster (i=1,2,..., n; k=1,2,..., c). The common distance used is the Euclidean distance.

f. Calculates the objective function (J_m) in the tth iteration by using an equation:

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

where J_m is the objective function (m= weighting exponent, $m \in [1, \infty)$), $\sum_{k=1}^{c} \mu_{ik}$ is the element of the partition matrix $\boldsymbol{U}$ (i=1,2,..., n; k=1,2,..., c).

g. Fix the degree of membership of each object in each cluster (repair partition matrix) by using equations

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{p} \left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{m-1}}}{\sum_{k=1}^{c} \left[\sum_{j=1}^{p} \left(X_{ij} - V_{kj}\right)^2\right]^{\frac{-1}{m-1}}} \quad (6)$$

h. Determine a criterion for stopping automatically that is the objective function of the last iteration minus the objective function of the previous iteration. If $|J_{mt} - J_{mt-1}| < \varepsilon$ or $t >$ *MaxIter,* then it stops. If it does not stop, then repeat the steps in point (d to h).

Cluster Validity Index

The clustering results are then validated with the cluster validity index. Three cluster validity indexes used in this study are explained as follows.

a. Xie Beni (XB) index

Xie Beni (XB) index aims to calculate the ratio of the total variation within the group and the separation of groups. The lower the XB value, the better the group partition. The XB index formula is defined below (Xie & Beni, 1991),

$$XB(c) = \frac{\sum_{i=1}^{n}\sum_{k=1}^{c}(\mu_{ik})^m d_{ik}^2(x_k - v_i)}{n\ \min_{ik} d_{ik}^2(v_i - v_k)} \quad (7)$$

where μ_{ik} is the membership value ($i = 1,2, \dots, n; k = 1,2, \dots, c$); m is the weighting exponent $m \in [1, \infty)$, for example m=2; n = number of objects; c = number of clusters; $d_{ik}^2(x_k - v_i)$ is the distance between objects to the cluster center (i=1,2,..., n; k=1,2,..., c).

b. Partition Coefficient (PC)

The Partition Coefficient aims to evaluate the membership value of each cluster regardless of the data. The PC value ranges between 0 and 1. The PC value that is close to 1 indicates the better cluster. The PC formula is defined below (Bezdek, 1973),

$$PC(c) = \frac{1}{N}\sum_{i=1}^{c}\sum_{k=1}^{N} u_{ik}^2 \quad (8)$$

where μ_{ik} is the membership value ($i = 1,2, \dots, n; k = 1,2, \dots, c$); n is the number of objects; and c is the number of clusters.

c. Modification Partition Coefficient (MPC)

Modification Partition Coefficient aims to overcome the shortcomings of the PC, which has a tendency to a monotone value or dependence on c, so the PC is modified to MPC (Dave, 1996) by using the equation below,

$$MPC = 1 - \frac{c}{c-1}(1 - PC) \quad (9)$$

where c is the number of clusters. The range of MPC value is between 0 and 1. The higher the MPC value, the better the cluster (Dave, 1996).

