



Available online at <http://scik.org>

Commun. Math. Biol. Neurosci. 2023, 2023:82

<https://doi.org/10.28919/cmbn/8088>

ISSN: 2052-2541

MAPPING INDONESIAN POTENTIAL FISHING ZONE USING HIERARCHICAL AND NON-HIERARCHICAL CLUSTERING

RESA SEPTIANI PONTOH*, SOFFY MULYANI, SALMA ZHAHIRA, OCTAVIA AULIA WIRATAMA,
MOHAMAD NAUFAL FARRAS, RESTU ARISANTI

Department of Statistics, Padjadjaran University, Bandung 45363, Indonesia

Copyright © 2023 the author(s). This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract: Indonesia, a maritime nation whose ocean area exceeds its land area, has an abundance of ocean-based natural resources, such as fish, seaweed, coral reefs, and other marine organisms. The fisheries industry is one of the potential sources of extraordinary marine resources for the Indonesian economy. The annual increase or decrease in fish production in Indonesia can be attributed to several factors, including natural influences such as climate and ocean waves, inadequate management of marine resources, unequal distribution of facilities to support increased fish production in Indonesia, and the characteristics of areas that have a significant impact on the resulting fish production. Consequently, the objective of this research is to classify provinces in Indonesia using clustering analysis so that government policy programs can be more focused and directed according to the characteristics of the clusters formed. The application of cluster analysis was based on the development of fish production data for each province in Indonesia from 2017 to 2019 obtained from the website of the Central Statistics Agency (BPS). Clustering analysis using hierarchical and non-hierarchical methods produces a dendrogram using the average linkage DTW hierarchical method, indicating the formation of two optimal clusters. Non-hierarchical clustering with two clusters produces the

*Corresponding author

E-mail address: resa.septiani@unpad.ac.id

Received July 03, 2023

same distribution of province members as group members in hierarchical clustering. The average silhouette coefficient value for the two clusters formed was 0.64, indicating that the clustering is classified as Good Classification.

Keywords: fish production; hierarchical clustering; non-hierarchical clustering; Indonesia.

2020 AMS Subject Classification: 92B10.

1. INTRODUCTION

Indonesia, which is a maritime country with an ocean area that exceeds the land area, certainly has a lot of natural potential that comes from the oceans such as fish, seaweed, coral reefs, and other marine biota. One of the potential sources of extraordinary marine resources to support the Indonesian economy is in the fisheries sector. This is because Indonesia is the largest archipelagic country in the world, with 17,508 islands and 81,000 km of coastline, and around 70% of its territory is sea. With marine waters covering a total area of 5.8 million km², Indonesia has abundant potential for biological and non-biological resources [1].



Figure 1. Territory of the Republic of Indonesia

Source: Indonesia.go.id

Fisheries are all activities related to the management and utilization of fish resources and their environment, starting from pre-production, production, and processing to marketing, which are carried out in a fishery business system [2]. Fish production in Indonesia can increase or decrease

every year. This can be influenced by several factors, namely natural influences such as climate, sea waves, inadequate management of marine resources, and the lack of an equal distribution of facilities to support increased fish production in Indonesia [3]. Figure 2 shows a graph showing the results of capture fish production in Indonesia from year to year.

