Analysis of Rice Field Cluster in Indonesia as an Evaluation of Food Production Availability Using Fuzzy C-Means

H Setiono[1], T M Dianto[2]

[1]Statistic of Mojokerto Regency
[2]Statistic of Sumenep Regency

*Corresponding author's e-mail: herusetiono@stit-au.ac.id

Abstract. Rice fields area in Indonesia is getting narrower every year with the rampant construction of housing and buildings. It results in lower availability of food production hence to meet the needs we have to import rice from other countries. By clustering rice fields, it can be used as an evaluation material to increase food production in Indonesia so that the need for rice imports can be minimized. The method used in the grouping of Rice Fields is the Fuzzy C-means method, implementation of the Knime Tool with data training and testing. The Fuzzy C-Means program produces three data groups/clusters, namely wide, moderate, and narrow rice fields. The results of the clustering show that the most potential areas for food production from rice fields are East Java, Central Java, and West Java.

1. Introduction

In the past, during the leadership of the late President Soeharto during the New Order era, Indonesia was known as a country with rice self-sufficiency in 1984 [1], Rice production reached around 27 million tons, while domestic rice consumption was below 25 million tons, meaning that Indonesia had a surplus of 2 million tons. The Rice Self-Sufficiency Program received an award from FAO in 1985 in rice export activities[2]. It was caused by the large rice fields area for rice production in Indonesia. However, every year the amount of rice production decreases. One of the reasons is the decrease in the development of rice fields every year. Data from the Central Statistics Agency (2001), Indonesia's rice fields in 1993 was 8,500,000 Ha, then in 2000 (7 years later) it greatly shrunk to only 7,790,000 Ha or 710,000 Ha shrinkage or every year rice fields in Indonesia was shrinking 59,167 ha[3]. The decrease was caused by the transfer of function from rice fields to residential and industrial development. Hence the decrease has an impact on the reduced availability of food production. To overcome the availability, the government is intensifying rice imports to meet domestic food needs.

Rice fields clustering in each province can be used as an evaluation material in increasing the availability of food production in an area by finding out the rice fields area in Indonesia. With technology, clustering can minimize rice import activities in meeting domestic food needs.

Fuzzy C-Means is a method widely used to solve problems related to clusters/classification. Among them are used to cluster rice varieties[4], diagnose heart disease[5], analysis of data confidence based on classification methods[6]. Hence to solve the problem of rice fields clusters in this study, the Fuzzy C-Means method can be used. The Fuzzy C-Means method is a data clustering technique in which the existence of each data point in a cluster is determined by the degree of membership[7] so that the cluster of rice fields can be determined

2. Literature Review

The fuzzy set theory will provide answers to a problem that contains uncertainty[8]. One of the sections in fuzzy is fuzzy clustering, this fuzzy is used to cluster data. This clustering is divided into four methods, namely Fuzzy Subtractive Clustering, Mountain, K-means and C-Means methods. Fuzzy C-Means (FCM) is a data clustering technique in which the existence of each data in a cluster is determined by the membership value[9]. In the initial conditions, the center of the cluster is still not accurate, so it is necessary to improve the center of the cluster repeatedly until it is at the right point. Each data will have a degree of membership for each cluster. The algorithm of Fuzzy C-Means is as follows[10]

2.1 Input data to the cluster x, in the form of a matrix of a size of n x m (n = number of data samples, m = attributes of each data). Xij=data number i (i=1,2,..,n) , attribute number j(j=1,2,..,m)

2.2 Determine:

- Number of clusters (c)
- Rank (w)
- Maximum iterations
- Smallest error expected
- Initial objective function (P0=0)
- Initial iteration (t=1)

2.3 Generate a random number as an element of the initial partition matrix U.
Count the number of each column:

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

Then

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

2.4. Calculate the center of the cluster number k:V_{kj} with k=1,2,…,c and j=1,2,..,m

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

2.5. Calculate the objective function at the iteration number t, P_t :

$$P_t = \sum_{i=1}^{n}\sum_{k=1}^{c}\left(\left[\sum_{j=1}^{m}(x_{ij} - V_{kj})\right](\mu_{ik})^w\right) \quad (4)$$

2.6. Calculate the change in the partition matrix

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right]^{\frac{-1}{w-1}}}{\sum_{k=1}^{c}\left[\sum_{j=1}^{m}(X_{ij} - V_{kj})^2\right]^{\frac{-1}{w-1}}} \quad (5)$$

