

# Quantitative approach for reclassification of the spatial cluster of archipelagos in Maluku Province for the basis of forest development

Patrich Papilaya, Endang Suhendang, I Nengah Surati Jaya, Teddy Rusolono

Department of Forest Management, Faculty of Forestry, IPB University, Indonesia

## Article Info

### Article history:

Received Jun 21, 2020

Revised Jul 17, 2021

Accepted Jul 30, 2021

### Keywords:

Accuracy clustering  
Archeipelago clustering  
Discriminant analysis  
Maluku

## ABSTRACT

In natural resource management, it is necessary to group regions based on the similarity of their spatial and non-spatial characteristics, to efficiency and effectiveness. Therefore, this study describes the re-grouping of the twelve island clusters established by the provincial government of Maluku into more homogeneous classes. The re-grouping was carried out based on the biophysical conditions of the regions, therefore, it could be used as the basis for determining the forest management units. The results showed that the twelve designated island clusters could be simplified to eight more homogeneous island clusters with 86.4% accuracy and 82.2% validation. It also showed that there were thirteen significant changes in the grouping of clusters of the island, including the horticultural crop area (Bf) and horticultural crop production (E). Moreover, when the island cluster is reclassified into 5 classes, the grouping would be more accurate, with 94.9% accuracy and 92.4% validation. This study concludes that there are two dominant factors in the classification of the island cluster in Maluku province namely, biophysical and social.

*This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.*



## Corresponding Author:

I Nengah Surati Jaya  
Department of Forest Management, Faculty of Forestry  
IPB University  
Ulin road, IPB University, Dramaga, Bogor 16680  
Email: ins-jaya@apps.ipb.ac.id

## 1. INTRODUCTION

The establishment of the island cluster is a strategic step taken by the Local Government of Maluku Province to accelerate the process of equalization and improving the welfare of the community. This island cluster consists of a collection of large and small islands with unique biophysical, economic, and socio-cultural characters, including rich natural resource potential, spread over an area of  $\pm 712\,480\text{ km}^2$  [1]-[3]. According to the Decree of the Minister of Affairs and Fisheries Number Kep.34/Men/2002, a cluster of islands is a group of islands that are geographically close to each other, without a close connection. However, they have interacting ecosystems, including socio-economic and cultural conditions, both individually and in groups. The twelve clusters formed are part of the area of 10 districts and 2 cities in Maluku Province, and they are expected to promote equalization of development in the province. Meanwhile, this province is dominated by small islands that are geographically separated by vast oceans and have a unique diversity of natural resources potential [4]-[9].

The National Human Development Index (HDI) describes the level of well-being of the Indonesian people [10], [11]. According to the index, Maluku Province ranks 26th out of 34 provinces in Indonesia. This

is because the province has the potential of natural wealth that could encourage the improvement of the welfare of people that inhabit 1,412 islands. Furthermore, according to the human development index (HDI) and poverty depth index (PDI), the lowest indexes were in 5 island clusters namely Island Clusters 4, 9, 10, 11, and 12, followed by 4 and 2. Meanwhile, the highest indexes were in Cluster 7 with an HDI value of 80.24 and PDI 0.93 [12]. The establishment of the island cluster is expected to be one of the solutions to improving the welfare of the community, where socioeconomic and cultural conditions could be clearly mapped. This makes it easier for local governments to implement development strategies more effectively and efficiently.

From the point of view of forest management, the current conditions in the twelve clusters of islands are not appropriate for the establishment of good forestry management units in supporting the welfare of the community. This is due to the uneven socio-economic conditions of the people in the islands, as described above. One of the best alternatives to this problem is to group the island clusters into more homogeneous areas. This could be carried out through the reclassification process [13]. The creation of homogeneous areas would facilitate development decision-making, specifically Forestry Development in Maluku Province. Discrimination analysis was used to obtain more homogeneous management units through the reclassification process, using biological, physical, economic, and social data from twelve island clusters in Maluku Province. Moreover, this approach has never been used in development planning in Indonesia, especially in other island provinces.

Discriminant analysis is a multivariate technique that allows the differentiation of objects with different populations and allocates new objects into previously defined populations [14]. The analysis has two very nice features namely 1) parsimony of description and 2) clarity of interpretation [15]. Verbel *et al.* [16] studied discriminant analysis has been widely used in answering various problems related to multivariate data. Le *et al.* [17] it was used in the data evaluation process to generate positron emission tomography (PET) brain image as a classification material in Alzheimer's disease patients. Furthermore, [18] the analysis was also used in comparing the partial least squares (PLS) discrete analysis and sparse PLS discrimination analysis in the detection and mapping of *Solanum mauritanum* in commercial plant forests using image textures. The results showed that least squares discriminant analysis (SPLS-DA) successfully performed simultaneous variable selection and dimension reduction to produce an overall classification accuracy of 77%. In contrast, the partial least squares-discriminant analysis (PLS-DA) model along with variable interest in projection (VIP) resulted in an overall classification accuracy of 67%. Onumanyi *et al.* [19] proposed the principle of discriminant analysis was applied in addressing the sorted statistic scheme. Additionally, the results obtained through the Monte Carlo simulation showed that the DA-OS scheme achieved a small CFAR loss of approximately 0.392 dB, relative to the average cell scheme (CA) in homogeneous radar return conditions with a possible detection of 0.5. These results outperform previous results and are in line with [20], [21], which used technological approaches in securing wildlife from human activity.

The main purpose of this research is to obtain or build homogeneous regions/clusters from the biophysical and socioeconomic aspects of island clusters, including their spatial patterns. Furthermore, the procedure used which involves quantitatively reclassifying the cluster of islands in Maluku Province as the basis for forestry development is described in the flow chart as shown in Figure 1.

