Advances in Dynamical Systems and Applications [ADSA]
ISSN 2250-3234 Volume 16, Number 2 (2021), pp. 639-656

http://www.ripublication.com

Landslide Susceptibility Mapping using Fuzzy C-Means and Fuzzy C-Modes Methods in Wonosobo Regency, Indonesia

Dhoriva Urwatul Wutsqa[1], Erdina Kurniawati[2], Sri Andayani[2], and Kismiantini[1]

[1]Department of Statistics, Universitas Negeri Yogyakarta, Indonesia
[2]Department of Mathematics, Universitas Negeri Yogyakarta, Indonesia.

Abstract

Wonosobo has a topography that is susceptible to a landslide. Mapping the level of landslide susceptibility is essential to provide information on areas with a high level of landslide susceptibility. This study aims to map the level of landslide susceptibility in Wonosobo Regency from the clustering results. The variables used as the basis for clustering are slope, rainfall, land use, and soil type. The original data of these variables are numerical but can be categorized based on the susceptibility level. Clustering with numerical data uses the fuzzy C-means (FCM) method, and categorical data is processed using the fuzzy C-modes (FCMO) method. FCMO has an algorithm that is almost the same as FCM. The main differences are in the dissimilarity measurement and the mode use as the cluster center instead of the cluster mean. Two cluster validation indices are used to determine the optimal number of clusters, namely the Xie Beni index and the modified partition coefficient index. The validity index results indicate that the best cluster is yielded by FCMO method with 4 clusters. Thus, the landslide susceptibility mapping in Wonosobo can be divided into 4 levels, namely not susceptible, slightly susceptible, highly susceptible.

Keywords: Wonosobo Regency**,** landslide**,** fuzzy C-means, fuzzy C-modes

1 INTRODUCTION

Indonesia is a tropical country with very high rainfall. Most areas in Indonesia experience very heavy rains for a long duration during the rainy season. Wonosobo Regency is one of 35 regencies/cities in Central Java Province with a rainfall level above 2500 mm/year, so it is included in the very high rainfall category. Wonosobo

Regency is a mountainous area with an altitude ranging from 275 meters to 2250 meters above sea level. Most areas in Wonosobo Regency have a slope of 15 - 40%. In terms of its geological structure, Wonosobo Regency is a young mountain type located in breakwater rocks that often experience natural disasters, especially in the rainy season such as landslides, collapsed soil movements, and creep mass movements. Based on the published data from the Regional Disaster Management Agency of Wonosobo Regency, there were 136 landslides in 2017 [1] and 49 incidents in 2018 [2].

The high risk for landslides is also triggered by soil types and land use in Wonosobo Regency. Soil types in Wonosobo Regency are mostly andosol and regosol. These two types of soil are categorized as susceptible and highly susceptible to landslides. Meanwhile, the land use can be divided into 12 categories, namely technical irrigated rice fields, semi-technical irrigated rice fields, simple irrigated rice fields, rainfed rice fields, buildings with yards, moorlands, ponds, reservoirs, state forest, community forest, plantations, and other uses. The largest area is used for moor. The moor in Wonosobo Regency is employed for the cultivation of potatoes, vegetables, and tobacco. Moor land, especially in areas with high slopes, causes a high risk of landslides in some areas in the Wonosobo Regency. Several studies have shown that the variables of rainfall, land slope, land use, and soil type [3-10] influence the occurrence of landslides.

Landslides cause economic and environmental losses and even threaten lives. Mapping the level of landslides susceptibility is very important to reduce losses and become a consideration for land management decision-makers. Landslide maps are created based on clustering results. The data on the variables for the landslide clustering are the areas of the variables dealing with the fuzzy numbers since they are not measured in an exact way but by approximation. Besides that, each object or the sub-district could belong in more than one cluster with different degrees of memberships. For those reasons, we choose FCM for landslide clustering. The FCM is a clustering method that is quite popular. Several studies related to the application of FCM in various studies have been carried out, including for data clustering of meteorological drought in India [11], urban air pollution in Liverpool, England [12], the status of bankruptcy and non-bankruptcy in India [13], and soil taxonomic units [14], hydrology [15], and soil map in Iran [16]. The applications of FCM have also been reported for landslide in Taiwan [17] and in Columbia [18], In FCM, the existence of each data point in a cluster is determined by the degree of membership. FCM provides a definite membership value for all clusters. This method has a relatively lighter computational load, so clustering can be done quickly. Even so, it relatively depends on a large amount of data and the number of desired clusters. The FCM has been revealed more superior than traditional methods [19-20] and a flexible method [21].

The variables used in this study are rainfall, soil slope, land use, and soil type. Each variable used in landslide clustering contains several types. For example, soil type consists of latosol, mediterranean, andosol, and regosol. The data of each sub-district except rainfall are the areas of entire types included in the variables, so the FCM calculation involved a greater number of attributes. If each data in the sub-district is categorized by determining which type is the most dominant in each variable, then the number of the attributes used in the FCM process will be equal to the number of variables. Certainly, with fewer attributes, the calculations will be simpler and more efficient. Therefore, we also propose clustering by using variables with category values. The version of FCM for categorical data is fuzzy c-modes (FCMO), which is introduced by [22]. Fuzzy C-modes is an expansion of the fuzzy C-means method. It uses modes in the matching process of the dissimilarity measure to update the cluster center point, to obtain the optimal clusters. There are very few studies on the application of FCMO. The application of the FCMO method includes the soybean disease database [23-24], and in a supplier segmentation problem on automobile parts suppliers in Taiwan [25]. The application of FCMO to map disasters such as landslides, floods, and earthquakes have not yet received researchers' attention.

This research aims to map landslide-susceptible areas in Wonosobo Regency by considering both numerical and categorical variables. The available data are numeric. However, the data can be categorized based on the level of susceptibility to landslides, so that they become categorical type. The method used for numerical data clustering is FCM and categorical data is FCMO. The clustering results of both methods are compared based on their clustering validity indices. From the clustering results, a map of landslide-susceptible areas in Wonosobo Regency is created. The contribution of this study is the use of the FCMO method for landslide clustering. We also propose the determination of the landslide susceptibility level on the clustering results and describe the characteristics of each cluster.

