PARAMETER
JURNAL MATEMATIKA, STATISTIKA DAN TERAPANNYA

e-ISSN. 2829-6303
Vol. 04, No. 3, December 2025
Page 441-458

Check for updates

Application of Hierarchical Cluster Analysis for Sub-District Grouping Based on Plantation Production

Egis Natasantika Latar[1], Yulia Betania Pusung[2], Nurvia Imran Wael[3], Paskalina Maspaitella[4], Joni Gaitian[5], Henry Junus Wattimanela[6]*

[1,2,3,4,5,6]Statistics Study Program, Faculty of Science and Technology, Pattimura University

Ir. M. Putuhena St, Ambon. 97233, Maluku, Indonesia

***E-mail Correspondence Author:** *hjwattimanela1003@gmail.com*

Abstract

Plantation production is the result of plantation cultivation which includes the of commodities such as coconut, cloves, nutmeg, cocoa, and coffee which have a high sales and economic value and are good for income, as well as being the main source of income for the community's life. Central Maluku Regency is one of the areas in Maluku Province that has great potential in the plantation sector, but the production level between sub-districts shows considerable variation due to differences in geographical conditions, infrastructure, and market access. This study aims to group sub-districts in Central Maluku Regency based on plantation production using a hierarchical cluster analysis approach. Distance measurement uses Single linkage, Average linkage, Complete linkage, and Ward linkage methods. In addition, descriptive analysis, standardization, KMO tests, and multicollinearity tests were carried out with the data used is in the form of plantation production data per sub-district in 2024 sourced from the Central Statistics Agency. Data processing using Microsoft Excel, SPSS, and R applications. The results of the study showed the formation of several sub-district clusters with similar production characteristics, where the Average linkage and Complete linkage methods with the number of clusters as many as three produced the best grouping with the highest Silhouette Score value of 0.477. Each cluster shows the basis for making policies for the development of the plantation sector. These results are expected to be the basis for local governments in formulating plantation development policies that are more targeted and effective.

Keywords: *Central maluku, hierarchical cluster analysis, linkage method, plantation production, silhouette score.*

: https://doi.org/10.30598/parameterv4i3pp441-458

Submitted: July 2025
Accepted: December 2025

1. INTRODUCTION

Central Maluku Regency is an area in Maluku Province that has great potential in the plantation sector. Plantation is a form of agroindustrial activity that begins with the cultivation of certain crops, then the results are processed into industrial raw materials, semi-finished products, or finished products that are ready to be used by consumers. Meanwhile, production refers to activities to create or increase the use value of an item [1]. Its strategic geographical location and supportive tropical climate make this area ideal for the development of superior commodities such as coconut, cloves, nutmeg, cocoa, and coffee. These commodities play an important role in the economy of the local community, both as a main source of income and as a regional export product. This study aims to classify sub-districts in Central Maluku Regency based on plantation commodity production using analysis Cluster Hierarchy. Interpret the results of grouping to understand the structure of regional production characteristics. Group analysis is a method of analysis that aims to select objects in a number of groups that have different characteristics from each other. Purpose of the analysis Cluster is to ensure that the variation of objects in each group is as small as possible or that all members of the group are homogeneous. As well as determining the method Linkage based on evaluation with Silhouette Score [2].

However, plantation production in each sub-district in Central Maluku Regency shows quite significant differences, which are influenced by various factors, such as soil conditions, infrastructure availability, market access, and applicable local policies. To deal with these problems, statistical analysis based on regional grouping is important. One of the relevant approaches is analysis Cluster Hierarchy. This method is used to group objects in sub-districts based on similarity in production data, so that areas with similar characteristics will be incorporated into one Cluster [3].

Analysis Cluster is the process of grouping objects into several small groups, where each group consists of objects that have similarities to each other [4]. Features Cluster good is having a high homogeneity (similarity) between members in one Cluster and has a high heterogeneity (difference) between one Cluster with Cluster other [5]. There are two assumptions that must be met in the analysis Cluster that is, the sample represents the population and does not multicollinearity. To ensure that the sample used can be representative of the population, a value calculation is required Kaiser-Meyer Olkin (KMO). Meanwhile, to find out the existence of multicollinearity, it can be done by calculating Pearson correlation or simple correlation [6]. This process can be done using several methods Linkage, among others single linkage, average linkage, Complete linkage and Ward Linkage [7]. Method single linkage measures the distance between two nearby objects and tends to produce Cluster that is more elongated [8]. Complete linkage measure the distance between the farthest objects, resulting in Cluster which is more compact and clearly separate [9]. While Average Linkage measure the average distance between objects in one Cluster, providing a solution between the two Cluster aforementioned [10].

Several studies have shown the effectiveness of this method in the context of agriculture and plantations. According to Widodo, in implementing the analysis Cluster hierarchy to group districts/cities in Central Java based on data on the production of coconut, coffee, cocoa, and other commodities. As a result, three main groups were formed that described regions with different production characteristics, which are useful for supporting commodity-based regional policies [11]. Another study by Taufik applied the Cluster hierarchy to map potential areas for chili cultivation in Indonesia. The study showed that the grouping of areas by the hierarchy very effective to support horticultural crop cultivation policies nationally [12].

Other research uses the Cluster Partitioning and Hierarchy to group provinces in Indonesia based on their poverty characteristics. The findings of this study indicate that the Cluster Hierarchy with two Cluster provide the most optimal grouping results. In addition, an analysis has been carried out Cluster Hierarchy to compare methods Complete linkage and the Ward Linkage in grouping the human development index in the South Sulawesi region. The results of the study showed that the Ward Linkage with five Cluster gives the best grouping results, indicated by the value of the smallest standard deviation ratio of 0.282 [5]. According to research conducted by [13] Comparison of standard deviation ratios, method Complete linkage produces the smallest value of 0.222, which indicates that this method is superior to single linkage and Average Linkage in the grouping of sub-districts in Sidoarjo Regency based on the variable type of livestock.