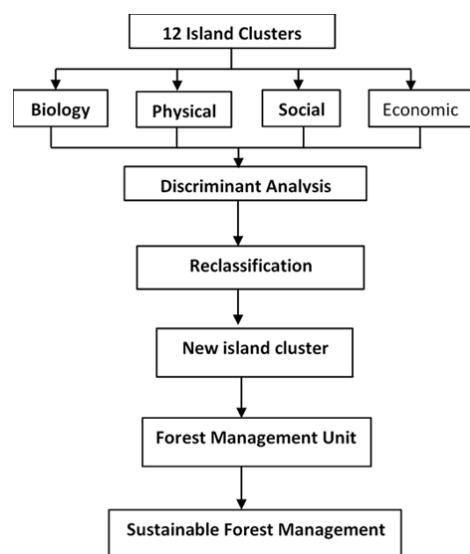


Figure 1. Flowchart of research framework

2. RESEARCH METHODS

2.1. Date and study site

This research was conducted from January to April 2019 on twelve island groups in Maluku Province. Biophysical, social and economic data from 118 sub-districts located in twelve island clusters in Maluku Province were used as the main source. Shapefile data for district, city and sub-district boundaries in Maluku Province, 2018 forest area shapefile data in Maluku Province obtained from the Ministry of Environment and Forestry (KLH). The study covers twelve island clusters in Maluku Province (Table 1). The research location can be seen in Figure 2.

Table 1. Twelve island clusters in Maluku Province

No.	Clusters	Number of districts
1	Buru Island	16
2	West Seram	11
3	North Seram	15
4	East Seram	4
5	South Seram	6
6	Banda dan Teon Nila Serua	1
7	Ambon dan Lease	12
8	Kei	16
9	Aru	10
10	Tanimbar	10
11	Babar	7
12	Terselatan	10

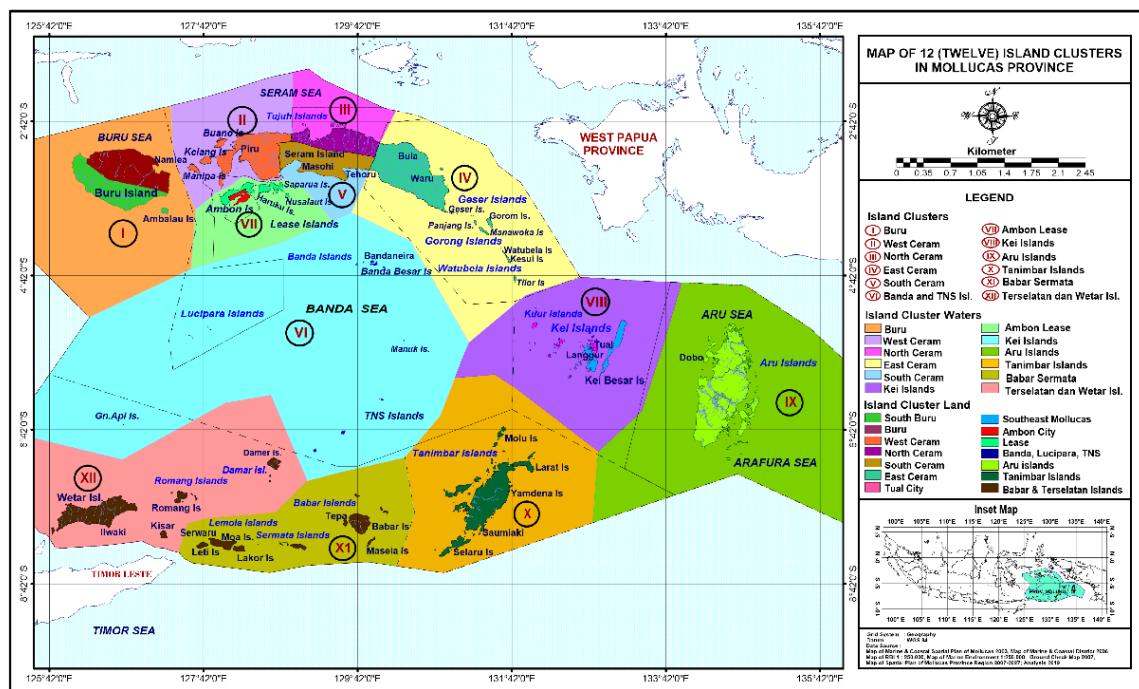


Figure 2. Twelve clusters of islands in Maluku Province

2.2. Software and hardware and data

Spatial analysis was carried out using the ArcGIS® version 10.6 software to extract forest cover data at the district to the sub-district levels. Furthermore, the discriminant analysis was carried out using statistical package for the social sciences (SPSS) version 25 [22], [23]. The main data in this study is the biophysical, social, and economic data from 118 sub-districts on twelve clusters of islands acquired from Maluku Province in the year 2018. Data Shapefile limit of the towns and districts in this province, and Landsat image of data Shapefile Forest area in year 2016 in the province obtained from the Ministry of Environment and Forestry (KLHK) were also used.

Supporting data was obtained from several sources including the results of discussions with the Maluku Provincial Regional Leaders, report on the Study of the Development of Island Clusters Based on Maluku [24], coordinates of geographic border districts, and districts in the province.

### 2.3. Discriminant analysis

Discriminatory analysis is "the dependent technique in which the independent variable is non-metric" [25]. Furthermore, according to [13], [14] this analysis is "grouping each object into two or more based on the criteria of independent variables. Determination of the twelve island clusters was carried through a typology approach, using socio-cultural, economic, and biophysical parameters. The data source was the statistical data of the Subdistrict in Maluku Province in 2018, and data on forest cover in the province in 2016, from the Ministry of Environment and Forestry.