2 RESEARCH METHOD

2. 1 Data and Analysis

The data used in this study included variables as factors causing landslides, namely slope, rainfall, land use, and soil type. Each variable is divided into several categories based on the level of susceptibility to landslides. Related to the slope variable, the higher the land slope, the more susceptible it is to landslides. The slope category is divided into five from 0% to 40%. The category of slope susceptible to landslides can be seen in Table 1.

Table 1. Category of landslide susceptibility based on slope

No.	Slope (%)	Category
1.	0 - 8	Not susceptible
2.	8 - 15	Slightly susceptible
3.	15 - 25	Moderately susceptible
4.	25 - 40	Susceptible
5.	> 40	Highly susceptible

In the category of land slope in the Wonosobo Regency, the land slope data from 0-8% are 18,201.07 hectares spread over 14 sub-districts (except Watumalang Sub-District) with a not susceptible level to erosion. The land slope of 8-15% covering an area of 28,846.69 hectares is spread over 14 sub-districts (except Watumalang Sub-District) with a slightly susceptible level of erosion. The land slope of 15-25% with an area of 39,044.84 hectares is spread across all sub-districts in Wonosobo Regency with a moderately susceptible level to erosion. The land slope of 25-40% covering an area of 6,241,133 hectares has spread across 11 sub-districts in the Wonosobo Regency with a susceptible level of erosion. The land slope of > 40% with an area of 4,019,312 hectares is spread over ten sub-districts with a highly susceptible level of erosion.

High rainfall can cause landslides. Wadaslintang Sub-District has a 1,000-2,500 mm rainfall per year, which is considered susceptible to erosion. Meanwhile, the rest of the sub-districts have more than 2,500 mm rainfall per year, which is classified as highly susceptible to erosion. Landslide susceptibility based on the rainfall category is presented in Table 2.

Table 2. Category of landslide susceptibility based on rainfall

No.	Rainfall	Category
1.	<1,000 mm/year	Not susceptible
2.	1,000-2,500 mm/year	Susceptible
3.	>2,500 mm/year	Highly susceptible

The types of land use can be divided into four categories based on their level of susceptibility to erosion. The more susceptible to erosion, the higher the potential for landslides to occur. There are seven categories of land use in Wonosobo Regency, such as moor, rice fields, shrubs, fields, plantations, forests, and waters (reservoirs and ponds). This land uses has various susceptibility levels to erosion from not susceptible, slightly susceptible, susceptible, and highly susceptible. The category of landslides susceptibility based on land use can be seen in Table 3.

Table 3. Category of landslides susceptibility based on land use

No.	Land use	Category
1.	Waters, fields, plantation areas, forests	Not susceptible
2.	Shrubs	Slightly susceptible
3.	Rice fields	Susceptible
4.	Moor	Highly susceptible

Soil types that easily absorb water can cause landslides. The category of landslide susceptibility based on soil types is presented in Table 4. Particularly in Wonosobo Regency, there are four soil types, namely latosol, mediterranean, andosol, and regosol. The area with latosol soil type includes three sub-districts, namely Kepil, Kalibawang, and Selomerto, with slightly susceptible to a landslide. Mediterranean soil is spread over three sub-districts in Wonosobo Regency, namely Kepil, Sapuran, and Kertek, with moderately susceptible to a landslide. Andosol soil covers 8 sub-districts in Wonosobo Regency which is susceptible to the landslide. The regosol soil is spread over 13 sub-districts in Wonosobo Regency with a highly susceptible to landslide.

.

Table 4. Category of landslide susceptibility based on soil types

No.	Soil Type	Category
1.	Alluvial, hydromorphous gray glei planosol soil, groundwater literita	Not susceptible
2.	Latosol	Slightly susceptible
3.	Brown forest soil, non-calcis brown, mediterranean	Moderately Susceptible
4.	Andosol, laterite, grumusol, podsol, podzolic	Susceptible
5.	Regosol, lithosol, organosol, renzina	Highly susceptible

The slopes, land use, soil types, and rainfall (in mm/year) are referred to as research attributes or input variables in the cluster. The object of this study is the 15 sub-districts in Wonosobo Regency. The data including rainfall, land use, and soil types are taken from the website of the Central Statistics Agency of Wonosobo Regency with the following address https://wonosobokab.bps.go.id/ [26]. Meanwhile, the land slope data is taken from the Ina-geoportal website (https://tanahair.indonesia.go.id/portal-web). The numerical data are in the form of the area of the research attributes except for rainfall. Next, the data are the amount of rainfall in mm/year. Furthermore, the numerical data are processed as inputs to the FCM clustering method according to the FCM algorithm steps. For the FCMO method, the data are the categorization results of all attributes. The results of the categorization are taken from the data of the most dominant area from each sub-district in the intended attribute, then categorized based on the level of susceptibility.

In both methods, the number of clusters is predefined in the early steps of the algorithm. To determine the optimal number of clusters, two cluster validation indices are used, namely the Xie Beni index (XBI) and the modified partition coefficient (MPC). The clustering results of the two methods were compared based on the three indices, which one was better. The fuzzy C-means clustering method, fuzzy C-Modes, and cluster validity index are explained in the following sub-section.

2. 2 Fuzzy C-Means

Fuzzy clustering is a method of analysis by considering the degree of membership that includes fuzzy sets as a weighting basis for grouping [27]. This is a development method of partitioning data with fuzzy weighting. The main advantage of fuzzy clustering is that it can provide clustering results for objects that are distributed irregularly. If there are data whose distribution is irregular, there is a possibility that a data point has characteristics from other clusters, so it is necessary to weigh the tendency of data points to a cluster.

In the FCM method, the number of clusters is initially determined before calculations. The determination of the cluster center is done iteratively. The cluster center that is still inaccurate at the initial conditions and the degree of membership of each object will be repaired repeatedly so that the cluster center will move to the right location. This iteration is based on minimizing the objective function that describes the distance from a given object to the cluster center weighted by the degree of membership of that object.