Thus, the novelty of this research lies in the specific regional focus, namely Central Maluku Regency, in the application of hierarchical cluster analysis for plantation production data, as well as in the comparative approach to four linkage methods evaluated using the Silhouette Score quantitatively. This research makes a significant contribution in supporting plantation development policy making that adopts a cluster approach and is responsive to local conditions.

2. RESEARCH METHODS

2.1. Types of research

The type of research used is quantitative approach, which is by utilizing data in the form of numbers that have been available, then analyzed and compiled in accordance with the research objectives regarding the analysis Cluster hierarchy in grouping plantation production according to sub-districts in Central Maluku Regency.

2.2. Tools and Materials

The tools used in this study are laptops and supporting software, namely SPSS, R and Microsoft excel. Meanwhile, the materials used in the study are in the form of secondary data obtained from the Central Statistics Agency (CSA) of Central Maluku Regency in 2024. The data taken is data on Plantation Production by District and Plant Type in Central Maluku Regency. There are six variables that represent plantation production and in each variable there are as many as 18 sub-districts in Central Maluku Regency. The data in the study can be seen in Figure 1.

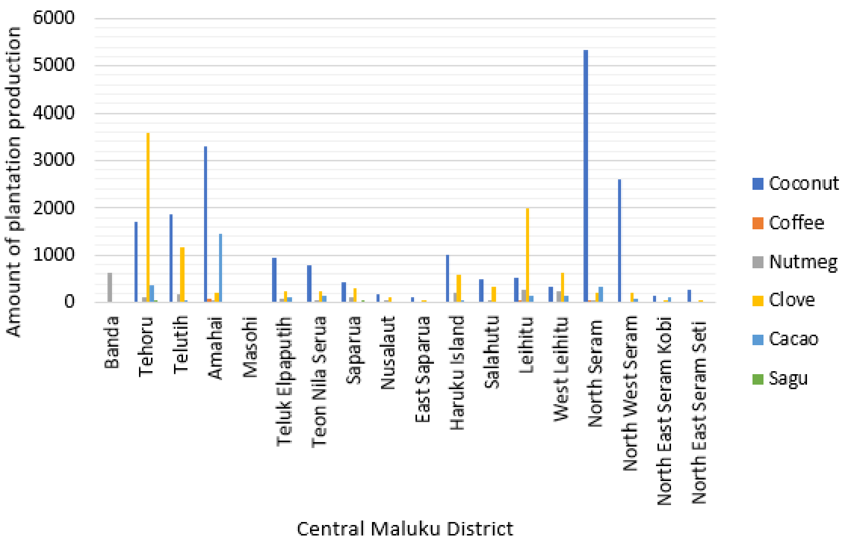

Figure 1. Plantation Production Research Data

The following are the variables of plantation production According to 18 Districts and Plant Types in Central Maluku Regency (Ton) used in this study shown in **Table 1**.

Table 1. Research Variables

Variable	Description of Plantation Production
X_1	Coconut
X_2	Coffee
X_3	Nutmeg
X_4	Clove
X_5	Cacao
X_6	Sagu

2.3. Research Procedure

The procedure used in this study is an analysis Cluster Hierarchy, with the steps of implementing the research described as follows **[14]**:

1. Collecting plantation production data by sub-district and plant type in Central Maluku Regency in 2024.
2. Analyze descriptive statistics based on research variables
   The form of presentation of the results of descriptive analysis depends on the type or scale of data from the variables being analyzed **[15].**
3. Standardizing data using Z-Score values
   The way to determine the standardization value is by calculating the mean and standard deviation values of each variable.
   $$Z = \frac{x_i - \bar{x}}{s} \tag{1}$$
4. Testing the cluster analysis assumptions where there are two tests, namely:
   a. KMO Test
   This test is done to find out if the sample should be representative of the population
   $$\mathrm{KMO} = \frac{\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij}^2}{\sum_{i=1}^{p}\sum_{j=1}^{p} r_{ij}^2 + \sum_{i=1}^{p}\sum_{j=1}^{p} a_{ij}^2} \tag{2}$$
   b. Multicollinearity Test
   This test was carried out to find out if there was a high linear relationship between the research variables.
   $$\mathrm{VIF} = \frac{1}{\mathrm{Tol}} \tag{3}$$
   $$\mathrm{Tol} = (1 - R^2) \tag{4}$$
   If there are no symptoms of multicollinearity, the VIF value < 10 and the tolerance value > 0,10 **[16]**.
5. Measuring distances between objects.
   Distance measurement was carried out using Euclidean distance distancce
   $$d_{ik} = \sqrt{\sum_{j=1}^{p}(x_{ij} - x_{kj})^2} \tag{5}$$
6. Perform the stages of hierarchical cluster analysis
   In this study, an analysis was used Cluster hierarchy by using the Single linkage, Average linkage, Complete linkage and Ward linkage as follows **[17].**

Single linkage

$$d_{(UV)W} = \min(d_{UW}, d_{VW}) \quad (6)$$

Where the values $\boldsymbol{d_{UW}}$ **and** $\boldsymbol{d_{VW}}$ are the minimum distance between cluster U and cluster W and cluster V and cluster W [18].

Average linkage

$$d_{(UV)W} = \frac{\sum a \sum b\, d_{ab}}{N_{(UV)} N_W} \quad (7)$$

Where d_{ab} is the distance between object a in cluster (UV) and object b in cluster W, $N_{(UV)}$ is the number of objects in cluster (UV) and N_W is the number of objects in cluster W [19].

Complete linkage

$$d_{(UV)W} = \max(d_{UW}, d_{VW}) \quad (8)$$

Where the values d_{UW} and d_{VW} are the distances between the furthest neighbors of cluster U and cluster W and cluster V and cluster W [20].