The typology study of the island clusters consists of twelve groups (clusters) which were dependent variables (response variables) with fourteen independent variables sourced from subdistrict in figures 2018 as follows (Table 2). The result of the calculation of discriminant function generated using fourteen variables was tested for accuracy by calculating the values of overall accuracy, user accuracy, producer accuracy, and cross-validation.

Table 2. Twelve island clusters independent variables

Variable	Information	Unit	Source
	Biophysics		
X <sub>1</sub>	: Forest Area Cover	ha	KLHK
X <sub>2</sub>	: Cluster Area	ha	BPS
X <sub>3</sub>	: Food Crop Area,	Ha	BPS
X <sub>4</sub>	: Horticultural Plant Area	Ha	BPS
X <sub>5</sub>	: Plantation Plant Area	Ha	BPS
	Social		
X <sub>6</sub>	: The Population of People	People	BPS
X <sub>7</sub>	: Population Sex Ratio	Percent	BPS
X <sub>8</sub>	: School Age Population	People	BPS
X <sub>9</sub>	: Number of Capture Fisheries Households,	Family	BPS
	Economy		
X <sub>10</sub>	: Food Crop Production	Ton/ha	BPS
X <sub>11</sub>	: Horticultural Crop Production,	Ton/ha	BPS
X <sub>12</sub>	: Plantation Crop Production	Ton/ha	BPS
X <sub>13</sub>	: Capture Fisheries Production	ton	BPS
X <sub>14</sub>	: Animal Husbandry Potential	Numbers	BPS

[BPS] Central Bureau of Statistics [11]

## 3. RESULTS AND DISCUSSION

### 3.1. The homogeneity of island clusters

The fundamental requirement of the disinterest analysis test is that the free variable variance of each cluster of islands needs to be the same, in which case, fourteen independent variables. Moreover, variances between the fourteen independent variables should be equal. The results of the twelve homogeneity tests of the island clusters indicate that the significant value of the test's M's (SIG) was  $0.0001 < \alpha < 0.05$ . Therefore, it could be concluded that the matrix of variance-covariant was not homogeneous, meaning that the assumption of discriminatory analysis was not fulfilled. According to [13] the assumption of the matrix of variance-covariant in practice is often violated. While [26] stated that discriminant analysis is not particularly sensitive to the violation of the assumption of the matrix of variance-covariance. According to [27] the analysis of discriminant function remains robust although, the assumption of homogeneity of variance is not met with the required data. This means that data from twelve island clusters are eligible for use in discriminant analysis.

### 3.2. The similarity between the twelve variables in the island cluster

The test of the average similarity of the island cluster group or the test of equality of groups means is used to ascertain whether it is univariate and if there is a discrepancy in setting the twelve clusters when viewed from the fourteen variables, with respect to the biophysical and socio-economic parameters used as a free variable (independent variable). This assessment was carried out in two ways such as by investigating the value of Wilks' Lambda and a significant value on the F test. When the value of Wilks' Lambda approaches the value 0, it indicates an increasingly significant condition. Meanwhile, when the value approaches 1, it means it is not significant. The results of the average group similarity test are presented in Tables 3 and 4. Table 3 shows fourteen independent variables of the twelve island clusters that were tested and shown to have a value close to 0, which means that they were all significant at (sig.)  $< 0.05$ . This shows that the data was univariate,

meaning there are different groupings of 12 island clusters formed using fourteen independent variables. The independent variables that contributed the most from the discriminant function are shown in the next analysis (Table 4).

Table 3. Similarities in the variables of the twelve island clusters

No.	Variable	Wilks' Lambda	F	Sig.
1	Forest Cover (Bf)	0.693	4.279	0.000
2	Cluster (Bf)	0.318	20.667	0.000
3	Food Crop Area (Bf)	0.562	7.513	0.000
4	Horticulture Plant Area (Bf)	0.114	74.987	0.000
5	Plantation Area (Bf)	0.498	9.702	0.000
6	Total Population (S)	0.648	5.244	0.000
7	Sex Ratio (S)	0.793	2.510	0.000
8	Total School Age Population (S)	0.677	4.608	0.000
9	Number of Capture fisheries Households (S)	0.422	4.279	0.000
10	Food Crop Production (E)	0.603	20.667	0.000
11	Horticultural Crop Production (E)	0.635	7.513	0.000
12	Plantation Plant Production (E)	0.519	74.987	0.000
13	Capture Fisheries Production (E)	0.667	9.702	0.000
14	Animal Husbandry Potential (E)	0.504	5.244	0.000

Information: Bf = Biophysical S = Social E = Economy

Table 4. Variable selection process according to Wilks' Lambda

Variable	Wilks' Lambda	df1	df2	Sig.
Forest Cover (Bf)	0.235	33	307.107	0.000
Cluster (Bf)	0.001	88	658.646	0.000
Food Crop Area (Bf)	0.128	44	396.007	0.000
Horticulture Plant Area (Bf)	0.019	55	475.722	0.000
Plantation Area (Bf)	0.004	77	606.685	0.000
Total School Age Population (S)	0.677	11	106.000	0.000
Number of Capture fisheries Households (S)	0.008	66	545.892	0.000
Plantation Plant Production (E)	0.348	22	210.000	0.000
Capture Fisheries Production (E)	0.001	99	702.540	0.000
Horticultural Crop Production (E)	0.000	110	739.231	0.000
Animal Husbandry Potential (E)	0.000	121	769.595	0.000

### 3.3. Variable forming discriminant function of the twelve clusters of islands

The discriminant function was formed from the dependent and independent variables that were examined. Furthermore, analysis of the variables forming the discriminant function was used to ascertain the independent variables that form the discriminant function of the twelve island clusters. The results of the analysis are shown in Table 4. Significant tests between two or more variables were carried out in stages using the Mahalanobis Distance method. This method would produce data distribution based on the average distance between the mean value of the examined data group [26], [27].