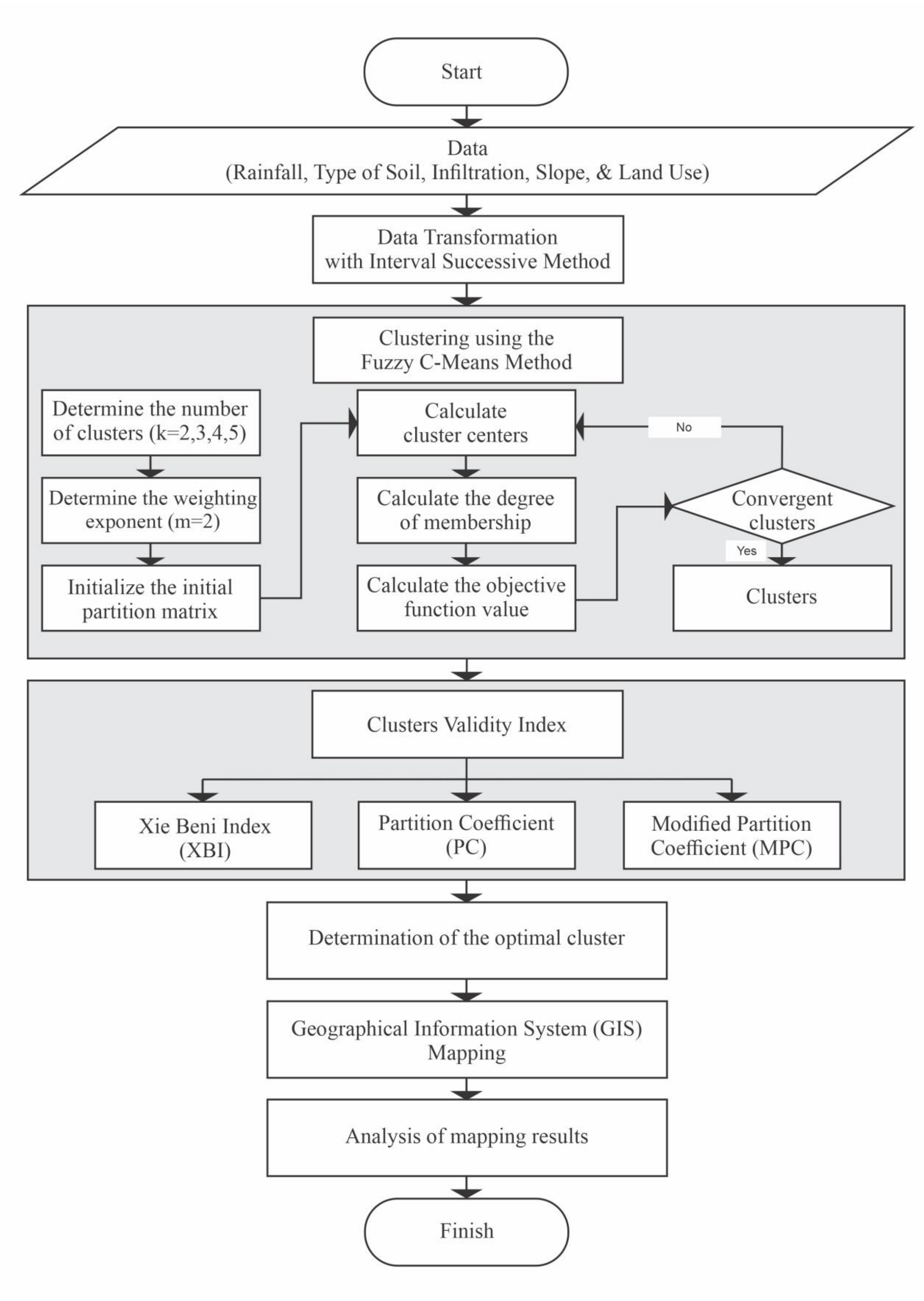

Figure 1. Research flow chart

RESULTS AND DISCUSSION

This study maps drought-prone areas in Gunungkidul district using a geographic information system based on the results of FCM clustering. The data used in this study were transformed from the ordinal to interval scale using the successive interval method. The transformed data in Table 1 were used as the input matrix. The clustering uses the FCM algorithm with the conditions of $k = 3, m = 2, \varepsilon = 10^9, t = 1$, and $MaxIter = 1000$. The input data is then processed using R Studio software with packages including ppclust (Cebeci et al., 2020), factoextra (Kassambara & Mundt, 2020), dplyr (Wickham et al., 2021), cluster (Maechler et al., 2021), fclust (Ferraro et al., 2019), psych (Revelle, 2021) and use the functions res.fcm <- fcm(x, centers = 3); res.fcm2<-ppclust2(res.fcm, "kmeans"); fviz_cluster(res.fcm2, data = x, ellipse.type = "convex", palette = "jco", repel = TRUE); res.fcm4<-ppclust2(res.fcm, "fclust"); idxpc <- PC(res.fcm4$U); paste("Partition Coefficient : ",idxpc); idxmpc <- MPC(res.fcm4$U); paste("Modified Partition Coefficient : ",idxmpc); clust=FKM(Data[,1:(ncol(Data))], k=3,m=2, stand = 0); xb=XB(clust$Xca, clust$U, clust$H, clust$m). Table 2 shows the cluster center at $k = 3$.

Table 1. The transformed data

No	Rainfall	Type of soil	Infiltration	Slope	Land use
1	4.155642524	3.050264212	2.669887026	1.0000000000	2.704510519
2	2.409099764	2.473856954	2.669887026	2.217452948	3.240261314
3	2.409099764	3.050264212	2.669887026	2.187400206	3.130638148
4	3.567739811	2.025132106	1.834943513	1.608726474	3.130638148
5	4.155642524	2.025132106	1.834943513	2.217452948	2.825074917
6	2.325613192	2.025132106	1.834943513	2.187400206	2.843040420
7	3.573461604	1.982571302	1.834943513	1.608726474	2.520895764
8	2.988419788	2.025132106	1.834943513	2.172373835	3.130638148
9	4.155642524	2.473856954	2.669887026	1.608726474	3.130638148
10	2.988419788	3.050264212	2.669887026	2.217452948	2.825074917
11	2.988419788	2.473856954	2.669887026	2.187400206	2.967354760
12	2.988419788	3.050264212	2.669887026	2.172373835	3.258501326
13	2.988419788	3.050264212	2.669887026	1.608726474	3.130638148
14	2.988419788	2.025132106	1.834943513	1.608726474	3.457387226
15	4.155642524	2.473856954	2.669887026	2.172373835	3.130638148
16	3.573461604	2.473856954	2.669887026	2.781100309	3.240261314
17	3.573461604	2.025132106	1.834943513	1.000000000	2.704510519
18	2.988419788	2.473856954	2.669887026	2.217452948	2.866884429

Table 2. Cluster center on $k = 3$

Cluster Center (V_{kj})	Rainfall	Type of soil	Infiltration	Slope	Land use
Cluster 1	3.367480552	2.056968171	1.878413041	1.687326721	2.976191125
Cluster 2	4.029142950	2.513800922	2.563860384	1.795481220	3.037035210
Cluster 3	2.858006164	2.736974429	2.622414865	2.154549193	3.050404935

The cluster centers in Table 2 are used to repair the partition matrix ***U*** (membership degrees) with the equation (6), resulting in Table 3. The objective function value is calculated using equation (5). The objective function value in the 62nd iteration (last iteration) is 4.713489170 and the previous iteration is 4.713489170, so the difference is 0.0000000 or less than 0.000000001. According to Table 3, cluster 1 has 7 districts, cluster 2 has 3 districts, and cluster 3 has 8 districts.

Clustering is performed on $k = 2,3,4,5$ in order to compare and determine the optimal k. The clustering results were then validated using the XB index, the PC, and the MPC. Validation results on $k = 2,3,4,5$ are shown in Table 4, and the validation results obtain that the k optimal is 3. Figure 2 shows the results of the clustering method using the FCM with 3 clusters.