Ward linkage

The Ward method aims to minimize the increase in total variance in the *Sum of Squares Error* (SSE) cluster each time two clusters are merged. Two objects will be merged if they have the smailest object among the existing possibilities [5].

$$\Delta SSE_{ij} = \frac{n_i \times n_j}{n_i + n_j} \times d(i,j)^2 \quad (9)$$

Where ΔSSE_{ij} is the values of sum of squares error is cluster i and cluster j are merged, n_i is the number of objects in cluster i, n_j is the number of objects in cluster j and $d(i, j)$ is the euclidean distance of cluster i and j.

7. Determine the number of clusters and cluster members that are formed.
8. Determine the best method with the optimal number of clusters using the Silhouette Score method.

$$\boldsymbol{s(i) = \frac{b(i) - a(i)}{\max\,(a(i), b(i))}} \quad (10)$$

Where s(i) is the value *silhoutte* for the ith data point, a(i) is the average distance between the ith point to all other points in the same cluster, and b(i) is the average distance between the ith point to points in the other nearest cluster [2].

9. Define cluster characteristics.
10. Interpreting the results of the analysis.
11. Draw conclusions and suggestions.

3. RESULTS AND DISCUSSION

3.1. Descriptive Statistical Analysis

The data analyzed were plantation production data by sub-district and crop type in Central Maluku Regency in 2024, six variables represented plantation production, each representing 18 sub-districts in Central Maluku Regency. The following is a description of the 2024 plantation production data.

Descriptive statistics can be seen in **Table 2**, as follows:

Table 2. Descriptive Statistics

Variable	N	Min	Max	Average	Baku Junction
X_1	18	8.000	5346.000	1118.561	1410.111
X_2	18	0.000	77.600	13.779	23.637
X_3	18	0.870	626.290	121.403	150.984
X_4	18	2.750	3597.000	554.400	905.266
X_5	18	0.000	1450.000	168.021	337.913
X_6	18	0.000	38.100	6.894	12.274

Referring to **Table 2**, the variables X_1 show a minimum value of 8.000 and a maximum of 5346.000, with an average value of 1118.561 and a standard deviation of 1410.111. The variables X_2 have a range of values between 0.000 to 77.600, with an average of 13.779 and a standard deviation of 23.637. For the variable X_3, the minimum value is 0.870 and the maximum is 626.290, with an average of 121.403 and a standard deviation of 150.984. On the variable X_4, the minimum value is 2.750 and the maximum is 3597.000, with an average of 554.400 and a standard deviation of 905.266. The variables X_5 have a minimum value of 0.000 and a maximum of 1450.000, with an average value of 168.021 and a standard deviation of 337.913. Meanwhile, the variables X_6 show a minimum value of 0.000 and a maximum of 38.100, with an average of 6.894 and a standard deviation of 12.274.

3.2. Data Standardization

The standardized value of all observation data on the variables $\mathbf{X_i}$ can be seen in **Table 3**, as follows:

Table 3. The Value of Standardization of Plantation Production Data

No	ZX_1	ZX_2	...	ZX_6
1.	-0.787	-0.583	...	-0.561
2.	0.423	-0.583	...	2.534
⋮	⋮	⋮	⋱	⋮
18.	-0.604	-0.583	...	-0.561

Based on **Table 3**, if the standardized values of plantation production data using the z-score method, where each variable is expressed as its deviation from the mean. Positive values indicate that the observation is above the average, while negative values show that it is below the average. This standardization ensures that all variables are on the same scale, making the data more comparable and suitable for further analysis, such as clustering or multivariate analysis.

3.3. Cluster Analysis Assumption Test

The assumption test carried out is in the KMO test, as follows:

The KMO test result showed a value of 0.576. This value is within the range of 0.5 – 1, so the sample is considered sufficiently representative for further analysis. Next, the assumption test of cluster analysis was carried out using the multicollinearity test. The results of the multicollinearity test using VIF and Tolerance values can be seen in Table 5.

Table 5. Multicollinearity Test

Type	Collinearity Statistics	
	Tolerance	VIF
(Constant)		
X_1	0.560	1.786
X_2	0.354	2.823
X_3	0.860	1.162
X_4	0.703	1.423
X_5	0.373	2.683
X_6	0.604	1.656

Based on Table 5, if the variables have a tolerance value of > 0.10 and VIF < 10, then it can be concluded that there are no symptoms of multicollinearity between variables.

3.4. Euclidean Distance Calculation

Euclidean distances are obtained with the help of SPSS software and one example of the calculation is as follows:

$$d_{1,2} = \sqrt{\sum_{j=1}^{6}(x_{ij} - x_{kj})^2}$$

$$d_{1,2} = \sqrt{(x_{11} - x_{21})^2 + \cdots + (x_{16} - x_{26})^2}$$

$$d_{1,2} = \sqrt{(-0.78757 - 0.42368)^2 + \cdots + (-0.56173 - 2.53432)^2} = 6.225$$

The following can be seen the value of Euclidean distance in Table 6.

Table 6. Euclidean Distance

No	District	Euclidean Distance				
		Banda	Tehoru	Telutih	...	North East Seram Seti
1.	Banda	0.000	6.225	3.472	...	4.049
2.	Tehoru	6.225	0.000	4.218	...	5.265
3.	Telutih	3.472	4.218	0.000	...	2.011
⋮	⋮	⋮	⋮	⋮	⋱	⋮
18.	North East Seram Seti	4.049	5.265	2.011	...	0.000

Based on Table 6, it can be seen that the closest characteristic distance between Banda District and Tehoru District is 6.225, while the closest distance between Banda District and Telutih District is 3.472. Likewise for other nearby ranges. In this case, it can be concluded that Banda District and Tehoru District are more similar in their proximity compared to Telutih District. The reference is that the smaller the Euclidean distance value, the more similar the characteristics.