2.7. Check stop conditions:

- if t>max iteration then stop
- Otherwise, t=t+1 repeat step 4

3. Research Method

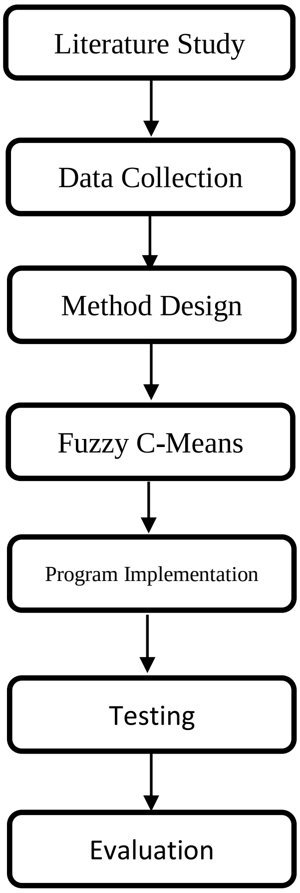
3.1. Literature Study

The object of research are rice fields from all provinces in Indonesia. Data on rice field area was obtained from the Ministry of National Land Agency (2015-2019) and Rice Production from the Central Statistics Agency.

Table 1. Sample data figure Land Area and Rice Production.

Province	Rice Fields					Rice Production	
	2015	2016	2017	2018	2019	2018	2019
1 Aceh	290.337	293.067	294.483	193.308	213.997	1.861.567	1.714.437
2 North Sumatera	423.465	423.029	415.675	245.801	308.667	2.108.284	2.078.901
3 West Sumatera	226.377	222.482	222.021	197.800	194.282	1.483.076	1.482.996
4 Riau	72.005	72.151	70.016	86.247	62.689	266.375	230.873
5 Jambi	94.735	96.588	97.690	111.147	68.349	383.045	309.932
6 South Sumatera	620.632	615.184	621.903	387.237	470.602	2.994.191	2.603.396
7 Bengkulu	85.130	83.449	82.429	47.968	50.840	288.810	296.472
8 Lampung	377.463	390.799	396.599	253.583	361.699	2.488.641	2.164.089
9 Bangka Belitung	10.668	13.820	13.679	5.409	22.402	45.724	48.805
10 Riau Islands	246	286	310	1.220	1.394	1097	1150
11 Greater Jakarta	650	581	571	451	414	4.899	3.359
12 West Java	912.794	913.976	911.817	930.334	928.218	9.647.358	9.084.957
13 Central Java	965.261	963.665	951.752	980.618	1.049.661	10.499.588	9.655.653
14 Special Region of Yogyakarta	53.553	53.985	51.343	75.990	76.273	514.935	533.477
15 East Java	1.091.752	1.087.018	1.081.873	1.287.356	1.214.909	10.203.213	9.580.933
16 Banten	199.492	203.123	199.408	198.284	204.335	1.687.783	1.470.503
17 Bali	75.922	76.096	74.278	69.078	70.996	667.069	579.320
18 West Nusa Tenggara	264.666	276.230	276.306	227.786	234.542	1.460.338	1.402.182
19 East Nusa Tenggara	177.238	181.039	184.346	146.071	155.520	899.935	811.724
20 West Kalimantan	330.724	356.741	368.728	155.818	242.972	799.715	847.875
21 Central Kalimantan	196.813	194.782	180.034	187.008	136.916	514.769	443.561
22 South Kalimantan	450.152	454.121	447.918	252.972	290.716	1.327.492	1.342.861
23 East Kalimantan	57.087	56.505	59.425	36.399	41.406	262.773	253.818
24 North Kalimantan	21.448	20.520	18.189	14.265	11.922	45.063	33.357
25 North Sulawesi	55.825	60.562	59.886	52.236	47.043	326.929	277.776
26 Central Sulawesi	129.014	132.489	136.541	119.670	116.828	926.978	844.904
27 South Sulawesi	628.148	649.190	646.611	641.457	654.818	5.952.616	5.054.166
28 Southeast Sulawesi	103.812	109.854	108.466	79.910	82.117	538.876	519.706
29 Gorontalo	32.054	32.749	32.681	29.067	33.056	269.540	231.211
30 West Sulawesi	61.292	63.671	64.232	42.216	39.485	316.478	300.142
31 Maluku	13.394	14.354	16.732	13.660	18.283	116.228	98.254
32 North Maluku	11.801	13.068	13.221	9.041	13.542	49.047	37.945
33 West Papua	10.193	10.680	11.340	4.239	8.860	24.967	29.943
34 Papua	48.764	51.880	53.543	21.498	36.195	223.119	235.339
	8.092.907	8.187.734	8.164.046	7.105.144	7.463.948	59.200.518	54.604.017

3.2. Data processing
From the rice fields data for each province, 34 data will be clustered into 3 levels of areas, namely high, medium, and low. The data processing process is divided into several stages. Here are the flow stages of the data processing as shown in Figure 2 below



This clustering is carried out using the Fuzzy C-Means algorithm and the Knime application with the Fuzzy c-means function.