Suppose the data matrix X= $[X_{ij}]$ has the size $n \times p$, n is the number of observations to be clustered and p is the number of variables, X_{ij} is the i-th observation (i = 1, 2, …, n) on the j-th variable (j= 1, 2, …, p), c is the number of cluster, V_{ki} is the center of k-th cluster of j-th variable, μ_{ik} is the membership function of the i-th observation and k-th cluster, m is the weighted power of the membership function and is called the fuzzifier, d_{ik} is the Euclidean distance between the i-th observation and the k-th cluster center. The minimized objective function for updating the membership function and cluster center is

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

where

$$0 \leq \mu_{ik} \leq 1 , \sum_{k=1}^{c} \mu_{ik} = 1 , 0 < \sum_{i=1}^{n} \mu_{ik} \leq n , i = 1, 2, \ldots, n; k = 1, 2, \ldots, c.$$

The complete FCM algorithm is as follows:

Step 1. Input data as matrix X with the size of $n \times p$

Step 2. Determine:

a. Number of clusters that will be generated c ($2 \leq c \leq n$),
b. Fuzzifier $m > 1$,
c. Maximum iteration number,
d. Initial iteration $t = 1$,
e. Smallest error ((ε = very small positive value).

Step 3. Determine partition matrix from membership degree of each data in initial cluster U randomly

$$U^{(0)} = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_1) & \cdots & \mu_{1c}(x_1) \\ \mu_{21}(x_2) & \mu_{22}(x_2) & \cdots & \mu_{2c}(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ \mu_{n1}(x_n) & \mu_{n2}(x_n) & \cdots & \mu_{nc}(x_n) \end{bmatrix} \tag{2}$$

Step 4. At step t, t = 0, 1, 2, …Calculate k-th cluster center on j-th variable, V_{kj} with the formula

$$V_{kj} = \frac{\sum_{i=1}^{n} (\mu_{ik})^m X_{ij}}{\sum_{i=1}^{n} (\mu_{ik})^m} \tag{3}$$

Step 5. Repair the membership degree $U^{(t)}$; Calculate the membership degree of each object at each cluster (repair partition matrix $U^{(t)}$). For i = 1 to n

a. Calculate I_i and $\overline{I_i}$:

$$I_i = \left\{ k | 1 \leq k \leq c, \ d_{ik} = \left[\sum_{j=1}^{p} \left(X_{ij} - V_{kj} \right)^2 \right]^{\frac{1}{2}} = 0 \right\}$$

$$\overline{I_i} = \{1, 2, \ldots, c\} - I_i$$

b. Calculate membership degree of *i*-th object of *k*-th cluster

(i) If $I_i = \emptyset$,

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{p} (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{m-1}}}{\sum_{k=1}^{c} \left[\sum_{j=1}^{p} (X_{ij} - V_{kj})^2 \right]^{\frac{-1}{m-1}}} \tag{4}$$

(ii) If $I_i \neq \emptyset$, $\mu_{ik} = 0$ for all $k \in \overline{I_i}$ and $\sum_{k \in I_i} \mu_{ik} = 1$; Next *i*.

Step 6. Determine the criteria of iteration termination by comparing matrix norm $U^{(t)}$with $U^{(t+1)}$.

If $\left\| U^{(t)} - U^{(t+1)} \right\| < \varepsilon$ or $t > MaxIter$, then terminate. If it is not terminated, then repeat step 4 at $t = t + 1$.

2. 3 Fuzzy C-Modes

Fuzzy C-modes algorithm is an extension of the hard C-modes algorithm and it was introduced in order to incorporate the idea of fuzziness and uncertainty in datasets [28]. Fuzzy C-means method performs clusterization using the mean as the cluster center so that it is only suitable for numerical data. For categorical data, the fuzzy c-modes (FCMO) method has been developed. The FCMO method is a modification of the FCM method with a simpler dissimilarity measure for categorical data and uses modes to replace mean in the FCM method.

Fuzzy C-modes was first introduced by Huang dan Ng [22]. It has a similar framework with FCM in generating partition matrices. The main difference between the two methods is in the definition of the distances specified for categorical data. In FCMO, the distance between *i*-th object $(X_{i1}, X_{i2}, \ldots, X_{ip})$ and the *k*-th cluster center $(V_{k1}, V_{k2}, \ldots, V_{kp})$ is defined as

$$d_{ik} = \sum_{j=1}^{p} \delta(X_{ij}, V_{kj}) \tag{5}$$

with

$$\delta(X_{ij}, V_{kj}) = \begin{cases} 1, & X_{ij} \neq V_{kj} \\ 0, & \text{otherwise} \end{cases}$$

Based on distance (5), the objective function is defined at FCM analogously with the FCM method as

$$F_\alpha(\boldsymbol{U}, \boldsymbol{V}) = \sum_{k=1}^{c} \sum_{i=1}^{n} (\mu_{ik})^m \, d_{ik} \tag{6}$$

Updating the fuzzy partition matrix that minimizes function (6) is as follows

$$\mu_{ik} = \begin{cases} 1, & \text{if } \boldsymbol{X}_i = \boldsymbol{V}_{\boldsymbol{k}}; \\ 0, & \text{if } \boldsymbol{X}_i = \boldsymbol{V}_{\boldsymbol{k'}}, k \neq k'; \\ \dfrac{\left[\sum_{j=1}^{p}(x_{ij}-V_{kj})^2\right]^{\frac{-1}{m-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{p}(x_{ij}-V_{kj})^2\right]^{\frac{-1}{m-1}}} & \text{, otherwise} \end{cases} \quad (7)$$

In the computational process, the FCMO algorithm is, in principle, analogous to the FCM algorithm, with differences in the mode as a substitute for the mean, distance function (5), and the membership function (7).