Table 3. Clustering FCM results on $k = 3$

Sub-district	U matrix (degree of membership)			*Cluster*
	Cluster 1	*Cluster* 2	*Cluster* 3	
Semin	0.222911900	**0.585606350**	0.191481760	2
Purwosari	0.119959770	0.086901670	**0.793138560**	3
Girisubo	0.091111740	0.082831580	**0.826056680**	3
Wonosari	**0.901802070**	0.064156830	0.034041090	1
Paliyan	**0.445742040**	0.409903520	0.144354440	1
Ponjong	**0.437915370**	0.153725540	0.408359090	1
Playen	**0.760663350**	0.15146260	0.087874040	1
Semanu	**0.643060830**	0.130187230	0.226751940	1
Ngawen	0.046017540	**0.921554230**	0.032428220	2
Saptosari	0.070154510	0.090062770	**0.839782730**	3
Patuk	0.069785300	0.066428890	**0.863785810**	3
Tepus	0.065867360	0.086737900	**0.847394740**	3
Tanjungsari	0.154093020	0.193109590	**0.652797390**	3
Rongkop	**0.701644880**	0.130596350	0.167758770	1
Nglipar	0.088430110	**0.827232240**	0.084337640	2
Gedangsari	0.208606840	0.356702220	**0.434690940**	3
Karangmojo	**0.651437870**	0.223902550	0.124659580	1
Panggang	0.085063560	0.080751010	**0.834185430**	3

Table 4. Comparison of cluster validation results

Validity index	$k = 2$	$k = 3$	$k = 4$	$k = 5$	Optimal
XB	0.350432089	0.2313805727	0.685426759	0.545276664	$k = 3$
PC	0.687972943	0.599809747	0.523820193	0.501842336	$k = 2$
MPC	0.375945886	0.399714621	0.365093591	0.377302920	$k = 3$

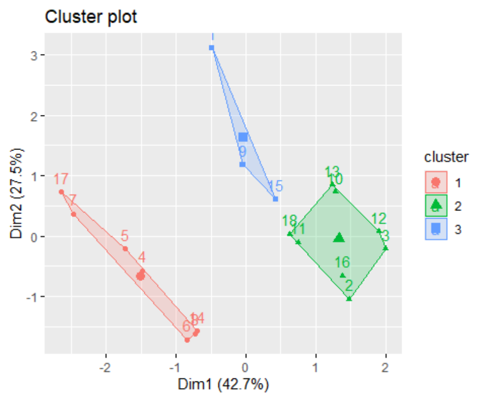

Figure 2. Plot clustering FCM method with $k = 3$

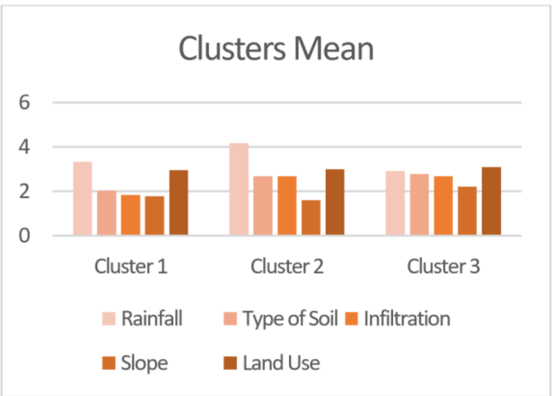

Figure 3. Clusters mean

The hazard category could be determined based on the cluster mean value and the characteristics of the variables in each cluster. The cluster mean is obtained from the average degree of membership in each cluster. Drought hazard factors with the greatest to lowest effect are rainfall, land use, distance to water sources, soil texture, and soil surface temperature (Darojati et al., 2015). According to Darojati et al. (2015) and the results of the clusters mean in Figure 3, the order of the drought hazard factors that have the greatest to lowest effect in this study are rainfall with a weight of 5, land use with a weight of 4, soil type with a weight of 3, infiltration with a weight of 2, slope with a weight of 1. The hazard category can be calculated from the multiplication between the cluster mean value and the weight of each variable, then added together to produce the hazard weight as in Table 5. The percentage of each category of variables can be used to analyze cluster characteristics (Table 6).

Table 5. The hazard category

Cluster center or cluster mean (V_{kj})	Rainfall	Type of soil	Infiltration	Slope	Land use	Sum
Cluster 1	3.367480552	2.056968171	1.878413041	1.687326721	2.976191125	11.96638
Cluster 2	4.029142950	2.513800922	2.563860384	1.795481220	3.037035210	13.93932
Cluster 3	2.858006164	2.736974429	2.622414865	2.154549193	3.050404935	13.42235
The weight of each variable	5	3	2	1	4	
The hazard category	16.8374	10.2848	9.39206	8.43663	14.8810	59.8319
	20.1457	12.5690	12.8193	8.97741	15.1852	69.6966
	14.2900	13.6849	13.1121	10.7728	15.2520	67.1118

Cluster 1

Cluster 1 includes Wonosari, Paliyan, Ponjong, Playen, Semanu, Rongkop, and Karangmojo sub-districts. Cluster 1 is a non-drought prone area. Table 6 shows that the most frequent rainfall vulnerability has a percentage of 44%, with an average of 1000-1500 mm/year. The area is dominated by Mediterranean soil types of 47% and grumusol of 47%. The infiltration rate in the area in cluster 1 has a comparable percentage with slow category of 50% and moderate 50%. The slope in this area has a gentle slope because it has a percentage of 46%. The land use in cluster 1 is 10% protected forest and 6% conservation forest, so there is still a place for water absorption and storage. As a result, the sub-districts in Cluster 1 are classified as not prone to drought disaster.

Cluster 2

The sub-districts in cluster 2 include Semin, Ngawen, and Nglipar. Table 6 shows that the percentage of rainfall prone is < 1000 mm/year by 50%, and 1000-1500 mm/year by 50%. The area is dominated by Mediterranean soil types by 60%. The infiltration rate in cluster 2 has a 100% slow infiltration. The slope in this area is dominated by gentle slopes with the percentage of 60%. The land use in cluster 2 is 8% protected forest. Therefore, the sub-districts in cluster 2 are classified as prone to drought disaster.