Wilk's Lamda value showed that only twelve variables out of the fourteen were included in the discriminant function, as the remaining two, namely total population (S) and food crop production (E) variables were not included. Furthermore, the twelve variables were obtained from the discriminant analysis using a stepwise process. This process started with the variable with the smallest statistical number, namely the number of school-age population, followed by the other variables.

Wilk's Lamda is in principle a total variance in discriminant scores that cannot be explained by the differences between the island cluster groups tested [22]. The initial stage by entering the variable, total school age population (S) with the figure explaining Wilk's Lamda is 0.677. This means 67.7% variance cannot be explained by differences between the groups tested. The second stage was by adding the Variable Plantation Plant Production with Wilk's Lamda value of 0.348, meaning 34.8% variance cannot be explained by differences between the groups tested. Subsequent results showed that Wilk's Lamda variable value was smaller, meaning that the smaller variance could not be explained by the differences between the test groups. The results of the analysis of the formation of discretionary variables were adequate, as could be seen from the rapid decline in the value of Wilk's Lamda from the initial stage. This indicates that most discriminant scores could be explained on the compost or cluster tested.

### 3.4. Distance difference between island cluster group

The distance between the island cluster groups analysed was tested to ascertain the furthest distance of a cluster from another cluster. This distance determines the differences between the cluster groups in the analysis. Paiwise group comparisons facility was used to examine the furthest distance between groups of twelve island clusters. The results of the analysis are shown in Table 5. Furthermore, the result of the test on the average distance between groups gives quite uniform results. Meanwhile, the farthest value was obtained from cluster 3 with 8 pairs of 133 (rounding value), followed by a 3-to-4 cluster pair of 129, pair 3 by 12, and so on. The results of this analysis also showed that the closest distance from couples to all groups existed in several groups of 1, 4, 8, 9, 11, and 12. Therefore, it could be concluded that there was a similarity of data-forming discriminant functions in the five groups (clusters).

Table 5. Pairwise group comparisons

Clusters	1	2	3	4	5	6	7	8	9	10	11	12
1		16	115	5	19	10	22	7	5	19	4	5
2	16		66	17	23	10	25	17	14	26	12	16
3	115	66		117	90	29	112	121	106	110	97	113
4	5	17	117		17	10	20	2	3	13	1	2
5	19	23	90	17		5	27	16	19	27	15	19
6	10	10	29	10	5		14	11	12	14	11	12
7	22	25	112	20	27	14		15	16	21	10	13
8	7	17	121	2	16	11	15		3	16	1	1
9	5	14	106	3	19	12	16	3		14	2	2
10	19	26	110	13	27	14	21	16	14		9	12
11	4	12	97	1	15	11	10	1	2	9		1
12	5	16	113	2	19	12	13	1	2	12	1	

The grouping of island clusters described in Table 5 is based on the average value (centroid) of the discriminant functions shown in Figure 3. Furthermore, centroids were used to ascertain how the dissemination of each data was carried out and how close the centroids of each group were formed. The formed distribution pattern consists of cluster groupings. The first group consists of a combined cluster of 1, 4, and 8, while the second consists of island clusters 9, 11, and 12. Moreover, the third group consists of Clusters 9, 11, and 12, followed by the fourth Cluster group of 7 and 10, and the fifth, which consists of Clusters 3, 5, and 6. Therefore, it could be concluded that the six Clusters (1, 4, 8, 9, 11, and 12) formed basically have more homogeneous data compared to the other six Clusters. Thus, 8 large cluster groups were formed with clearly separated centroid distances. The first group consists of a combined Cluster of 1, 4, 8, and 11, while the second consists of clusters 12 and 7, as well as other Cluster groups namely 2, 3, 5, 6, 7, and 10 respectively.

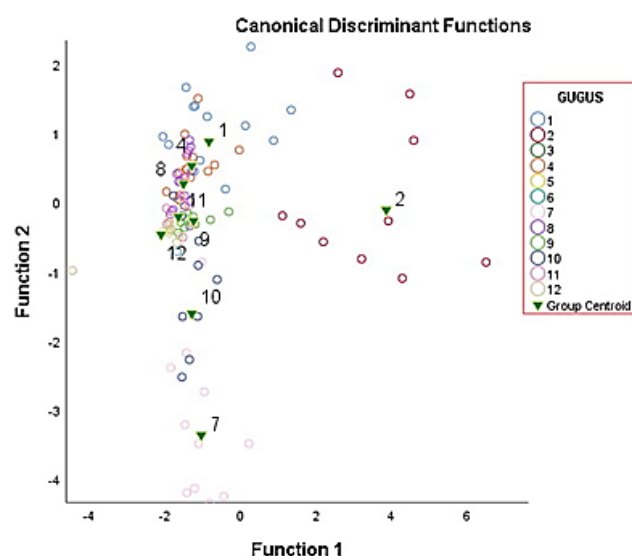


Figure 3. Centroid spread patterns of the twelve island clusters

### 3.5. Discriminant model of twelve island clusters

The model is considered good when the formed variables have high accuracy in describing the real conditions in the field. Furthermore, the discriminant function that is built would give maximum results when it can accurately answer the objectives of the analysis. Analysis of the accuracy of the discriminant function was carried out by observing the eigenvalues and Wilks's Lamda values. Moreover, the function followed the n-1 rule and the number formed was 11 functions. This could be seen in Table 6.