2. 4 Cluster Validation

The cluster validity index can be used as the criteria to determine the optimal number of clusters. Cluster validation is required to measure clustering accuracy. Measurement to determine the quality of fuzzy clusters must be carried out using the appropriate cluster validity and based on to the respective criteria. Some of the measurements that can be used for cluster validation are as follows.

a. Xie Beni index (XBI) aims to calculate the ratio of the total variation within the cluster and cluster separation. A low XBI value indicates good cluster results [29]. This index is written as follows,

$$XBI(c) = \frac{\sum_{i=1}^{n}\sum_{k=1}^{c}(\mu_{ik})^m \, d_{ik}{}^2(x_k - v_i)}{n \, min_{ik} d_{ik}{}^2(v_i - v_k)} \quad (8)$$

where μ_{ik} is membership value, m is the weighted power, $d_{ik}{}^2(x_k - v_i)$ is the square of the distance between the object to the center of the cluster and $d_{ik}{}^2(v_i - v_k)$ is the square of the distance between cluster centers, *n* is the number of objects.

b. Partition coefficient index (PCI) was proposed by Bezdek in 1981. The PCI values are within [0, 1]. The greater the value (closer to 1), the better the cluster results. The PC index is written as following [28]

$$PCI(c) = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{c}\mu_{ik}{}^2 \quad (9)$$

c. Modified partition coefficient index (MPCI) proposed by Dave in 1996 reduces the tendency of PCI to change monotonously as the number of clusters changes. This index has the values within [0, 1]. The greater the value (closer to 1), the better the cluster results. The MPCI index is written as [30]

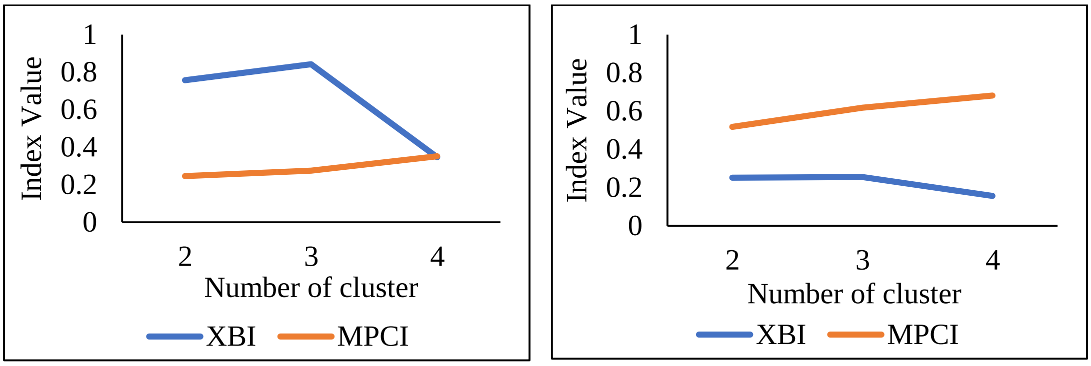
$$MPCI(c) = 1 - \frac{c}{c-1}(1 - PC) \quad (10)$$

In this study, only two validity indices are used to find and compare the best clustering results, namely XBI and MPCI. The selection of the MPCI is to monotonously anticipate changes in value as the number of clusters changes.

3 RESULTS AND ANALYSIS

3. 1 Validity Index Results

The FCM and FCMO methods are used to analyze landslide disaster data in 15 sub-districts in Wonosobo Regency, including variables of slope, rainfall, land use, and soil type. The data in Wonosobo Regency contain five levels of the slope, seven types of land use, four types of soil, and rainfall data (mm/year), so there are 17 research attributes. In the clustering calculation process, the attribute is stated as variables. The numerical data are analyzed using FCM, and categorical data are analyzed using FCMO. The optimal number of clusters is determined using the Xie Beni index and the modified partition coefficient index. The results of XBI and MPCI with clusters 2, 3, and 4 for FCM and FCMO are presented in Figure 1.



a b

Figure 1. XBI and MPCI values in a. FCM method and b. FCMO method

Figure 1.a shows that the minimum value on XBI and PMCI results from clustering using FCM is resulted in the formation of 4 clusters with the values of 0.3454865 and 0.351425. Therefore, the best cluster for the two indices of clustering results using FCM is the formation of 4 clusters. Figure 1.b shows that the minimum values of XBI and MPCI of the clustering using FCMO have resulted in the formation of 4 clusters with the values of 0.156127 and 0.68104. Therefore, the best cluster for the two indices of the clustering results using FCMO is the formation of 4 clusters. In other words, the FCMO method is better than the FCM method. It can be seen from the lower XBI value and higher MPCI in FCMO when compared to those in FCM.

3. 2 Clustering Results using FCM Method

In FCM, the clustering results are membership degrees of the sub-districts in each cluster formed. The objects are categorized in the cluster with the largest degree of membership. Table 5 provides clustering results with 4 clusters, according to the optimal clustering results. It can be seen that of the 15 sub-districts in Wonosobo Regency, there are two sub-districts, namely Kalikajar and Kertek, which are included in Cluster 1; three sub-districts namely Sukoharjo, Garung, and Kejajar are included in Cluster 3; three sub-districts are included in Cluster 4, namely Wadaslintang, Kepil, and Sapuran, and the rest is in Cluster 2.

Table 5. Best clustering results using FCM

No	Sub-district	Membership degree				Cluster
		Cluster 1	Cluster 2	Cluster 3	Cluster 4	
1	Wadaslintang	0.1749	0.2329	0.2788	**0.3134**	4
2	Kepil	0.1198	0.1200	0.0645	**0.6957**	4
3	Sapuran	0.1830	0.1438	0.0762	**0.5971**	4
4	Kalibawang	0.1505	**0.6914**	0.0609	0.0973	2
5	Kaliwiro	0.2597	**0.3175**	0.123	0.2998	2
6	Leksono	0.0474	**0.8822**	0.0333	0.037	2
7	Sukoharjo	0.0397	0.0908	**0.8184**	0.0511	3
8	Selomerto	0.2045	**0.3403**	0.2307	0.2245	2
9	Kalikajar	**0.7383**	0.1181	0.0428	0.1007	1
10	Kertek	**0.9597**	0.0217	0.0051	0.0135	1
11	Wonosobo	0.2485	**0.4945**	0.1175	0.1395	2
12	Watumalang	0.1779	**0.4261**	0.2166	0.1793	2
13	Mojotengah	0.1163	**0.7838**	0.0404	0.0594	2
14	Garung	0.0502	0.0962	**0.7901**	0.0636	3
15	Kejajar	0.0610	0.1054	**0.7602**	0.0734	3

The categories in the cluster do not refer to the level of landslide susceptibility. As a determinant of the level of landslide susceptibility in Wonosobo Regency, the FCM method calculates the percentage of the area of all sub-districts in a cluster on the input variable. The highest percentage result is considered to be the variable that characterizes or dominates the cluster. Then, these variables are adjusted for the susceptibility level based on Table 1, Table 2, Table 3, and Table 4. Figure 2 is the categorization result of the susceptibility level of input data, which has the highest percentage.