Cluster 3

Sub-districts in cluster 3 include Purwosari, Girisubo, Saptosari, Patuk, Tepus, Tanjungsari, Gedangsari, and Panggang. These sub-districts are located in quite drought-prone areas. Table 6 shows that the rainfall was dominated by the vulnerable of 1500-2000 mm/year by 53%. The area is also dominated by Mediterranean soil types of 67%. The infiltration rate in cluster 3 has a 100% slow infiltration. The slope in this area is dominated by a moderate slope with the percentage of 47%. The land use in cluster 2 is 3% protected forest and 5% conservation forest. Therefore, the area in cluster 3 is classified as a quite prone to drought disaster.

Table 6. The percentage of each category of variables

Rainfall	*Cluster* 1		*Cluster* 2		*Cluster* 3	
	n	%	n	%	n	%
< 1000 mm/year	3	19	3	50	1	6
1000 - 1500 mm/year	7	44	3	50	6	40
1500 - 2000 mm/year	5	31	0	0	8	53
2000 - 2500 mm/year	1	6	0	0	0	0
Sum	16	100	6	100	15	100
Type of soil	*Cluster* 1		*Cluster* 2		*Cluster* 3	
	n	%	n	%	n	%
Mediteran	7	47	3	60	8	67
Latosol	1	7	2	40	4	33
Grumusol	7	47	0	0	0	0
Sum	15	100	5	100	12	100
Infiltration	*Cluster* 1		*Cluster* 2		*Cluster* 3	
	n	%	n	%	n	%
Slow	7	50	3	100	8	100
Medium	7	50	0	0	0	0
Sum	14	100	3	100	8	100
Slope	*Cluster* 1		*Cluster* 2		*Cluster* 3	
	n	%	n	%	n	%
Steep	2	15	1	20	4	27
Moderate	5	38	1	20	7	47
Gentle	6	46	3	60	4	27
Sum	13	100	5	100	15	100
Land use	*Cluster* 1		*Cluster* 2		*Cluster* 3	
	n	%	n	%	n	%
Settlement	7	23	3	23	8	21
Rainfed rice fields	0	0	0	0	4	11
Irrigation paddy	7	23	3	23	8	21
Garden	7	23	3	23	7	18
Production forest	5	16	3	23	8	21
Conservation forest	2	6	0	0	2	5
Protected forest	3	10	1	8	1	3
Sum	31	100	13	100	38	100

The results of the hazard category calculation and variable characteristics show that cluster 2 is the most dangerous level of vulnerability, followed by cluster 3, and cluster 1. Figure 4 is the clustering results as visualized by using a Geographic Information System (GIS). The Semin, Ngawen, and Nglipar sub-districts are particularly having high level of vulnerability. This result was in line with Sigit (2016a) that these three sub-districts were frequently affected by water scarcity, resulting in frequent droughts. The Purwosari, Girisubo, Saptosari, Patuk, Tepus, Tanjungsari, Gedangsari, and Panggang sub-districts have a fairly high level of vulnerability to drought. The Wonosari, Paliyan, Ponjong, Playen, Semanu, Rongkop, and Karangmojo sub-districts are the areas with low vulnerability to drought. A part of this result was supported by Fahmi (2016) that the Wonosari and Karangmojo sub-districts had good water absorption which might cause the low vulnerability level to drought.

Table 7. Comparison of fuzzy c-means method clustering results with drought news in 2016 and drought events in 2019

No	Sub-district	FCM	Category	Drought news in 2016	Drought events in 2019
1	Semin	2	High	No drought	Drought
2	Purwosari	3	Medium	Drought	Drought
3	Girisubo	3	Medium	Drought	Drought
4	Wonosari	1	Low	No drought	No drought
5	Paliyan	1	Low	No drought	Drought
6	Ponjong	1	Low	No drought	Drought
7	Playen	1	Low	No drought	No drought
8	Semanu	1	Low	No drought	Drought
9	Ngawen	2	High	Drought	Drought
10	Saptosari	3	Medium	No drought	Drought
11	Patuk	3	Medium	No drought	Drought
12	Tepus	3	Medium	Drought	Drought
13	Tanjungsari	3	Medium	No drought	Drought
14	Rongkop	1	Low	Drought	Drought
15	Nglipar	2	High	No drought	Drought
16	Gedangsari	3	Medium	No drought	Drought
17	Karangmojo	1	Low	No drought	Drought
18	Panggang	3	Medium	Drought	Drought

Source: 1. Drought news 2016 (Kurniawan, 2016) and (Ika, 2016)
2. Drought incident in 2019 from social service water dropping data (Dinas Sosial DIY, 2019)

Table 7 shows the differences in the results of FCM clustering with the news of drought events in 2016. The FCM method shows that Semin and Nglipar sub-districts have a high level of drought vulnerability, despite the fact that the news states that there is no drought (Kurniawan (2016) and (Ika, 2016)). The FCM method indicated that the sub-districts of Saptosari, Patuk, Tanjungsari, and Gedangsari have a medium level of vulnerability, but the news indicated that the areas did not experience drought. The FCM found that Rongkop sub-district has a low level of drought vulnerability, which contradicts the news, which states that the area is prone to drought. Table 7 also shows that the results of FCM clustering and drought news in 2016 have changed when compared to drought events based on social service dropping data. Some areas did not experience drought in 2016, but there was a demand for water dropping in 2019 and a drought occurred.