The canonical correlation value measures the closeness of the relationship between the discriminant function and the group (twelve island clusters). Furthermore, Table 6 provides the variation value in the relationship between the function and the clusters. The highest variation in the discriminant function 1 was 0.976 for a scale of 0 to 1. Meanwhile, the lowest value on the 11-discriminant function was 0.047, which was still used for further analysis (Wilks's Lamda value). The discriminant function with a canonical correlation value of 0.976 when squared would produce a R-squared value of 0.9525, which means that 95.25% of the variation in the dependent variable (twelve island clusters) could be explained by fourteen independent variables. Furthermore, for the second discriminant function with a canonical correlation value of 0.903, it was squared to 0.8154, which means that 81.54% of the variation in the dependent variable (twelve island clusters) could be explained by fourteen independent variables, followed by the third to eleventh discriminant functions. Further testing was carried out using the chi-square value and the significant value in the analysis result table.

The results of the analysis showed that the nine island cluster functions had a sig value  $<0.05$ , which means that there was a significant difference in the centroid of the nine discriminant functions produced against the twelve clusters produced. Thus, it could be concluded that there are only nine island clusters that differ in statistical analysis. These results will have an impact on the level of classification accuracy and validation of the resulting discriminant functions.

Table 6. Eigenvalue and Wilks' Lambda value

Function	Eigenvalue	Canonical Correlation	CR <sup>2</sup>	Wilks' Lambda	Sig.
1	19.758	0.976	95.26	0.000	0.000
2	4.413	0.903	81.36	0.003	0.000
3	3.083	0.869	75.69	0.016	0.000
4	2.579	0.849	61	0.065	0.000
5	0.978	0.703	49.14	0.166	0.000
6	0.575	0.604	35.76	0.326	0.000
7	0.546	0.594	34.22	0.507	0.000
8	0.217	0.422	17.89	0.770	0.036
9	0.189	0.399	5.15	0.938	0.666
10	0.036	0.187	1	0.989	0.887
11	0.002	0.047	0.2	0.998	0.675

### 3.6. Eligibility to function in the twelve clusters of islands

The resulting discriminant function needs to be tested for its feasibility and whether it fulfills the statistical rules. The rule that is commonly used is to observe the resulting value of the classification and validation results. When the value is high, it could be concluded that the resulting discriminant function could answer the objectives of the research which include 1) are there any differences in the formed island cluster groups?, 2) if any, which island cluster ?, 3) what variables form it ?, 4) whether the formed discriminant function has a proper level of accuracy, and 5) if it turns out that the resulting discriminant function has not met the eligibility, it is necessary to carry out further testing or reclassification.

The results of the classification and validation test of the discriminant function using fourteen independent variables were 79.7 and 70.3% for the cross-validation test (Table 7). This means that 79.9% of the 1652 data have been entered into groups or clusters according to the original data, with a high level of validation at 70.3%. On the other hand, one of the objectives of this discriminant analysis is to test the ideal grouping, where the cluster grouping truly represents the real conditions in the field (biophysical, socioeconomic). The classification results illustrate that some of the variables forming the discriminant function are classified into other groups, meaning that the differences between groups are uneven and give misclassification to certain groups. Thus, the negative impact that would occur when this function is applied to real conditions is that the planning for equitable development in Maluku Province cannot be maximized. This is because the assessment of the level of similarity of biophysical and socio-economic variables is not correct. This understanding is very important because the basis for the formation of twelve island clusters was to create equitable development in the province.

Referring to the canonical correlation value which measures the closeness of the relationship between the discriminant function and the group (twelve island clusters), there was only a maximum of 8 different island

clusters in discriminant functions (Table 6). Contrary to the results of this analysis, classification could be continued to ascertain more homogeneous cluster grouping. Discrimination analysis could be carried out in several stages, including: 1) identifying clusters that have the lowest classification value and are characterized by errors in their classification results (classified in other groups), 2) combine the low cluster data of the classification results into a new cluster, and 3) reclassify the newly formed grouping until the desired results are obtained.

The first stage of reclassification was by combining Clusters 4 and 11, and the value of classification accuracy for both clusters was 60 and 100%, respectively. This was carried out because 26.7% of Cluster 4 data were classified into Cluster 11. The results of this reclassification produced eleven clusters where the position of Cluster 4 was a combination of itself and Cluster 11. Furthermore, the clusters were reordered starting from 5 to 11. The reclassification results are shown in Table 7.

Table 7. The result of the twelve-island cluster reclassification

Number of clusters	Number of Variable	Overall Accuracy	User Accuracy	Producer Accuracy	Validation
12	12	81.4	87.71	83/36	70.30
11	12	83.90	86.84	85.48	78
9	12	85.60	89.01	86.14	82.2
8	12	87.3	90.32	91.48	83.1
5	10	95.8	93.24	97.00	93.2

The reclassification results obtained by combining two clusters namely Clusters 4 and 11 provide significant effects in increasing the classification accuracy value to 83.9% and the validation test by 79.7% (Table 7). However, these results had no effect on the analysis of discriminant functions, where there was no change in the number of discretionary functions with sig values  $<0.05$  on Wilk's Lambda and canonical correlation. Therefore, it was concluded that there was a better change from the reclassification of this first stage towards improving the accuracy of the resulting classification. The reclassification process continues by combining several clusters at once, namely Clusters 1, 8, and 11 (former cluster 12).