For example, Cluster 1 with a slope variable has a value of 2, meaning that a slope of 8-15% dominates the area in Cluster 1 with a slightly susceptibility of landslide.The order of factors that influence the occurrence of landslides is the slope of the soil with

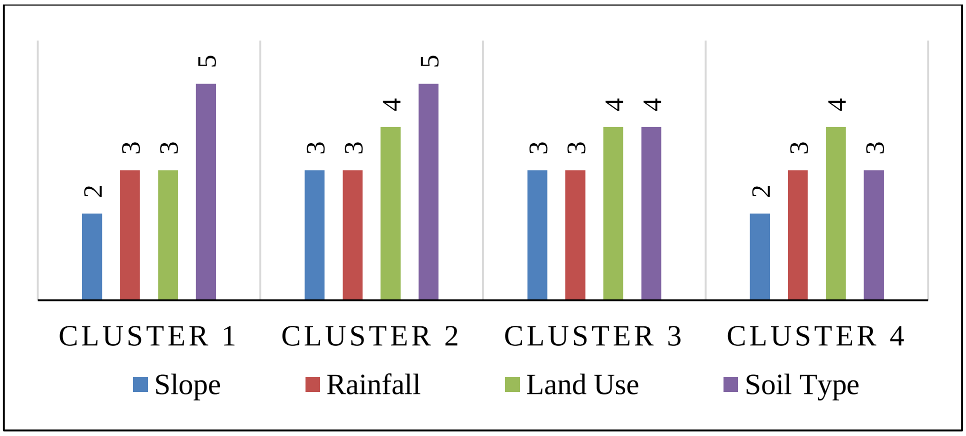


Figure 2. Susceptibility level at highest percentage of input data using FCM

a weight of 4, rainfall with a weight of 3, land use with a weight of 2, and soil type with a weight of 1 [6]. To identify the level of landslide susceptibility of each cluster, these weights are multiplied by the values that have been generated in Figure 2.

The calculation results show that the highest score is generated by cluster 2, with a weight of 34. Thus, cluster 2 becomes the cluster with a highly susceptibility level to landslides. Cluster 3 weighs 33, so it is classified as a susceptibility level to landslides. Cluster 1 and cluster 4 weigh 28. The cluster that is more prone to landslide is determined by the order of the higher factors that influence landslide. In clusters 1 and 4, slope and rainfall have the same values. For land use, cluster 1 weighs 3 while cluster 4 weighs 4, so cluster 4 has a slightly susceptibility level and cluster 1 has a not susceptibility level. The clustering results, together with the level of landslide susceptibility, are presented in a map created using ArcGIS software (Figure 3).

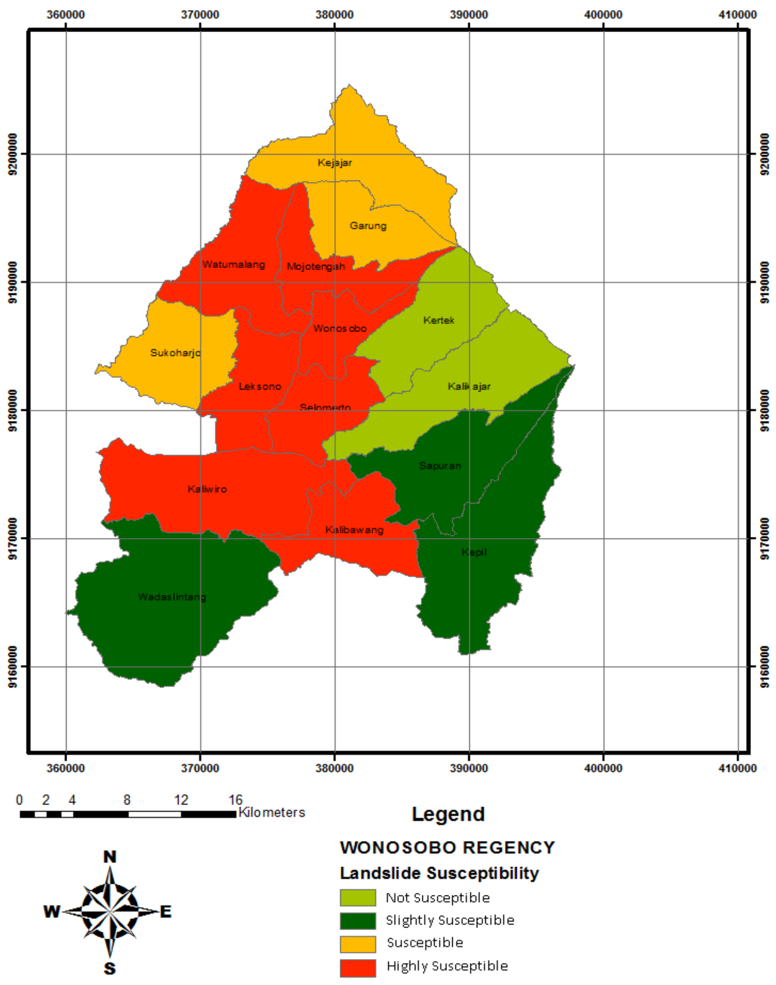


Figure 3. Map of landslide susceptibility of Wonosobo regency using FCM

Figure 3 shows that the cluster with a highly susceptibility level to landslides consists of seven sub-districts, which are Kalibawang, Kaliwiro, Leksono, Watumalang, Mojotengah, Selomerto, and Wonosobo. The cluster with a susceptibility to landslides has 3 members, which are Sukoharjo, Kejajar, and Garung. The cluster with a slightly susceptibility level has 3 members, which are Wadaslintang, Sapuran, and Kepil. The cluster with a not susceptibility level has 2 members, which are Kalikajar and Kertek.

3. 3 Clustering Results using FCMO Method

The clustering results using the FCMO method are also in the form of membership degrees. Table 6 provides the results of the FCMO cluster with 4 as the number of clusters. Two sub-districts have two maximum membership functions in more than one different cluster. In such conditions, cluster selection is carried out by choosing the shortest distance between the input data and each cluster center. If the closest distance has the same value, it prioritizes the match between the input data variables according to the order of the landslide susceptibility level with each cluster center. However, if the distance in the input data variable does not match with the cluster center, then a cluster that has the closest susceptibility level with input data will be selected.
Table 6 represents the distribution of 15 sub-districts in Wonosobo Regency in each cluster. Three sub-districts, which are Wadaslintang, Sapuran, and Garung, are included in cluster 1. Three sub-districts, namely Kalikajar, Kertek, and Wonosobo, are included in cluster 2. Three sub-districts are included in cluster 3, including Sukoharjo, Selomerto, and Kejajar, and the rest are included in cluster 4.