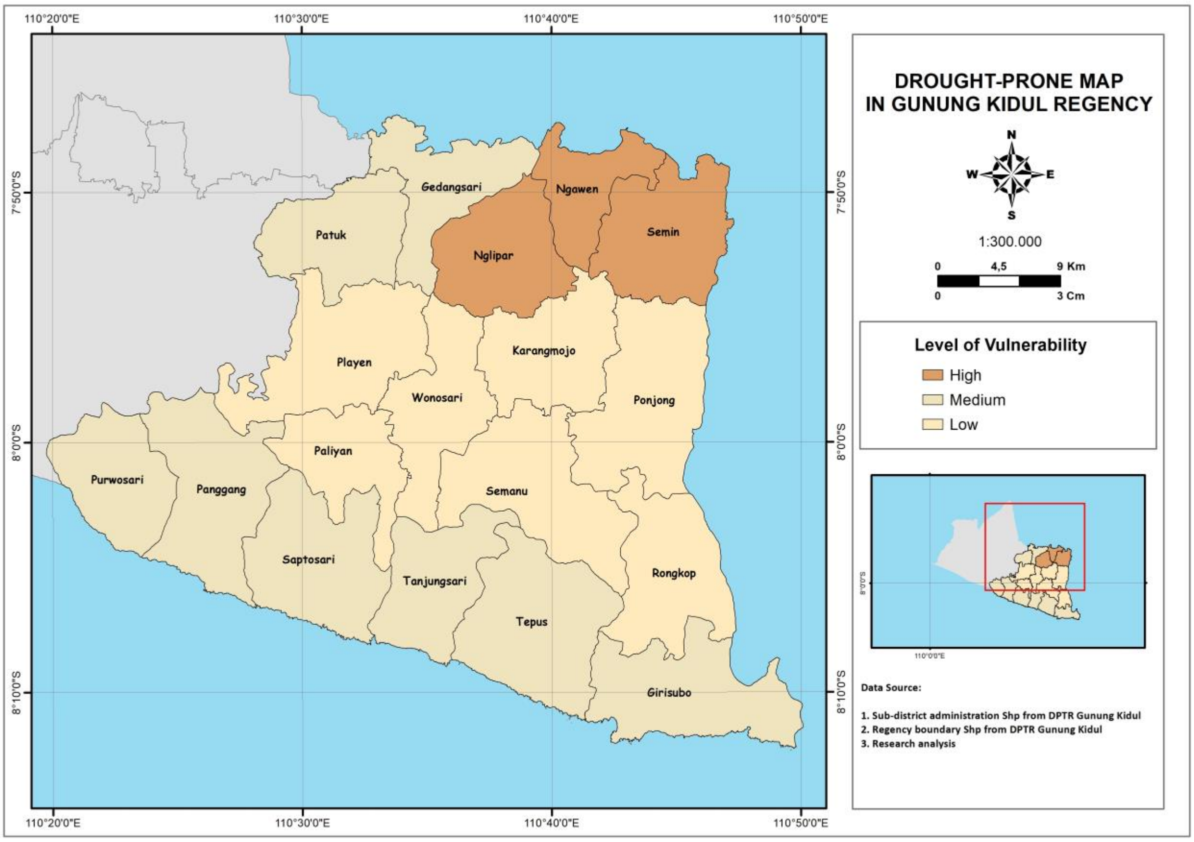

Figure 4. Visualization using a geographic information system (GIS)

CONCLUSION

Cluster validation on the Fuzzy C-Means (FCM) method yielded a k optimal of 3, so it is used to map the drought-prone sub-districts in Gunungkidul district. The hazard category in each cluster is determined by the cluster mean value and the variable characteristics. Cluster 1 has a low level of susceptibility to drought with 7 sub-districts of Wonosari, Playen, Semanu, Rongkop, Pojong, Karangmojo, and Paliyan. Cluster 2 has a high level of vulnerability consisting of 3 sub-districts of Semin, Ngawen, and Nglipar. Cluster 3 has a moderate level of vulnerability consisting of 8 districts of Gedangsari, Girisubo, Tepus, Saptosari, Panggang, Patuk, Purwosari, and Tanjungsari. However, the findings of this study do not rule out the possibility that the level of accuracy is still not optimal because drought was becoming increasingly difficult to predict. There are many factors that can cause drought, so it is hoped that more variables will be used in future research.

ACKNOWLEDGEMENTS

The authors acknowledge the Gunungkidul Land and Spatial Planning Agency for providing an SHP file for sub-district administration in Gunungkidul district, the Gunungkidul Central Statistics Agency for providing 2016 slope data, and the Land and Spatial Planning Agency of the Special Region of Yogyakarta for providing valid maps for 2018-2038 and SHP files of rainfall, soil types, infiltration, slope, and land use in 2016 sourced from the Development Planning Agency at Sub-National Level of Special Region of Yogyakarta. Finally, the authors would like to thank the anonymous reviewers and the editor of the journal for their valuable comments in improving the quality of the paper.

REFERENCES

Al-Bakri, J. T., Jameel, M., Naimat, A., Al-Karablieh, E., Al-Bakri Journal, J. T., Al-Bakri, J. T., Alnaimat, M. J., & Qaryouti, E. A. (2019). *Assessment of combined drought index and mapping of drought vulnerability in Jordan*. *9*(3), 59–68. http://www.ijera.com/papers/vol9no3/Series-1/G0903015968.pdf

Alam, M. S., & Paul, S. (2020). A comparative analysis of clustering algorithms to identify the homogeneous rainfall gauge stations of Bangladesh. *Journal of Applied Statistics*, *47*(8), 1460–1481. https://doi.org/10.1080/02664763.2019.1675606

Apriani, F., Setianingsih, Y. D., Arum, U. M. P., Susanti, K. A., Wicaksono, S. I., & Faruk, A. (2014). Analisis curah hujan sebagai upaya meminimalisasi dampak kekeringan di kabupaten Gunung Kidul tahun 2014. *Khazanah*, *6*(2), 13–22. https://doi.org/10.20885/khazanah.vol6.iss2.art2