3.5. Hierarchy Cluster Analysis

The next stage is to carry out the hierarchical clustering process using the Single linkage, Average linkage, Complete linkage, and Ward linkage methods. In the single linkage method, the dendogram form is obtained in Figure 2. As follows:

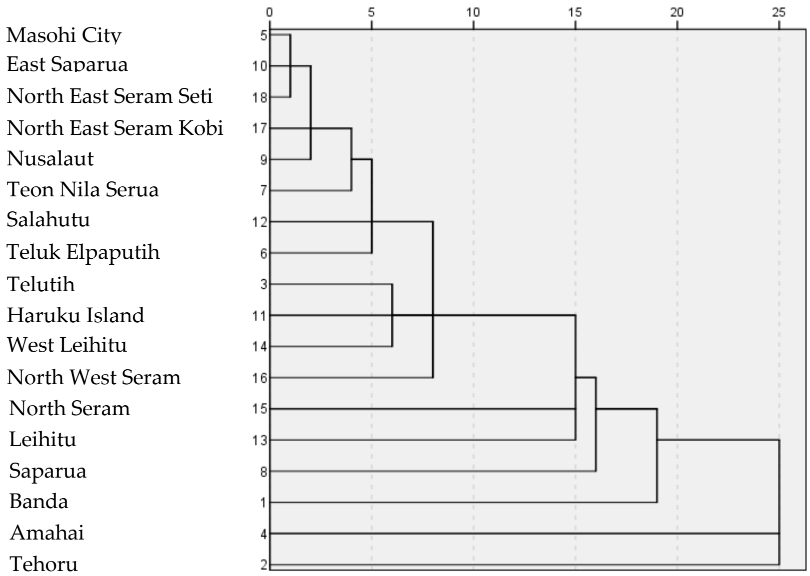

Figure 2. **Dendogram using the Single linkage method**

Based on the dendogram results of the single linkage method, the grouping process is carried out in stages based on the closest distance between objects. The dendogram is cut off at a specific point that results in the main cluster. The results of the process of determining the cluster members of the Single linkage method can be seen in **Table 7**.

Table 7. **Cluster Single linkage Members**

No	District	Clusters		
		4	3	2
1.	Banda	1	1	1
2.	Tehoru	2	2	2
3.	Telutih	3	1	1
4.	Amahai	4	3	1
5.	Masohi	3	1	1
6.	Teluk Elpaputih	3	1	1
7.	Teon Nila Serua	3	1	1
8.	Saparua	3	1	1
9.	Nusalaut	3	1	1
10.	East Saparua	3	1	1
11.	Haruku Island	3	1	1
12.	Salahutu	3	1	1
13.	Leihitu	3	1	1
14.	West Leihitu	3	1	1
15.	North Seram	3	1	1
16.	North West Seram	3	1	1
17.	North East Seram Kobi	3	1	1
18.	North East Seram Seti	3	1	1

Based on **Table 7**, it can be seen that the clustering of sub-districts in Central Maluku Regency based on Plantation Production in 2024 with Single linkage is:

a. If 4 clusters are formed, the sub-districts that are included in cluster 1 according to plantation production are only 1 sub-district namely Banda. Only Tehoru District is included in cluster 2. Cluster 3 includes 15 sub-districts, including Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, North Seram, North West Seram, North East Seram Kobi, and North East Seram Seti. Furthermore, only Amahai District is included in cluster 4.

b. If formed in 3 clusters, which are included in cluster 1 according to plantation production, there are 16 sub-districts, including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, North Seram, North West Seram, North East Seram Kobi, and North East Seram Seti. Cluster 2 is only Tehoru District and Cluster 3 is only Amahai District.

c. If formed in 2 clusters, then included in cluster 1 according to plantation production there are 17 sub-districts including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, North Seram, West Seram, North West Seram, North East Seram Kobi, and North East Seram Seti. What is included in cluster 2 is only Tehoru District.

Furthermore, in cluster analysis using the average linkage method, the form of a dendogram was obtained in **Figure 3**. As follows:

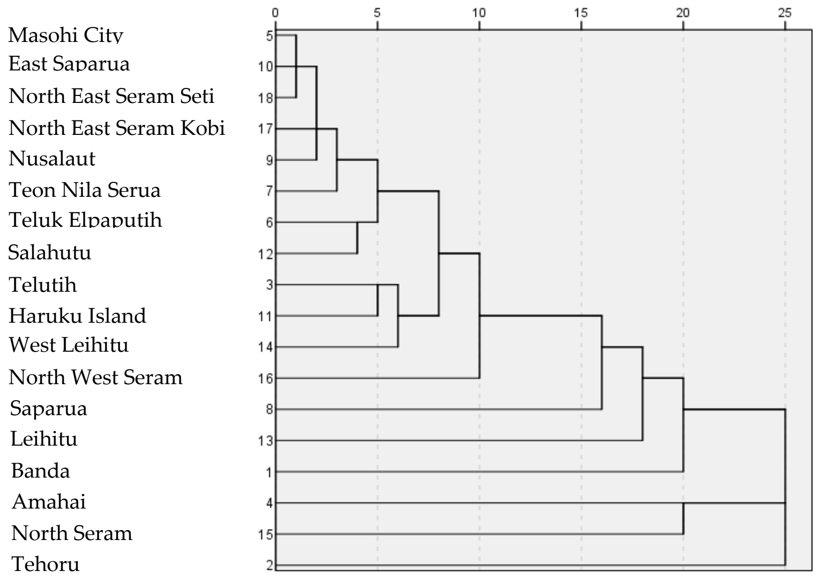

Figure 3. Dendogram using the Average linkage method

Based on the dendogram of the average linkage method, the grouping process is carried out based on the closest distance between objects. The dendogram is cut off at a specific point that results in the main cluster. The results of defining the cluster member process of the Average linkage method can be seen in **Table 8**.