Table 6. Best cluster results using FCMO

No	Sub-districts	Membership degree				Cluster
		Cluster 1	Cluster 2	Cluster 3	Cluster 4	
1	Wadaslintang	**1**	0	0	0	1
2	Kepil	0	0	0	**1**	4
3	Sapuran	**0.25**	0.25	0.25	0.25	1
4	Kalibawang	0	0	0	**1**	4
5	Kaliwiro	0.143	0.214	0.214	**0.429**	4
6	Leksono	0	0	0	**1**	4
7	Sukoharjo	0	0	**1**	0	3
8	Selomerto	0.2	0.2	**0.3**	0.3	3
9	Kalikajar	0	**1**	0	0	2
10	Kertek	0	**1**	0	0	2
11	Wonosobo	0.12	**0.48**	0.16	0.24	2
12	Watumalang	0	0	0	**1**	4
13	Mojotengah	0	0	0	**1**	4
14	Garung	**0.333**	0.1677	0.333	0.167	1
15	Kejajar	0	0	**1**	0	3

As a determinant of landslide susceptibility level in the Wonosobo Regency, this study uses the mode of each variable from all data (in the form of categories) of the sub-districts in one cluster. Then, the mode is adjusted to the categories in Table 1, Table 2, Table 3, and Table 4. Figure 4 illustrates the categorization result of the input data mode. For example, Cluster 1 with a variable slope has a value of 2, meaning that in Cluster 1, most of the sub-districts have a slope of 8-15% with slightly susceptible level of landslide.
The categorization results in Figure 4 are then multiplied by the weight of the landslide factor, namely land slope, rainfall, land use, and soil type, as in the FCM method. The calculation results show that the highest weight is generated by Cluster 4, i.e., 34, so that Cluster 4 has a highly susceptible level to landslides. Cluster 3 weighs 33, so it is a cluster with a susceptible level to landslides. Cluster 1 weighs 29, so it has a slightly susceptible level to landslides. Furthermore, Cluster 2 weighs 28, so it is not susceptible to landslides.

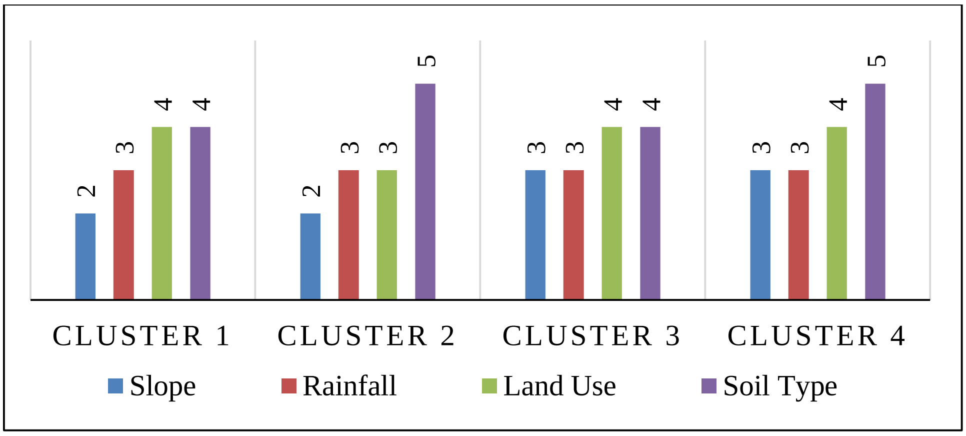


Figure 4. Modes diagram from input data using FCMO

The map of landslide susceptibility map resulted from FCMO is presented in Figure 5. According to this figure, cluster with highly susceptibility to landslide consists of six sub-districts, those are Kalibawang, Kaliwiro, Leksono, Watumalang, Mojotengah, and Kepil. The cluster with a susceptible level to landslide has three sub-districts, which are Sukoharjo, Kejajar, and Selomerto. Cluster with slightly susceptibility to landslide has three sub-districts including Wadaslintang, Sapuran, and Garung. Meanwhile, not susceptible cluster consists of Kalikajar, Kertek, and Wonosobo.