The result of the second stage of reclassification provides a high value of 85.6% and a revalidation of 80.5% (Table 7). Also, the results provide interesting information that even when the value of classification accuracy and validation are high, when viewed from the spread of classification result values, it could be seen that Cluster 5 provides a sufficient value of 66.7%, while other cluster classification results were above 80%, meaning there was a classified Cluster 5 data on Cluster 6 (initial cluster). As stated earlier, one of the purposes of discrimination analysis is to determine the potential of the right island cluster, in order that the reclassification process continues to obtain the ideal cluster for the establishment of the island cluster in Maluku Province.

The third stage of the reclassification which involved combining Clusters 5 and 6 resulted in 8 new island clusters with a classification accuracy of 86.4% and cross-validation of 82.2%. The eight newly formed clusters provide variations in classification accuracy above 80%. The results of the reclassification and distribution of the 8 Clusters could be seen in Table 7. Moreover, the reclassification results combining Clusters 5 and 6 provided high results although there was still a cluster grouping in a given location, which could be seen from the average spread of the value of its discretionary function (centroid). Referring to the centroid value spread of eight island clusters, some clusters are at a very close distance, meaning that several clusters tend to similarize data (homogeneously).

Departing from this situation, the reclassification process continued by combining several island clusters namely Island Clusters 1, 4, and 7 into one Cluster, and Clusters 6 and 8 into the next cluster. The reclassification result brought about five new clusters with an accuracy value of 94.9% and cross-validation of 92.4% (Table 7). Centroids spread the discrimination function of five island clusters and these five clusters could be seen in Figures 4 and 5. The five new Island Clusters consist of Cluster I (combined ex-Cluster Cluster 1, 4, 8, 9, 11, and 12), Cluster 2, 3, 4 (combined ex-Clusters 5 and 6), and Cluster 5 (combined ex-Clusters 7 and 10). Reclassification results provide the best classification level, such as overall classification, user accuracy, producer accuracy, validation, and number of variables. The discrimination function of five clusters with variable 10 could be seen in Table 8.