Table 8. Cluster Average linkage members

No	district	Clusters		
		4	**3**	**2**
1.	Banda	1	1	1
2.	Tehoru	2	2	2
3.	Telutih	1	1	1
4.	Amahai	3	3	1
5.	Masohi	1	1	1
6.	Teluk Elpaputih	1	1	1
7.	Teon Nila Serua	1	1	1
8.	Saparua	1	1	1
9.	Nusalaut	1	1	1
10.	East Saparua	1	1	1
11.	Haruku Island	1	1	1
12.	Salahutu	1	1	1
13.	Leihitu	1	1	1
14.	West Leihitu	1	1	1
15.	North Seram	4	3	1
16.	North West Seram	1	1	1
17.	North East Seram Kobi	1	1	1
18.	North East Seram Seti	1	1	1

Based on **Table 8**, it is seen that the clustering of sub-districts in Central Maluku Regency based on Plantation Production in 2024 with the average linkage is:

a. If 4 clusters are formed, then the sub-districts included in cluster 1 according to plantation production there are 15 sub-districts, including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, West North Seram, North East Seram Kobi, and North East Seram Seti. Only Tehoru District is included in cluster 2. Cluster 3 is only Amahai District and cluster 4 is only North Seram District.
b. If formed in 3 clusters, then those included in cluster 1 according to plantation production are 15 sub-districts, including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, North Seram, North West Seram, North East Seram Kobi, and North Seram East Seti. Cluster 2 is only Tehoru District and Cluster 3 is Amahai and North Seram Districts.
c. If formed in 2 clusters, then there are 17 sub-districts including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, North Seram, West Seram, North West Seram, North East Kobi Seram, and East North Seti. What is included in cluster 2 is only Tehoru District.

Furthermore, in the cluster analysis with the complete linkage method, the dendogram form was obtained in **Figure 4**. As follows:

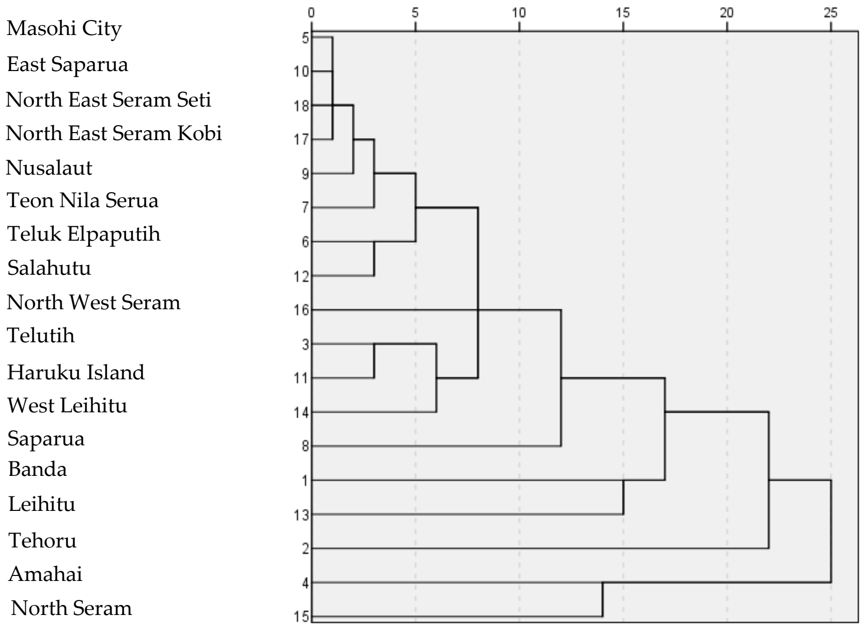

Figure 4. Dendogram using the Complete linkage method

Based on the dendogram results of the complete linkage method, the grouping process is carried out based on the closest distance between objects. The dendogram is cut off at a specific point that results in the main cluster. The results of determining the process of cluster members of the Complete linkage method can be seen in **Table 9**.

Table 9. Cluster Complete linkage members

No	District	Clusters		
		4	3	2
1.	Banda	1	1	1
2.	Tehoru	2	2	1
3.	Telutih	3	1	1
4.	Amahai	4	3	2
5.	Masohi	3	1	1
6.	Teluk Elpaputih	3	1	1
7.	Teon Nila Serua	3	1	1
8.	Saparua	3	1	1
9.	Nusalaut	3	1	1
10.	East Saparua	3	1	1
11.	Haruku Island	3	1	1
12.	Salahutu	3	1	1
13.	Leihitu	1	1	1
14.	West Leihitu	3	1	1
15.	North Seram	4	3	2
16.	North West Seram	3	1	1
17.	North East Seram Kobi	3	1	1
18.	North East Seram Seti	3	1	1

Based on **Table 9**, it is seen that the clustering of sub-districts in Central Maluku Regency in 2024 Plantation Production with Complete linkage is:

a. If 4 clusters are formed, the sub-districts that are included in cluster 1 according to plantation production there are 2 sub-districts, namely Banda and Leihitu Districts, which are included in cluster 2 only Tehoru District. Included in cluster 3 are 13 sub-districts, including Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, West Leihitu, West North Seram, North East Kobi Seram, and East North Seti. Which includes cluster 4 of Amahai and North Seram Districts.

b. If formed in 3 clusters, which are included in cluster 1 according to plantation production, there are 15 sub-districts, including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, West North Seram, North East Kobi Seram, and East North Seti. Cluster 2 is only Tehoru District and Cluster 3 is Amahai and North Seram Districts.

c. If formed in 2 clusters, then there are 16 sub-districts including Banda, Tehoru, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, West Seram, North est Seram, North East Kobi Seram, and East North Seti. Those included in cluster 2 are Amahai and North Seram Districts.