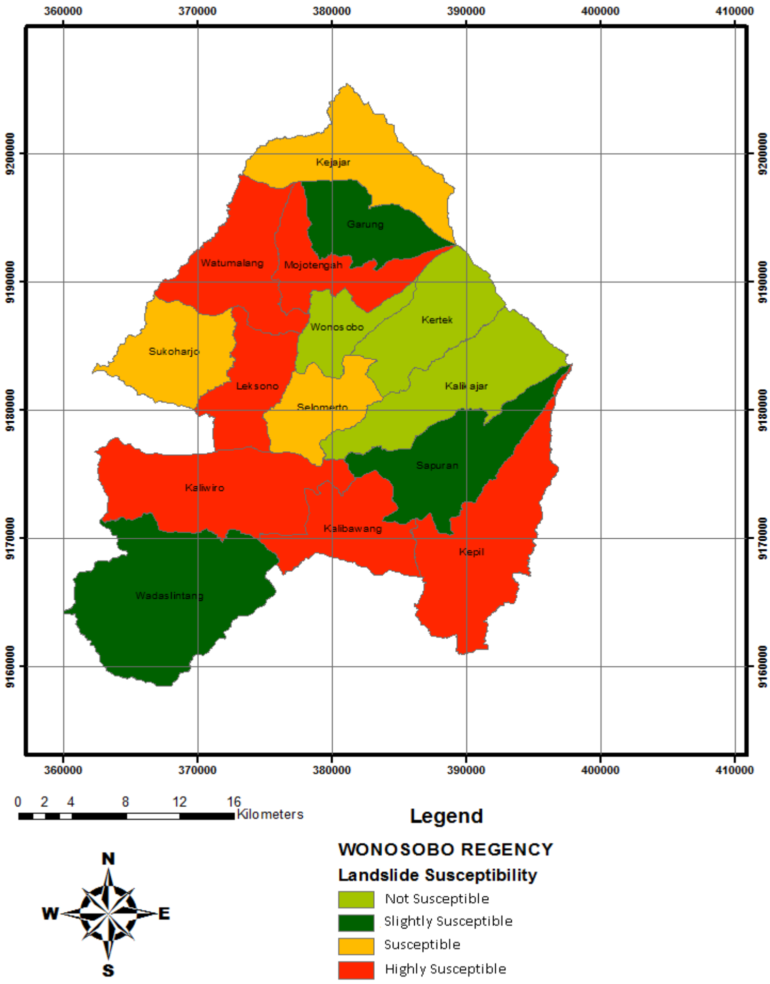


Figure 5. Map of landslide susceptibility in Wonosobo regency using FCMO

3. 4 Mapping Comparison of Best Clustering Results using FCM and FCMO

The validation value of the XBI and MPCI clusters of the FCMO method gives the same results, namely the four optimal number of clusters, as in the FCM method. However, the results of the placement of sub-districts in each cluster are not exactly the same. Table 7 shows several differences in cluster members' results and their characteristics based on the level of susceptibility to landslides. Cluster member differences occur in all clusters. Highly susceptible clusters, Selomerto and Wonosobo Sub-districts, are only created using the FCM method, while Kepil is only produced using the FCMO method; the other sub-districts remain the same in the two clusters. In the susceptible cluster, Garung Sub-District is only generated using the FCM method, while Selomerto is only created using the FCMO method.

Table 7. Mapping comparison of best clustering using FCM and FCMO

Category	*Fuzzy C-Means*	*Fuzzy C-Modes*
Highly susceptible	Members: • Kalibawang • Kaliwiro • Leksono • Watumalang • Mojotengah • Selomerto • Wonosobo	Members: • Kalibawang • Kaliwiro • Leksono • Watumalang • Mojotengah • Kepil
	Characteristics: • Slope 15-25% • Rainfall >2.500 mm/year • Land is used as moor • Regosol soil type	Characteristics: • Slope 15-25% • Rainfal >2.500 mm/year • Land is used as moor • Regosol soil type
Susceptible	Members: • Sukoharjo • Kejajar • Garung	Members: • Sukoharjo • Kejajar • Selomerto
	Characteristics: • Slope 15-25% • Rainfall >2.500 mm/year • Land is used as moor • Andosol soil type	Characteristics: • Slope 15-25% • Rainfall >2.500 mm/year • Land is used as moor • Andosol soil type
Slightly susceptible	Members: • Wadaslintang • Sapuran • Kepil	Members: • Wadaslintang • Sapuran • Garung
	Characteristics: • Slope 8-15% • Rainfall >2.500 mm/year • Land is used as moor • Mediteran soil type	Characteristics: • Slope 8-15% • Rainfall >2.500 mm/year • Land is used as moor • Andosol soil type
Not susceptible	Members: • Kalikajar • Kertek	Members: • Kalikajar • Kertek • Wonosobo
	Characteristics: • Slope 8-15% • Rainfall >2.500 mm/year • Land is used as rice fileds • Regosol soil type	Characteristics: • Slope 8-15% • Rainfall >2.500 mm/year • Land is used as rice fileds • Regosol soil type

In a slightly susceptible cluster, Kepil Sub-District is only produced using the FCM method, while Garung is only generated using the FCMO method. In not susceptible clusters, Wonosobo Sub-District is only generated using the FCMO method. The difference in characteristics is only in the slightly susceptible areas; based on the FCM method's results, this cluster is characterized by a mediterranean soil type, while the cluster of the FCMO method is characterized by an andosol soil type.

4 CONCLUSION

The FCM and FCMO methods have been applied to cluster landslide susceptibility in Wonosobo Regency. The numerical data are analyzed using FCM method. The categorization of numerical data as an input variable is carried out in the clustering process using the FCMO method. Based on the XBI and MPCI values, the two methods give the same optimal number of clusters, which is 4. Although the number is the same, the two methods give different results in the placement of sub-districts in each cluster. Most of the sub-district placement in each cluster of the two methods is the same. There are four sub-districts from 15 sub-districts whose placement in the clustering results is not the same for the two methods. Regarding the value of cluster validity, categorical data processed using FCMO are more suitable for clustering the landslide susceptibility in Wonosobo Regency than numerical data using the FCM method.

ACKNOWLEDGEMENTS

The authors would like to thank the grant from the Ministry of Research, Technology, and Higher Education, No. 080/SP2H/LT/DRPM/2019.

REFERENCES

[1] Regional Disaster Management Authority, 2017. "Recapitulation of Disaster Events in Wonosobo Regency (in Indonesian)," Wonosobo: BPBD.

[2] Regional Disaster Management Authority, 2018, "Recapitulation of Disaster Events in Wonosobo Regency (in Indonesian)," Wonosobo: BPBD.

[3] Basofi, A. Fariza, A. and Nailussa'ada, 2017, "Landslide Susceptibility Mapping using Ensemble Fuzzy Clustering: A Case Study in Ponorogo, East Java, Indonesia," *Proceedings of the 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE*).

[4] Zhang, T., Han, L., Zhang, H., Zhao, Y., Li, X. and Zhao, L., 2019, "GIS-Based Landslide Susceptibility Mapping using Hybrid Integration Approaches of Fractal Dimension with Index of Entropy and Support Vector Machine," *Journal of Mountain Science*, Vol. 16, No. 6, pp. 1275-1288.

[5] Aghdam, I. N., Varzandeh, M. H. M., and Pradhan, B., 2016, "Landslide Susceptibility Mapping using an Ensemble Statistical Index (Wi) and Adaptive

Neuro-Fuzzy Inference System (ANFIS) Model at Alborz Mountains (Iran)," *Environmental Earth Sciences*, Vol 75:553.

[6] Taufik, Q., Firdaus, and Deniyatno, 2012, "Mapping the Threat of Landslide Disasters in the Konawe District (in Indonesian)," *Jurnal Aplikasi Fisika*, Vol 8, pp. 36-40.