Avia, L. Q., & Sofiati, I. (2018). Analysis of El Niño and IOD phenomenon 2015/2016 and their impact on rainfall variability in Indonesia. *IOP Conference Series: Earth and Environmental Science*, *166*(1). https://doi.org/10.1088/1755-1315/166/1/012034

Banjarnahor, N., Hindarto, K. S., & Fahrurrozi, F. (2018). Hubungan kelerengan dengan kadar air tanah, Ph tanah, dan penampilan jeruk gerga di kabupaten Lebong. *Jurnal Ilmu-Ilmu Pertanian Indonesia*, *20*(1), 13–18. https://doi.org/10.31186/jipi.20.1.13-18

Baudoin, M. A., Vogel, C., Nortje, K., & Naik, M. (2017). Living with drought in South Africa: lessons learnt from the recent El Niño drought period. *International Journal of Disaster Risk Reduction*, *23*, 128–137. https://doi.org/10.1016/J.IJDRR.2017.05.005

Bezdek, J. C. (1973). Cluster validity with fuzzy sets. *Journal of Cybernetics*, *3*(3), 58–73. https://doi.org/10.1080/01969727308546047

Bezdek, J. C. (1981). Pattern Recognition with Fuzzy Objective Function Algorithms. *Pattern Recognition with Fuzzy Objective Function Algorithms*. https://doi.org/10.1007/978-1-4757-0450-1

BPBDDIY. (2019). *Data dropping air di Daerah Istimewa Yogyakarta*.

BPS Gunungkidul. (2016). *Data kemiringan lereng*.

Cebeci, Z., Yildiz, F., Kavlak, A. T., Cebeci, C., & Onder, H. (2020). *Package "ppclust": Probabilistic and possibilistic cluster analysis*. CRAN. https://cran.r-project.org/web//packages/ppclust/ppclust.pdf

Chang, C., Andreanus, J., Chan, W., & Verdian, I. (2019). Aplikasi sistem informasi geografis berbasis web pemetaan lokasi tempat makan vegetarian di kota Batam. *Jurnal Telematika*, *13*(1), 55–60. https://journal.ithb.ac.id/telematika/article/view/215/pdf

Chiew, F. H. S., Piechota, T. C., Dracup, J. A., & McMahon, T. A. (1998). El Nino/Southern Oscillation and Australian rainfall, streamflow and drought: Links and potential for forecasting. *Journal of Hydrology*, *204*(1–4), 138–149. https://doi.org/10.1016/S0022-1694(97)00121-2

Costa, M. da S., Oliveira-Júnior, J. F. de, Santos, P. J. dos, Correia Filho, W. L. F., Gois, G. de, Blanco, C. J. C., Teodoro, P. E., Silva Junior, C. A. da, Santiago, D. de B., Souza, E. de O., & Jardim, A. M. da R. F. (2021). Rainfall extremes and drought in Northeast Brazil and its relationship with El Niño–Southern Oscillation. *International Journal of Climatology*, *41*(S1), E2111–E2135. https://doi.org/10.1002/JOC.6835

Darojati, N. W., Barus, B., & Sunarti, E. (2015). Pemantauan bahaya kekeringan di kabupaten Indramayu. *Jurnal Ilmu Tanah Dan Lingkungan*, *17*(2), 60–68. https://doi.org/10.29244/jitl.17.2.60-68

Dave, R. N. (1996). Validating fuzzy partitions obtained through c-shells clustering. *Pattern Recognition Letters*, *17*(6), 613–623. https://doi.org/10.1016/0167-8655(96)00026-8

Dinas Sosial DIY. (2019). *Data dropping air di Kabupaten Gunung Kidul*.

DPTRDIY. (2016). *Data shp curah hujan, jenis tanah, infiltrasi, dan penggunaan lahan*.

Ewbank, R., Perez, C., Cornish, H., Worku, M., & Woldetsadik, S. (2019). Building resilience to El Niño-related drought: experiences in early warning and early action from Nicaragua and Ethiopia. *Disasters*, *43*(S3), S345–S367. https://doi.org/10.1111/DISA.12340

Fahmi, H. H. (2016). *Analisis kondisi resapan air dengan menggunakan sistem informasi geografis di kabupaten Gunungkidul*. Universitas Muhammadiyah Surakarta.

Febtriko, A. (2017). Pemakaian mobile robot dalam meningkatkan perkembangan kognitif anak usia dini di taman kanak-kanak. *RABIT : Jurnal Teknologi Dan Sistem Informasi Univrab*, *2*(2), 125–135. https://doi.org/https://doi.org/10.36341/rabit.v2i2.215

Feng, P., Wang, B., Liu, D. L., & Yu, Q. (2019). Machine learning-based integration of remotely-sensed drought factors can improve the estimation of agricultural drought in South-Eastern Australia. *Agricultural Systems*, *173*(December 2018), 303–316. https://doi.org/10.1016/j.agsy.2019.03.015

Ferraro, M. B., Giordani, P., & Serafini, A. (2019). fclust: An R Package for Fuzzy Clustering. *The R Journal*, *11*(01), 1–18. https://journal.r-project.org/archive/2019/RJ-2019-017/RJ-2019-017.pdf

Gader, K., Gara, A., Vanclooster, M., Khlifi, S., & Slimani, M. (2020). Drought assessment in a south Mediterranean transboundary catchment. *Hydrological Sciences Journal*, *65*(8), 1300–1315. https://doi.org/10.1080/02626667.2020.1747621