Furthermore, in the cluster analysis using the ward linkage method, the dendogram form was obtained in **Figure 5**. As follows:

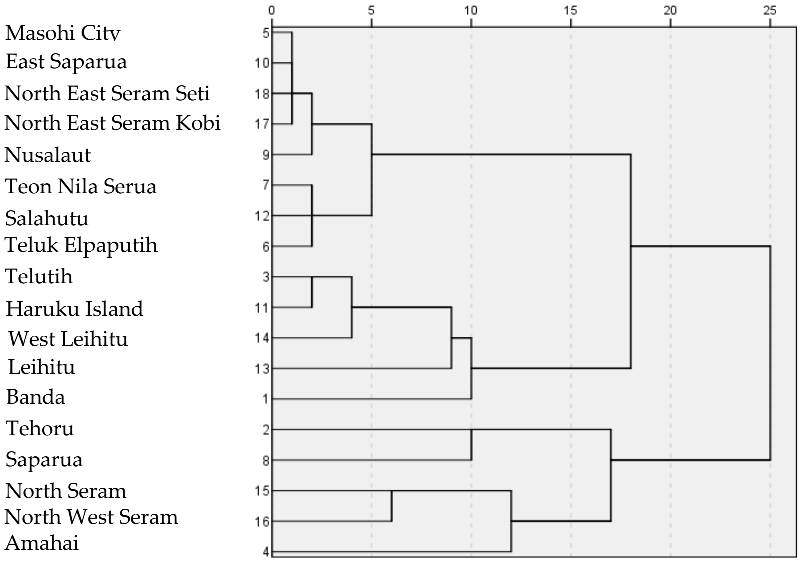

Figure 5. Dendogram using the Ward method

Based on the dendogram results of the ward linkage method, the grouping process is carried out based on the closest distance between objects. The dendogram is cut off at a specific point that results in the main cluster. The results of defining the Ward method cluster member process can be seen in **Table 10**.

Table 10. Ward Cluster Members

No	District	Clusters		
		4	**3**	**2**
1.	Banda	1	1	1
2.	Tehoru	2	2	2
3.	Telutih	1	1	1
4.	Amahai	3	2	2
5.	Masohi	4	3	1
6.	Teluk Elpaputih	4	3	1
7.	Teon Nila Serua	4	3	1
8.	Saparua	2	2	2
9.	Nusalaut	4	3	1
10.	East Saparua	4	3	1
11.	Haruku Island	1	1	1
12.	Salahutu	4	3	1
13.	Leihitu	1	1	1
14.	West Leihitu	1	1	1
15.	North Seram	3	2	2
16.	North West Seram	3	2	2
17.	North East Seram Kobi	4	3	1
18.	North East Seram Seti	4	3	1

Based on **Table 10**, it is seen that the clustering of sub-districts in Central Maluku Regency based on Plantation Production in 2024 with Ward linkage is:

a. If 4 clusters are formed, the sub-districts that are included in cluster 1 according to plantation production are 5 sub-districts, namely Banda, Telutih, Haruku Island, Leihitu and West Leihitu Districts, which are included in cluster 2 there are 2 sub-districts, namely Tehoru and Saparua Districts. Which includes cluster 3 there are 3 sub-districts, including Amahai, North Seram and West North Seram. Included in cluster 4 are 8 sub-districts, namely Masohi, Teluk Elpaputih, Teon Nila Serua, Nusalaut, East Saparua, Salahutu, North East Kobi District, and East North Seti.
b. If formed in 3 clusters, which are included in cluster 1 according to plantation production, there are 5 sub-districts, including Banda, Telutih, Haruku Island, Leihitu, and West Leihitu Districts. Included in cluster 2 are 5 sub-districts, namely Tehoru, Amahai, Saparua, North Seram and West North Seram. Cluster 3 includes 8 sub-districts, namely Masohi, Teluk Elpaputih, Teon Nila Serua, Nusalaut, East Saparua, Salahutu, North East Kobi District, and East East North Seti.
c. If formed in 2 clusters, then there are 13 sub-districts including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, East North Seram Kobi, and North East Seram Seti. Included in cluster 2 are 5 sub-districts, namely Tehoru, Amahai, Saparua, North Seram, and West North Seram.

3.6. Best Method with Silhoutte Score

The determination of the best method using the Silhoutte Score of each cluster using the single linkage, average linkage, complete linkage and ward linkage methods can be seen in **Table 11**.

Table 11. Silhoutte Score

Multiple Clusters	Silhoutte Score			
	Single Linkage	Average Linkage	Complete Linkage	Ward Linkage
2	0.432	0.432	0.475	0.372
3	0.446	0.477	0.477	0.263
4	0.368	0.394	0.410	0.305

Based on Table 11, the highest silhoutte score value of 0.477 was obtained in the average linkage and complete linkage methods with a large number of clusters, namely 3 clusters, thus showing that these two methods provide the most optimal cluster separation and good cohesiveness between data in the cluster. Therefore, the average linkage and complete linkage method with 3 clusters is the best choice in the data grouping process in this analysis.

3.7. Characteristics of each cluster

After obtaining the best method, namely average linkage, the next stage is to identify the characteristics of each cluster. The purpose of this step is to provide a special overview of the plantation production variables in each cluster formed, and calculate the mean value of each cluster on the average linkage method. The results of the mean calculation for each cluster can be seen in Table 12.

Table 12. Mean Variables of Each Cluster

Cluster	Mean Variables of Each Cluster						Total Mean Cluster
	$\mathbf{X_1}$	$\mathbf{X_2}$	$\mathbf{X_3}$	$\mathbf{X_4}$	$\mathbf{X_5}$	$\mathbf{X_6}$	
1	-0.332	-0.226	0.054	-0.172	-0.325	-0.243	-0.208
2	0.424	-0.583	0.057	3.361	0.624	2.534	1.070
3	2.277	1.990	-0.430	-0.390	2.130	0.558	1.022

Referring to Table 12, the characteristics of the three clusters formed show that cluster 1 has a lower average value on the variable $\mathbf{X_1}$, $\mathbf{X_5}$, and $\mathbf{X_6}$ compared to the other clusters. Meanwhile, cluster 2 shows a higher mean value on the variables $\mathbf{X_3}$, $\mathbf{X_4}$, and $\mathbf{X_6}$, but has the lowest mean value of the variable $\mathbf{X_2}$. On the other hand, cluster 3 recorded a higher average value for the variables $\mathbf{X_1}$, $\mathbf{X_2}$, and $\mathbf{X_5}$, but the mean value for the variable $\mathbf{X_3}$ and $\mathbf{X_4}$ was lower than the other cluster.