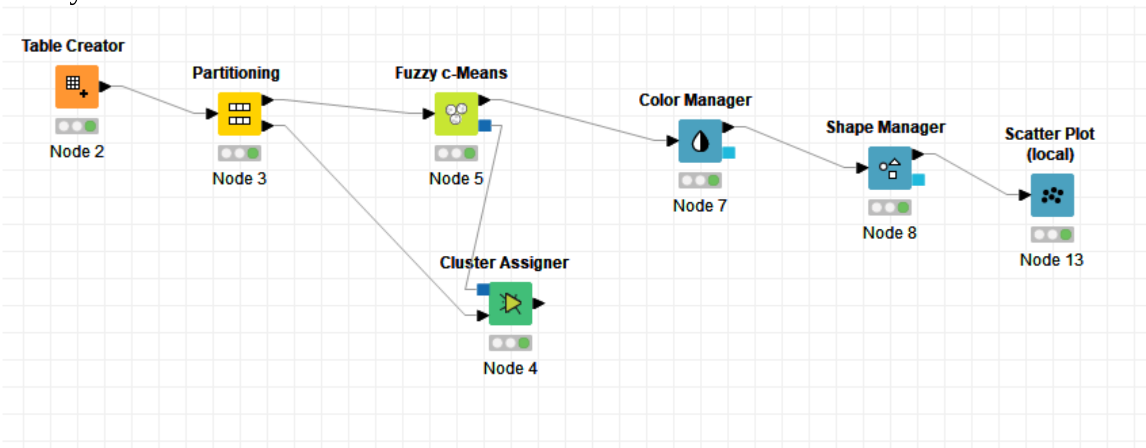


Figure 3. Implementation Model of the Fuzzy C-Means Algorithm on Knime.

For the calculation of clusters using the determination of the 3 number of clusters, namely narrow, moderate, and wide with a maximum iteration of 100 with 95% data participation.

The data entered in the creator table is connected to the partitioning data of 95% of the training data entered in an analytical Fuzzy C-Means, and the results of the cluster are displayed in the color management, shape management is shown as a scatter plot.

4. Results and Discussion

The program implementation of this research is divided into two, namely the clustering process and the testing process. In the clustering process, the initial data is used and three clusters are generated with degrees of membership. The black color indicates cluster 1 which is narrow, the green color indicates cluster 2 which is moderate, and red indicates cluster 3 which is wide. The clustering process produces output data. The results of the cluster can be seen in Figure 4 with different colors for each cluster. The cluster at the top is the center of the largest cluster, a large area of rice fields. Likewise, the cluster below that is an area with a narrow rice fields area. From the clustering process, the data for each cluster can be seen in Table 1.

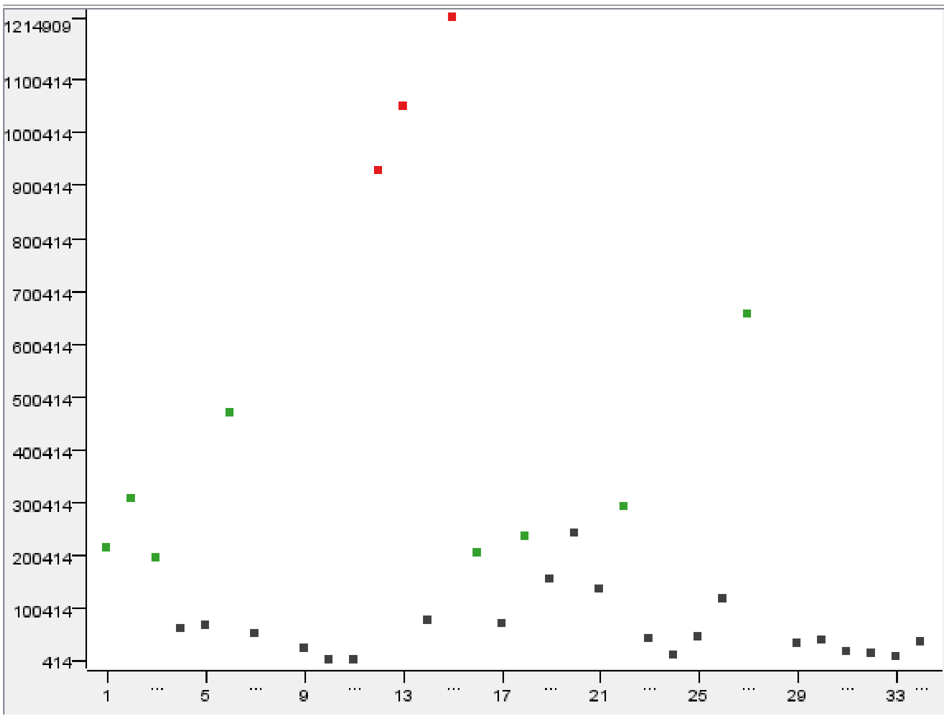


Figure 4. Clustering results.