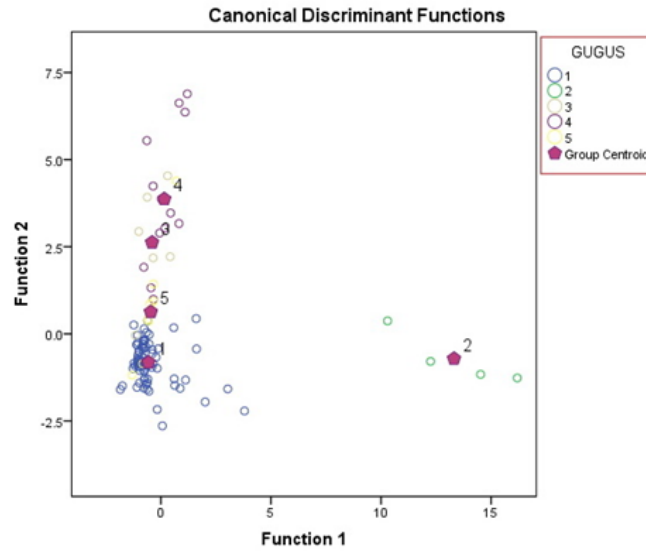


Figure 4. Island cluster five centroid spread pattern

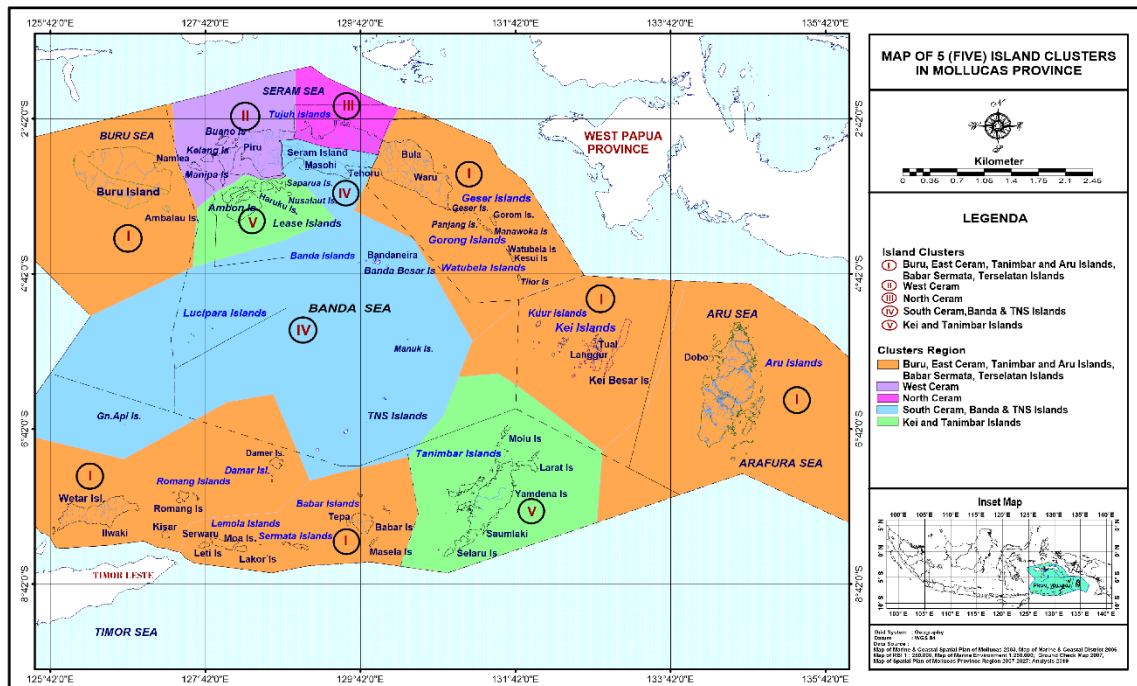


Figure 5. Five island clusters

Table 8. Canonical discriminant function coefficients

Variable	Function			
	1	2	3	4
Forest Cover (Bf)	0.0000032	-0.0000149	0.0000043	-0.0000040
Cluster (Bf)	0.0001230	0.0004237	-0.0001893	-0.0000371
Food Crop Area (Bf)	0.0009996	-0.0000356	-0.0000330	0.0001353
Horticulture Plant Area (Bf)	0.0024049	-0.0000893	0.0000674	-0.0000617
Plantation Area (Bf)	0.0000065	0.0001470	0.0002485	0.0000953
Total School Age Population (S)	-0.0000175	0.0000767	0.0000211	-0.0000629
Number of Capture fisheries Households (S)	0.0002300	0.0002818	0.0000679	-0.0001313
Horticultural Crop Production (E)	-0.0003313	-0.0000827	0.0000079	-0.0000274
Plantation Plant Production (E)	-0.0001608	-0.0002630	-0.0005297	0.0008972
Capture Fisheries Production (E)	0.0000000	0.0000001	0.0000001	0.0000000
(Constant)	-1.120	-0.941	-0.094	-0.532

### 3.7. Best variable selection for discriminant function builders

The final result of the reclassification gave rise to the five best island clusters. Each discriminant function is built as a result of the correlation between the dependent and independent variables. One of the objectives of the discrimination analysis was to find differences between groups or clusters of islands as well as the best variables in building discriminant functions. The result of the reclassification resulted in ten of the best variables of the five discriminant function builders Island clusters. The best variable selection as a result of the discriminant analysis was obtained from the Structure Matrix table and standard table conical linear function coefficients. The matrix structure explains the correlation between independent variables and discriminant functions, while the conical linear function standard shows partial contributions of each variable to the resulting discriminant function. The effect between discriminant functions and free variables is key to ascertaining the influence of the free variables in any discriminated function formed. The most variable correlation in the first discriminant function were variable broad horticultural plant (Bf), cluster area (Bf), and plantation crop production (E).

The discriminant function was built by the contribution of each dependent and independent variable. As mentioned earlier, the Conical linear Function standard aims to provide an overview of the best variables that make up each of the five discriminant functions of the island cluster. The best partial contribution variables with discriminant functions include the variables, broad horticultural crops (Bf), horticultural crop production (E), and plantation crop production (E). Furthermore, the highest partial contribution was from the two variables area of horticultural crops (Bf) and horticultural crop production (E) that have consecutive partial contribution values of 1.4271 and 1.093, respectively.

### 3.8. The role of forest in five island clusters

Forest cover variables were important in building a five-function disinterest in the island cluster. The final reclassification result consisted of 10 best variables and four discriminant functions. Furthermore, the forest cover variable was one of the variables. Results of the analysis on the tables of structures matrix and standardized canonical linear function coefficients of the five island clusters indicated that the forest land cover variable provided good correlation and contribution. The best partial contribution was to the second discriminant function, while the best correlation to the fourth discriminant function. The information of these two tables gives the conclusion that forest cover variables are one of the best for encouraging forestry development in Maluku Province. Therefore, forests become a leading sector or company that could be used as a primary or major economic force supporting economic development in the five island clusters in the province.

## 4. CONCLUSION

Islands clusters in Maluku Province could be grouped into 5 homogeneous clusters with an overall accuracy value of 95.80%, user accuracy 93.24%, Producer accuracy 97%, and cross-validation value of 92.4%. The final results of the reclassification produced the ten best variables that constitute the discriminant function of the clusters. Furthermore, the most dominant variables in the formation of the cluster areas are horticultural crop area (Bf), horticultural crop production (E), plantation plant production (E), and forest land cover (Bf). Finally, the similarity of the biophysical character of an island cluster is not the same as the geographic distribution of that cluster.

## REFERENCES

- [1] W. Beller and P. d'Ayala, "Sustainable development and environmental management of small islands," *Publisher Taylor & Francis*, 1990.
- [2] I. Kelman, T. R. Burns, and N. M. des Johansson, "Islander innovation: A research and action agenda on local responses to global issues," *Journal of Marine and Island Cultures*, vol. 4, no. 1, pp 34-41, 2015, doi: 10.1016/j.imic.2015.04.001.
- [3] A. Polido, E. João, and T. B. Ramos, "Sustainability approaches and strategic environmental assessment in small islands: An integrative review," *Ocean & Coastal Management*, vol. 96, pp 138-148, 2014, doi: 10.1016/j.ocecoaman.2014.05.005.
- [4] A. Romadhon, F. Yulianda, D. G. Bengen, and L. Adrianto, "Socio Ecological Assesment for Tourism at Sapeken Archipelago," *International Journal of Ecosystem*, vol. 3, no. 5, pp. 106-114, 2013, doi: 10.5923/j.ije.20130305.02.
- [5] A. Pamungkas, A. Sulisetyono, Z. Hidayah, and A. K. Sugiyono, "Mainstreaming Development Strategies for Independent Small Island: Case study of Poteran Island, Madura," *Paper presented at 2<sup>nd</sup> Binnale – International Conference on Indonesian Architecture and Planning*, Yogyakarta, 2014.
- [6] M. Scobie, "Policy coherence in climate governance in Caribbean Small Island Developing States," *Environmental Science & Policy*, vol. 58, pp. 16-28, 2016, doi: 10.1016/j.envsci.2015.12.008.