Based on the results of the research conducted, it is known that plantation production according to sub-districts in Central Maluku Regency is divided into three groups (clusters). So that there are three sub-district groups in Central Maluku Regency with different plantation production.

Cluster 1 consists of 15 sub-districts, namely Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, North West Seram, North East Seram Kobi and East North Seram Seti. In this cluster, it can be seen that plantation production has a very low mean total among other clusters with an mean value of -0.208. In cluster 2 only consists of 1 sub-district, namely Tehoru District has a very high average total plantation production among other clusters with an mean value of 1.070 and in cluster 3 consists of 2 sub-districts, namely Amahai and North Seram Districts which have a total mean of medium livestock production with an mean value of 1.022.

The results of this study are consistent with the findings of Djafar et al. [11], who classified districts and cities in Central Java based on plantation production and obtained three main clusters. From the methodological perspective, Sanusi and Hasanah [5] found that the Ward method provided the most optimal results for clustering the human development index in South Sulawesi, whereas this study shows that the Average Linkage and Complete Linkage methods were more suitable for the characteristics of plantation data in Central Maluku. Therefore, this study strengthens the existing literature while providing a new contribution by focusing on the plantation sector in Central Maluku, the results of which can be the basis for formulating regional development policies based on production clusters.

4. CONCLUSION

This study successfully applied the hierarchical cluster analysis method to group 18 sub-districts in Central Maluku Regency based on the production of six types of coconut, coffee, nutmeg, cloves, cocoa, and sago plantation commodities. The four linkage methods used are Single Linkage, Average Linkage, Complete Linkage, and Ward Linkage. The average linkage and complete linkage methods produced three clusters with the highest Silhouette Score value of 0.477, so it was considered the most optimal and representative. The three clusters that were formed, namely, cluster 1 consisted of 15 sub-districts including Banda, Telutih, Masohi, Teluk Elpaputih, Teon Nila Serua, Saparua, Nusalaut, East Saparua, Haruku Island, Salahutu, Leihitu, West Leihitu, West North Seram, North East Seram Kobi, and East North Seram Seti. Cluster 2 consists of Tehoru, and cluster 3 consists of Amahai and North Seram. In cluster 1, it has a very low mean total plantation production with a value of -0.208. In cluster 2 it has a very high mean plantation production total with a value of 1.070 and in cluster 3 it has a medium mean plantation production total value of 1.022. The results of this clustering can be used as a basis for planning and development policies for the plantation sector in areas that have similar production characteristics.

Acknowledgments

All praise and gratitude are given to the Lord Jesus for His grace, blessings, and guidance, which enabled the completion of this article. The authors also express their deepest appreciation to Mr. Henry Junus Wattimanela, lecturer of the Island Sea Analysis course and academic advisor, for his valuable input and suggestions in the preparation of this article. Special thanks are extended to the co-authors, Egis Natasantika Latar, Yulia Betania Pusung, Nurvia Imran Wael, Paskalina Maspaitella, and Joni Gaitian, for their collaboration, dedication, and contributions to this research. The authors also give special appreciation as this article represents our first published work, making it a valuable experience as well as motivation for future research. In addition, the authors sincerely thank the Central Maluku Statistics Agency (BPS Maluku Tengah) for providing the secondary data that served as the basis for this analysis.

Funding Information

The authors declare that there was no funding involvement. This research was conducted as part of a course assignment without external financial support.

Author Contributions Statement

Egis Natasantika Latar: Conceptualization, method, formal analysis, initial draft writing, software implementation. Yulia Betania Pusung: Investigation, resources, data processing, validation. Nurvia Imran Wael: Data curation, validation, writing – review & editing. Paskalina Maspaitella: Method, formal analysis, visualization, writing – review & editing. Joni Gaitian: Software, data analysis, validation, writing – review & editing. All authors discuss the Hierarchical Cluster method, analyze plantation production, and contribute to the interpretation of planning policies and the development of the plantation sector in regions with similar production characteristics.

Conflict of Interest Statement

The authors declare that they have no competing financial interests or personal relationships that could have influenced the work reported in this paper. This research was conducted purely for academic purposes as part of a course requirement. The authors state that there is no conflict of interest.

Informed Consent

Not applicable. This study used publicly available secondary data from the Central Bureau of Statistics (BPS) publication "Central Maluku Regency in Figures" and did not involve human subjects requiring informed consent.

Ethical Approval

Not applicable. This study uses publicly available plantation production data and applies Hierarchical Cluster methods without involving human or animal subjects. Ethical approval is not required for this type of secondary data analysis.

Data Availability

The plantation production data supporting this research finding is publicly available through the official publication "Central Maluku Regency in Figures" published by the Central Maluku Regency Statistics Agency (BPS), accessible at https://malukutengahkab.bps.go.id/id. Processed data, including matrices, transition probability calculations, and Hierarchical Cluster analysis results supporting the conclusions, are available from the corresponding author upon reasonable request. Implementations in Microsoft Excel and R software used for the calculation of steady-state distributions and plantation production can also be provided to researchers interested for reproduction purposes.

REFERENCES

[1] I. Arsyad, S. Maryam, I. Arsyad, S. Maryam, F. Pertanian, and U. Mulawarman, "[analisis faktor – faktor yang mempengaruhi produksi kelapa sawit pada kelompok tani sawit," vol. 14, no. 1, pp. 75–85, 2017.

[2] N. Afira and A. W. Wijayanto, "Analisis Cluster Kemiskinan Provinsi di Indonesia Tahun 2019 dengan Metode Partitioning dan Hierarki Cluster Analysis with Partitioning and Hierarchical Methods on Provincial Poverty Information Data in Indonesia in 2019," vol. 10, no. 28, 2021.