Govender, P., & Sivakumar, V. (2020). Application of k-means and hierarchical clustering techniques for analysis of air pollution: A review (1980–2019). *Atmospheric Pollution Research*, *11*(1), 40–56. https://doi.org/10.1016/j.apr.2019.09.009

Habibie, M. I., Noguchi, R., Matsushita, S., & Ahamed, T. (2020). Development of micro-level classifiers from land suitability analysis for drought-prone areas in Indonesia. *Remote Sensing Applications: Society and Environment*, *20*(September), 1–14. https://doi.org/10.1016/j.rsase.2020.100421

Halwatura, D., McIntyre, N., Lechner, A. M., & Arnold, S. (2017). Capability of meteorological drought indices for detecting soil moisture droughts. *Journal of Hydrology: Regional Studies*, *12*, 396–412. https://doi.org/10.1016/J.EJRH.2017.06.0012

Herawaty, N. (2014). Faktor-faktor yang mempengaruhi akuntabilitas kinerja instansi pemerintah daerah kota Jambi. *Jurnal Cakrawala Akuntansi*, *6*(2), 151–161.

Hidayati, I. N., & Suryanto, S. (2015). Pengaruh perubahan iklim terhadap produksi pertanian dan strategi adaptasi pada lahan rawan kekeringan. *Jurnal Ekonomi & Studi Pembangunan.*, *16*(1), 42–52. https://journal.umy.ac.id/index.php/esp/article/view/1217

Ika. (2016). *Water project, solusi kekeringan di Gunungkidul*. Universitas Gadjah Mada. https://ugm.ac.id/id/berita/12395-water-project-solusi-kekeringan-di-gunungkidul

Kassambara, A., & Mundt, F. (2020). *Package "factoextra": Extract and visualize the results of multivariate data analyses*. CRAN. https://github.com/kassambara/factoextra/issues

Khasanah, I. U., & Sastra, A. R. (2017). Pengaruh fenomena El-Nino dan La-Nina terhadap perairan Sumatera Barat. *Seminar Nasional Penginderaan Jauh Ke-4*, 499–510. http://sinasinderaja.lapan.go.id/files/sinasja2017/prosiding/59_PENGARUH FENOMENA EL-NINO DAN LA-NINA TERHADAP PERAIRAN SUMATERA BARAT.pdf

Kurniawan, D. (2016). *Kekeringan gunungkidul: lima kecamatan gunungkidul dilanda krisis air*. Solopos.Com. https://www.solopos.com/kekeringan-gunungkidul-lima-kecamatan-gunungkidul-dilanda-krisis-air-753608

Kusumadewi, S., & Purnomo, H. (2010). *Aplikasi logika fuzzy untuk pendukung keputusan*. Graha Ilmu.

Leitold, V., Morton, D. C., Longo, M., Nara Dos-Santos, M., Keller, M., & Scaranello, M. (2018). El Niño drought increased canopy turnover in Amazon forests. *New Phytologist*, *219*(3), 959–971. https://doi.org/10.1111/nph.15110

Maechler, M., Rousseeuw, P., Struyf, A., Hubert, M., & Hornik, K. (2021). cluster: Cluster analysis basics and extensions. *R Package Version 2.1.2*. https://cran.r-project.org/package=cluster

Maranell, G. (2017). *Scaling: A sourcebook for behavioral scientists*. Routledge.

Mardiana, S., Widhiastuti, R., & Erningpraja, L. (2020). Management and employees perception analysis on sugar industry waste management based on cleaner production. *Britain International of Exact Sciences Journal*, *2*(1), 442–449. https://doi.org/https://doi.org/10.33258/bioex.v2i1.201

Moghbeli, A., Delbari, M., & Amiri, M. (2020). Application of a standardized precipitation index for mapping drought severity in an arid climate region, southeastern Iran. *Arabian Journal of Geosciences*, *13*(5), 1–16. https://doi.org/10.1007/S12517-020-5201-7

Moradi, A. M., & Dariane, A. B. (2015). Evolving neural networks and fuzzy clustering for multireservoir operations. *Neural Computing and Applications*, *28*(5), 1149–1162. https://doi.org/10.1007/S00521-015-2130-6

Muharsyah, R., & Ratri, D. N. (2015). Durasi dan kekuatan kekeringan menggunakan indeks hujan terstandarisasi di pulau Bali. *Jurnal Meteorologi dan Geofisika*, *16*(2), 93–104. https://doi.org/10.31172/jmg.v16i2.272

Nabilah, F., Prasetyo, Y., & Sukmono, A. (2017). Analisis pengaruh fenomena El Nino dan La Nina terhadap curah hujan tahun 1998 - 2016 menggunakan indikator Oni (Oceanic Nino Index) (Studi kasus : Provinsi Jawa Barat). *Jurnal Geodesi Undip*, *6*(4), 402–412. https://ejournal3.undip.ac.id/index.php/geodesi/article/view/18170

Ningsih, S., & Dukalang, H. H. (2019). Penerapan metode suksesif interval pada analisis regresi linier berganda. *Jambura Journal of Mathematics*, *1*(1), 43–53. https://doi.org/10.34312/jjom.v1i1.1742

Nishom, M. (2019). Perbandingan akurasi Euclidean distance, Minkowski distance, dan Manhattan distance pada algoritma K-means clustering berbasis chi-square. *Jurnal Informatika: Jurnal Pengembangan IT*, *4*(1), 20–24. https://doi.org/10.30591/jpit.v4i1.1253

Nugroho, J., Zid, M., & Miarsyah, M. (2020). Potensi sumber air dan kearifan masyarakat dalam menghadapi risiko kekeringan di wilayah karst (Kabupaten Gunung Kidul, Provinsi Yogyakarta). *Jurnal Pengelolaan Lingkungan Berkelanjutan (Journal of Environmental Sustainability Management)*, *4*(1), 438–447. https://doi.org/10.36813/jplb.4.1.438-447