- [7] M. Crossley and T. Sprague, "Education for sustainable development: Implications for small island developing states (SIDS)," *International Journal of Educational Development*, vol. 35, pp. 86-95, 2014, doi: 10.1016/j.ijedudev.2013.03.002.
- [8] G. Scandurra, A. A. Romano, M. Ronghi, and Carfora, "Does climate finance reduce vulnerability in Small Island Developing States? An empirical investigation," *Journal of Cleaner Production*, vol. 256, 2020, doi: 10.1016/j.jclepro.2020.120330.
- [9] E. J. H. Yero, N. C. Sacco, and M. Nicoletti, "Effect of the Municipal Human Development Index on the results of the 2018 Brazilian presidential elections," *Expert Systems with Applications*, vol. 168, 2021, doi: 10.1016/j.eswa.2020.114305.
- [10] S. I. Martínez-Guido, J. B. González-Campos, and J. M. Ponce-Ortega, "Strategic planning to improve the Human Development Index in disenfranchised communities through satisfying food, water and energy needs," *Food and Bioproducts Processing*, vol. 117, pp. 14-29, 2019, doi: 10.1016/j.fbp.2019.06.007.
- [11] [BPS] Central Bureau of Statistics, Human Development Index (IPM) in 2020 by District/City and Province 2020.
- [12] N. Mohammad, A. M. Muad, R. Ahmad, and M. Y. P. Yusof, "Reclassification of Demirjian's mandibular premolars staging for age estimation based on semi-automated segmentation of deep convolutional neural network," *Forensic Imaging*, vol. 24, 2021, doi: 10.1016/j.fri.2021.200440.
- [13] J. F. Hair, R. E. Anderson, R. L. Tatham, and W. C. Black, "Multivariate Data Analysis, 7<sup>th</sup> Edition," Pearson Prentice Hall, 2010.
- [14] K. A. Pituch and J. A. Stevens, "Applied Multivariate Statistics For The Social Sciences-Analyses with SAS and IBM's SPSS," *Sixth edition published 2016 by Routledge 711 Third Avenue*, New York, NY 10017 and by Routledge, © 2016 Taylor & Francis, 2016.
- [15] J. O. Verbel, R. V. Reyes, L. C. P. Londono and B. Johnson, "Discriminant analysis for activation of the aryl hydrocarbon receptor by polychlorinated naphthalenes," *Journal of Molecular Structure: THEOCHEM*, vol. 678, pp. 157-161, 2004, doi: 10.1016/j.theochem.2004.01.048.
- [16] K. Le, C. Chaux, R. Frederic, and E. Guedj, "The adapted LDA and the variable selection procedure are both evaluated on synthetic data, and applied to real data from PET brain images for the classification of patients with Alzheimer's disease," *Computational Statistics & Data Analysis*, vol. 152, np. 1, 2020, doi: 10.1016/j.csda.2020.107031.
- [17] R. T. Lottering, M. Govender, K. Peerbhay, and S. Lottering, "Comparing partial least squares (PLS) discriminant analysis and sparse PLS discriminant analysis in detecting and mapping *Solanum mauritianum* in commercial forest plantations using image texture," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 159, pp. 271-280, 2020, doi: 10.1016/j.isprsjprs.2019.11.019.
- [18] A. Onumanyi *et al.*, "A discriminant analysis-based automatic ordered statistics scheme for radar systems," *Journal Physical Communication*, vol. 43, 2020, doi: 10.1016/j.phycom.2020.101215.
- [19] N. Datta, M. Ashutosh, S. Sarkar, and D. Punetha, "Eradication of elephant mortality and injury due to railway accidents through automatic tracking and alert system," In *2016 International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Spring)*, IEEE, 2016, pp. 1-4, doi: 10.1109/ICACCA.2016.7578856.
- [20] N. Datta, M. Ashutosh, S. Sarkar, and D. Punetha, "Automatic tracking and alarm system for eradication of wildlife injury and mortality," In *2016 2nd International Conference on Advances in Computing, Communication, & Automation (ICACCA)(Fall)*, IEEE, 2016, pp. 1-4, doi: 10.1109/ICACCAF.2016.7748972.
- [21] R. Maalman, Y. Donkor, A. Ayamba, J. Korpisah, and K. Ampong, "Sex estimation using discriminant analysis of hand dimensions among adult population in Ghana," *Forensic Science International: Reports*, vol. 3, 2021, doi: 10.1016/j.fsir.2021.100171.
- [22] M. Balasubramanian, A. Spencer, W. Sohn, and D. Brennan, "The Practice of Dentistry by Australian- and Overseas-Trained Dentists in Australia: Discriminant Analysis of key Predictors," *International Dental Journal*, 2021, doi: 10.1016/j.identj.2021.01.002.
- [23] S. L. Sampaio *et al.*, "Potato biodiversity: A linear discriminant analysis on the nutritional and physicochemical composition of fifty genotypes," *Food Chemistry*, vol. 345, 2021, doi: 10.1016/j.foodchem.2020.128853.
- [24] [BAPPEDA] Regional Development Planning Board, "Report on the Study of the Development Direction of Maluku Development based on Island Clusters. Provincial Government of Maluku, Ambon". (In Bahasa: Laporan Kajian Arah Pengembangan Pembangunan Maluku Berbasis Klaster Pulau. Pemerintah Provinsi Maluku, Ambon), 2014.
- [25] R. A. Johnson and D. W. Wichern, "Applied Multivariate Statistical Analysis," *New Jersey: Pearson Education Inc*, 2007.
- [26] I. Ghozali, "Multivariate Analysis Application with IBM SPSS 25 Edition 9 Program," *Diponegoro University Publishing Board*, 2018.
- [27] S. Santoso, "Proficient in Multivariate Statistics with SPSS," *Publisher PT. Elex Media Komputindo*, Kompas Gramedia, 2018.