[3] A. N. Riadhoh, G. E. Puspita, I. Rafidah, and E. Widodo, "Pengelompokan Kabupaten/Kota Berdasarkan Produksi Tanaman Pangan Sumatera Utara Tahun 2020 Menggunakan Pengelompokan Hirarki Aglomeratif," *KUBIK J. Publ. Ilm. Mat.*, vol. 6, no. 2, pp. 61–70, 2022.

[4] A. N. Fathia and R. Rahmawati, "Analisis klaster kecamatan di kabupaten semarang berdasarkan potensi desa menggunakan metode ward dan single linkage," vol. 5, pp. 801–810, 2016.

[5] W. Sanusi and A. Hasanah, "Perbandingan Analisis Cluster Metode Complete Linkage dan Metode Ward dalam Pengelompokkan Indeks Pembangunan Manusia di Sulawesi Selatan," vol. 7, no. 1, pp. 75–86, 2024.

[6] D. Ls, Y. A. Lesnussa, M. W. Talakua, and M. Y. Matdoan, "Analisis klaster untuk pengelompokkan Kabupaten/Kota di Provinsi Maluku berdasarkan indikator pendidikan dengan menggunakan metode Ward," *Stat. dan Apl.*, vol. 5, no. 1, pp. 51–60, 2021.

[7] M. Y. Matdoan and D. Noya Van, "Penerapan Analisis Cluster Dengan Metode Hierarki Untuk Klasifikasi Kabupaten/Kota Di Provinsi Maluku Berdasarkan Indikator Pembangunan Manusia," vol.

2, no. 2, pp. 123–130, 2020.

[8] E. L. Heatubun, A. Z. Wattimena, and H. Batkunde, "ANALISIS KLASTER DAERAH SAMPAH MENGGUNAKAN METODE SOM , SINGLE LINKAGE DAN AVERAGE LINKAGE" vol. 03, no. 01, pp. 33–48, 2024.

[9] Y. Reinaldi, N. Ulinnuha, T. Hartono, and M. Hafiyusholeh, "Perbandingan metode single linkage, complete linkage, dan average linkage pada kesejahteraan masyarakat di Jawa Timur," *Mat. Stat. Komputasi*, vol. 18, no. 1, pp. 130–140, 2021.

[10] A. L. Yusniyanti, F. Virgantari, and Y. E. Faridhan, "Comparison of Average Linkage and K-Means Methods in Clustering Indonesia's Provinces Based on Welfare Indicators," *J. Phys. Conf. Ser.*, vol. 1863, no. 1, 2021.

[11] N. M. Djafar, L. N. Wijayanti, A. R. Elprilita, and E. Widodo, "PENGELOMPOKAN PRODUKSI PERKEBUNAN MENURUT KABUPATEN/KOTA JAWA TENGAH TAHUN 2020 MENGGUNAKAN HIERARCHICAL CLUSTERING," vol. 4, no. 2, pp. 59–66, 2021.

[12] R. A. Rizko, E. Rahman, S. Sujiono, and T. Taufik, "Segmentasi Wilayah Indonesia Yang Berpotensi Budidaya Cabai Dengan Hierarki Clustering," *J. Tek. Ind. Terintegrasi*, vol. 6, no. 4, pp. 1223–1235, 2023.

[13] S. Fikri and N. Ulinnuha, "Perbandingan Metode Single Linkage , Complete Linkage Dan Average Linkage dalam Pengelompokan Kecamatan Berdasarkan Variabel Jenis Ternak Kabupaten Sidoarjo," vol. 4, no. 2, 2019.

[14] M. Gagolewski, A. Cena, S. James, and G. Beliakov, "Hierarchical clustering with OWA-based linkages , the Lance – Williams formula , and dendrogram inversions," *Fuzzy Sets Syst.*, vol. 473, no. October, p. 108740, 2023.

[15] L. Ramadhani and I. Purnamasari, "Penerapan Metode Complete Linkage dan Metode Hierarchical Clustering Multiscale Bootstrap (Studi Kasus : Kemiskinan Di Kalimantan Timur Tahun 2016) Application of Complete Linkage Method and Hierarchical Clustering Multiscale Bootstrap Method," vol. 9, no. 2016, pp. 1–10, 2018.

[16] A. V. Latuhimallo, W. M. Talakua, and Z. Leleury A, "Analisis Clustering Untuk Pengelompokan Kabupaten/Kota Berdasarkan Indikator Kesejahteraan Rakyat Di Wilayah Provinsi Maluku," *Param. J. Mat. Stat. dan Ter.*, vol. 01, no. 01, pp. 1–12, 2023.

[17] E. Widodo, P. Ermayani, L. N. Laila, and A. T. Madani, "Pengelompokkan Provinsi di Indonesia Berdasarkan Tingkat Kemiskinan Menggunakan Analisis Hierarchical Agglomerative Clustering," vol. 2, pp. 557–566, 2020.

[18] Suyanto, Syarippudini, and Wasono, "Analisis Cluster Single Linkage Berdasarkan Potensi Desa Di Kabupaten Kutai Kartanegara Tahun 2019," *J. Eksponensial*, vol. 12, pp. 59–64, 2021.

[19] N. S. Laamena and T. Talib, "PENERAPAN ANALISIS KLASTER HIERARKI UNTUK PENGELOMPOKAN KABUPATEN/KOTA DI PROVINSI MALUKU BERDASARKAN STATUS PENDIDIKAN," *Sci. Map J.*, vol. 5, pp. 10–18, 2023.

[20] M. M. Ramly and B. Poerwanto, "Analisis Hierarchical Clustering Multiscale Bootstrap (Kasus : Indikator Kemiskinan Di Provinsi Sulawesi Selatan Tahun 2020)," vol. 4, no. 3, pp. 142–152, 2022.