Omran, M. G. H., Engelbrecht, A. P., & Salman, A. (2007). An overview of clustering methods. *Intelligent Data Analysis*, *11*(6), 583–605. https://doi.org/10.3233/ida-2007-11602

Paparrizos, S., Maris, F., Weiler, M., & Matzarakis, A. (2018). Analysis and mapping of present and future drought conditions over Greek areas with different climate conditions. *Theoretical and Applied Climatology*, *131*(1–2), 259–270. https://doi.org/10.1007/s00704-016-1964-x

Pawitan, H. (2014). *Land use changes and their impacts on watershed hydrology*. https://www.researchgate.net/publication/237486643_PERUBAHAN_PENGGUNAAN_LAHAN_DAN_PENGARUHNYA_TERHADAP_HIDROLOGI_DAERAH_ALIRAN_SUNGAI_Land_Use_Changes_and_Their_Impacts_on_Watershed_Hydrology/link/0a85e53b54d92e7a0c000000/download

Pertiwi, A. P., & Kurniawan, R. (2017). Analisis pemetaan risiko bencana banjir di Indonesia tahun 2011 – 2015 menggunakan fuzzy c-means. *SEMINAR MATEMATIKA DAN PENDIDIKAN MATEMATIKA UNY 2017*, 97–104. https://doi.org/10.31227/OSF.IO/5KDVU

Pimentel, B. A., & de Souza, R. M. C. R. (2016). Multivariate Fuzzy C-Means algorithms with weighting. *Neurocomputing*, *174*, 946–965. https://doi.org/10.1016/j.neucom.2015.10.011

Rahakbauw, D. L., Ilwaru, V. Y. I., & Hahury, M. H. (2017). Implementasi fuzzy C-means clustering dalam penentuan beasiswa. *Barekeng: Jurnal Ilmu Matematika dan Terapan*, *11*(1), 1–11. https://www.neliti.com/id/publications/277582/implementasi-fuzzy-c-means-clustering-dalam-penentuan-beasiswa

Ramadhan, E. R., Sutoyo, E., Musnansyah, A., & Belgaman, H. A. (2020). Analysis of hotspot data for drought clustering using K-means algorithm. *ICONETSI: Proceedings of the International Conference on Engineering and Information Technology for Sustainable Industry*, 1–6. https://doi.org/10.1145/3429789.3429824

Raykov, Y. P., Boukouvalas, A., Baig, F., & Little, M. A. (2016). What to do when K-means clustering fails: A simple yet principled alternative algorithm. *PLoS ONE*, *11*(9), 1–28. https://doi.org/10.1371/journal.pone.0162259

Revelle, W. (2021). psych: Procedures for psychological, psychometric, and personality research. In *R package version 2.1.9*. https://cran.r-project.org/package=psych

Rizal, A. S., & Hakim, R. B. F. (2015). Metode k-means cluster dan fuzzy c-means cluster: Studi kasus indeks pembangunan manusia di kawasan Indonesia timur tahun 2012. *Prosiding Seminar Nasional Matematika dan Pendidikan Matematika UMS*, 643–657. https://publikasiilmiah.ums.ac.id/handle/11617/5803

Sharma, M., & Kamal Borana, M. (2014). Clustering in data mining : A brief review. *International Journal Of Core Engineering & Management (IJCEM)*, *1*(5), 117–126.

Sigit, A. (2016a). *Dinsos Gunungkidul mulai droping air*. Krjogja.Com. https://www.krjogja.com/berita-lokal/diy/gunungkidul/dinsos-gunungkidul-mulai-droping-air/

Sigit, A. A. (2016b). Analisis spasial kemampuan infiltrasi sebagai bagian dari indikasi bencana kekeringan hidrologis di DAS Wedi, kabupaten Klaten-Boyolali. *Prosiding Seminar Nasional Geografi UMS*, 102–111. https://publikasiilmiah.ums.ac.id/handle/11617/8224

Simhachalam, B., & Ganesan, G. (2015). Fuzzy clustering algorithms in medical diagnostics. *Wulfenia*, *22*(7), 308–317.https://www.researchgate.net/publication/279806430_Fuzzy_Clustering_Algorithms_in_Medical_Diagnostics

Sofiani, D., Maulida, A. S., Fadhillah, N., & Sihite, D. Y. (2017). Gender differences in students' attitude towards science. *IOP Conf. Series: Journal of Physics: Conf. Series*, *895*, 12168. https://doi.org/10.1088/1742-6596/895/1/012168

Surmaini, E. (2016). Pemantauan dan peringatan dini kekeringan pertanian di Indonesia. *Jurnal Sumberdaya Lahan*, *10*(1), 37–50. http://ejurnal.litbang.pertanian.go.id/index.php/jsl/article/view/6320

Wickham, H., François, R., Henry, L., & Müller, K. (2021). "dplyr": A grammar of data manipulation. *CRAN*. https://cran.r-project.org/web/packages/dplyr/index.html

Xie, X. L., & Beni, G. (1991). A validity measure for fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *13*(8), 841–847. https://ieeexplore.ieee.org/document/85677

Xulu, S., Peerbhay, K., Gebreslasie, M., & Ismail, R. (2019). Unsupervised clustering of forest response to drought dtress in Zululand region, South Africa. *Forests*, *10*(531), 1–15. https://doi.org/10.3390/f10070531

Yang, M. S., & Nataliani, Y. (2017). Robust-learning fuzzy c-means clustering algorithm with unknown number of clusters. *Pattern Recognition*, *71*, 45–59. https://doi.org/10.1016/J.PATCOG.2017.05.017