Table 2. Clustering result data.

Cluster	Data Number	Province
1	4	Riau
	5	Jambi
	7	Bengkulu
	9	Bangka Belitung
	10	Riau Islands
	11	Greater Jakarta
	14	Special Region of Yogyakarta
	17	17 Bali
	19	East Nusa Tenggara
	20	West Kalimantan

Cluster	Data Number	Province
	21	Central Kalimantan
	23	East Kalimantan
	24	North Kalimantan
	25	North Sulawesi
	26	Central Sulawesi
	28	Southeast Sulawesi
	29	Gorontalo
	30	West Sulawesi
	31	Maluku
	32	North Maluku
	33	West Papua
	34	Papua
2	1	Aceh
	2	Sumatera Utara
	3	Sumatera Barat
	6	Sumatera Selatan
	8	Lampung
	16	Banten
	18	Nusa Tenggara Barat
	22	Kalimantan Selatan
	27	Sulawesi Selatan
3	12	West Java
	13	Central Java
	15	East Java

Table 3. Clustering Results.

	Cluster		
	1	2	3
Area	55610	285541	1053786
Productivity	293475	1858708	9328262

5. Conclusion

The Fuzzy C-Means (FCM) method can be implemented in the clustering of rice fields in Indonesia. 3 clusters resulted from the clustering process, namely narrow, moderate, and wide rice fields. This clustering uses 34 data so this can be used as material for evaluation for the government in increasing the availability of food production to minimize rice imports, especially maximizing large rice fields, namely the Java area, specifically West Java, Central Java and East Java.

Based on land area, in moderate land areas, the government and the community must work together to suppress factors that narrowing of land such as the construction of housing, malls and high-rise buildings and the government and the community should cooperate in adding and using new methods to improve agricultural production.

Accelerate the implementation of agrarian reform. This acceleration can be done by, among others, updating the Agrarian Law. This law must place emphasis on saving agricultural land and prohibiting the conversion of productive agricultural land by strengthening the role of Village Institutions, increasing the area of rice fields of 22 million ha or 2.2 million ha per year for the next 10 years, and strengthening the role of the National Land Agency to control national land.

References

[1] Syahrir Ika, Hadi Setiawan, Sofia Arie Damayanty,March 2015 Vol 19 :Evaluation of Indonesian Food Politics and Fiscal Politics Support, Kajian Ekonomi Keuangan

[2] Syahrir Ika, Hadi Setiawan, Sofia Arie Damayanty,March 2015 Vol 19 :Evaluation of Indonesian Food Politics and Fiscal Politics Support, Kajian Ekonomi Keuangan

[3] R Uchyani, SW Ani, Jurnal SEPA. Februari 2012 Vol 8 (2): 54 61 Tren Alih Fungsi Lahan Pertanian Di Kabupaten Klaten

[4] AFO Gafar, I Tahyudin. 2017. Comparison between k-means and fuzzy C-means clustering in network traffic activities. International Conference on Management Science and Engineering Management

[5] Hammouda, Khaled. A Comparative Studi Of Data Clustering Techniques. Journal Universitas Of Waterloo. Ontario. Canada

[6] Asyali, Musa. Alci, Musa. Reliability Analysis Of Microarray Data Using Fuzzy C-MeansAnd Normal Mixture Modeling Based Classification Methods. Jurnal Bioinformatics Vol.21 No.5 2005, Pages 644-649

[7] Kusumadewi, Sri. Purnomo, Hadi. 2010. Aplikasi Logika Fuzzy Untuk Pendukung Keputusan Eds.2.Yogyakarta. Graha Ilmu

[8] RD Ariyanti, S Kusumadewi, IV Paputungan. Beck depression inventory test assessment using fuzzy inference system 2010 International Conference on Intelligent Systems, Modelling

[9] S Kusumadewi, S Hartati 2006. Neuro fuzzy: integrasi sistem fuzzy dan jaringan syaraf. Yogyakarta. Graha ilmu

[10] S Kusumadewi, S Hartati, A Harjoko, R Wardoyo 2006. Fuzzy multi-attribute decision making (fuzzy madm). Yogyakarta: Graha Ilmu, 78-79