[7] Dehnavi, A. , Aghdam, I. N., Pradhan, B., and Varzandeh. M. H. M., 2015, "A New Hybrid Model using Step-Wise Weight Assessment Ratio Analysis (SWARA) Technique and Adaptive Neuro-Fuzzy Inference System (ANFIS) for Regional Landslide Hazard Assessment in Iran," *Catena*, Vol. 135, pp. 122–148.

[8] Conforti, M., Pascale, S., Robustelli, G., and Sdao, F., 2014, "Evaluation of prediction capability of the artificial neural networks for mapping landslide susceptibility in the Turbolo River catchment (northern Calabria, Italy)," *Catena*, Vol. 113, pp. 236–250.

[9] Bahrami, Y., Hassani, B., and Maghsoudi, A., 2020, "Landslide Susceptibility Mapping using AHP and Fuzzy Methods in The Gilan Province, Iran," *GeoJournal*.

[10] Ali, S. A., Parvin, F., Vojteková, J., et al., 2020, "GIS-Based Landslide Susceptibility Modeling: A Comparison between Fuzzy Multi-Criteria and Machine Learning Algorithms."

[11] Goyal M. K. and Sharma, A., 2016, "A Fuzzy C-Means Approach Regionalization for Analysis of Meteorological Drought Homogeneous Regions in Western India, " *Natural Hazards*, Vol. 84, pp. 1831–1847.

[12] Aceves-Fernández, M. A., Sotomayor-Olmedo, A., Gorrostieta-Hurtado, E., Pedraza-Ortega, J. C., Tovar-Arriaga, S., and Ramos-Arreguin, J. M., 2011, "*Performance Assessment of Fuzzy Clustering Models Applied to Urban Airborne Pollution*," *CONIELECOMP 2011, 21st International Conference on Electrical Communications and Computers*, San Andres Cholula, pp. 212-216.

[13] Guha, A. and Veeranjaneyulu, N., 2019, "Prediction of Bankruptcy using Big Data Analytic Based on Fuzzy C-Means Algorithm," *IAES International Journal of Artificial Intelligence (IJ-AI)***,** Vol. 8, No. 2, pp. 168-174.

[14] Horaček, M., Samec, P., and Minar, J., 2018, "The Mapping of Soil Taxonomic Units via Fuzzy Clustering–A Case Study from the Outer Carpathians, Czechia," *Geoderma*, Vol. 326, pp. 111–122.

[15] Rajsekhar, D., Mishra, A.K., and Singh, V. P., 2012, "Regionalization of Drought Characteristics using an Entropy Approach," *Journal of Hydrologic Engineering*, 18(7), pp. 870–887.

[16] Zeraatpisheh, M., Ayoubi, S., Brungard, C. W., and Finke, P., 2019, "*Disaggregating and Updating a Legacy Soil Map using DSMART, Fuzzy C-Means and K-Means Clustering Algorithms in Central Iran*," *Geoderma*, Vol. 340, pp. 249-258.

[17] Wan, S., Yen, J.Y., Lin, C. Y., and Chou, T. Y., 2015, "Construction of Knowledge-Based Spatial Decision Support System for Landslide Mapping

using Fuzzy Clustering and KPSO Analysis," *Arabian Journal of Geosciences*, Vol. 8, pp. 1041–1055.

[18] Correa-Muñoz, N. A., Murillo-Feo, C. A., and Martínez-Martínez, L. J., 2018, "The Potential of PALSAR RTC Elevation Data for Landform Semi-Automatic Detection and Landslide Susceptibility Modeling," *European Journal of Remote Sensing*, Vol. 52, pp. 148-159.

[19] Hassan, A. A., Shah, W., Iskandar, M. F., and Hassan, H. A., 2020, "Evaluate the performance of K-Means and the Fuzzy C-Means Algorithms to Formation Balanced Clusters in Wireless Sensor Networks," *International Journal of Electrical and Computer Engineering (IJECE)* Vol. 10, No. 2, pp. 1515~1523.

[20] Suguna, J. and Arul Selvi, M., 2012, "Ensemble Fuzzy Clustering for Mixed Numeric and Categorical Data," *International Journal of Computer Applications*, Vol. 42, No.3, pp. 19-23.

[21] Mingoti, S. A. and Lima, J. O., 2006, "Comparing SOM Neural Network with Fuzzy C-Means, K-Means and Traditional Hierarchical Clustering Algorithm," *European Journal of Operational Research*, Vol. 174, pp. 1742–1759.

[22] Huang, Z. J. and Ng, M. K., 1999, "A Fuzzy K-Modes Algorithm for Clustering Categorical Data," *IEEE Transactions on Fuzzy* Systems, Vol. 7, pp. 446-452.

[23] Soliman, O. S., Saleh, D. A., and Rashwan, S., 2012, "A Bio Inspired Fuzzy K-Modes Clustering Algorithm," *ICONIP*, Vol. 7665, pp. 663–669.

[24] Khalid, 2018, " Hybridization Methods of Artificial Bee Colony and Fuzzy K-Modes for Categorical Data Clustering," *SYSTEMIC: Information System and Informatics Journal*, Vol. 4, pp. 36-42.

[25] Kuo, R. J., Potti, Y., and Zulvia, F. E., 2018, "Application of Metaheuristic based Fuzzy K-modes Algorithm to Supplier Clustering," *Computers & Industrial Engineering*, Vol. 120, pp. 298-307.

[26] Central Bureau of Statistics, 2019, "*Wonosobo Regency in Number* (in Indonesian), " Wonosobo: BPS.

[27] Bezdek, J. C., 1981, "*Pattern Recognition with Fuzzy Objective Function Algorithms*," New York: Plenum Press.

[28] Stylianou, C. and Andreou, A. S., 2007, "A Hybrid Software Component Clustering and Retrieval Scheme, using an Entropy-based Fuzzy *k*-Modes Algorithm," *19th IEEE International Conference on Tools with Artificial Intelligence*, Vol. 1, pp. 202–209.

[29] Xie, X. L. and Beni, G., 1991, "A Validity for Fuzzy Clustering," *Transactions on Pattern Analysis and Machine Intelligence*," Vol. 13, pp. 841-847.

[30] Wang, W. and Zhang, Y., 2007, "*On Fuzzy Cluster Validity Indices*," *Fuzzy Sets System*, Vol. 158, No. 19, pp. 2095-2117.