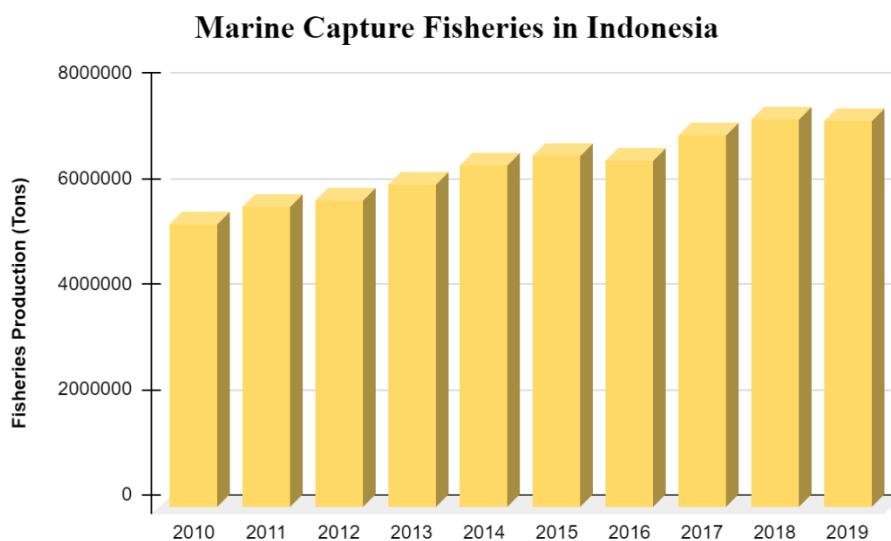


Figure 2. Graph of Captured Fish Production in Indonesia

Source: Statistics of the Ministry of Maritime Affairs and Fisheries of the Republic of Indonesia

Seeing how important the development of marine aquaculture in Indonesia is, data management is also very necessary. Therefore, against the background of these problems, the author will conduct a cluster analysis by clustering provinces in Indonesia that have the same characteristics based on the amount of fish production to determine the provinces with the best marine aquaculture opportunities, data management is based on annual production data (tons) in each province in Indonesia in 2017-2019. Data mining is a method used in large-scale information processing. Therefore, information mining has a very significant role in several areas of life, including industry, finance, weather, science, and technology [4].

Cluster analysis is a multiple variable analysis technique used to identify and classify individuals or objects that have similarities into certain groups or clusters. The characteristics of objects in a certain group will have a high level of similarity, while the characteristics of objects

in one group to another group will have a low level of similarity [5]. The K-Means algorithm will test each component in the data population and mark this component as one of the predetermined cluster centers, depending on the minimum distance between the components and each cluster center [6]. In addition, the location of the cluster center will be recalculated until all data components are classified into each cluster, and a new cluster will be formed at the end which can then be used as a 2-dimensional or 3-dimensional mapping [7][8].

Mapping provinces in Indonesia based on the amount of fish production is important to assist related parties, namely the government, in making appropriate plans and policies. Besides that, the results of this mapping help business people see a map of the potential of the marine fisheries sector. This mapping effort is expected to be able to combine provinces that have similar indicators. Thus, mapping the fisheries sector will help related parties make policies that are more focused because handling will be adjusted to the characteristics of each province and help make the right strategy for the development of processed marine fishery products and create a good ecosystem for population growth. fish production in every province in Indonesia, and can meet the needs of facilities that can support increased fish production in every province in Indonesia.

The objective of this study is to cluster fish production results (tons) per year in all provinces in Indonesia for the 2017-2019 period using the K-Means algorithm so that groups that are based on the best marine aquaculture opportunities are formed in terms of the amount of fish production by year. Another objective of this clustering is to find out which clusters contain provinces with the highest fish production. The benefits of this research include providing information for fisheries management in every province in Indonesia, as reference material and guidance for planning fishing operations for the fisheries sector, and also meeting the needs of facilities in the fisheries sector to increase the amount of fish production in the following year. This benefit is especially important for the Indonesian government and input material for universities throughout Indonesia, both public and private universities that have Fisheries study programs regarding clustering based on the amount of fish production in each province in Indonesia using the K-Means algorithm.

2. RESEARCH METHODOLOGY

2.1 DATA COLLECTION

The data used in this study is secondary data. Secondary data is data reported by an agency, that the agency does not collect but obtains from other parties [9]. Secondary data in this study was used from the Central Statistics Agency (BPS) website, namely capture fisheries production data at sea by main commodity (tons) with observation units in 34 provinces in Indonesia. The period for fish production was taken from 2017-2019 [10].

2.2 ANALYSIS METHODS

Data on Sea Capture Fisheries Production according to main commodities (tons) in Indonesia in 2017-2019 will be analyzed using the Non-Hierarchical and Hierarchical Clustering methods with the help of Software R. The package libraries used are “gridExtra”, “factoextra”, and “dendextend” [11].

The Non-Hierarchical Clustering method is used to determine the number of clusters to be formed. After that, the clustering process is carried out without regard to hierarchy or also known as the K-Means Cluster. Meanwhile, the Hierarchy method is carried out by first clustering two or more objects that have the closest similarity [12]. There is a clear level (hierarchy) between objects, from the most similar to the least similar. From the results of this clustering, it can only be known how many clusters are formed. The hierarchical graph that is formed is also called a dendrogram. The formation of the number of clusters can be seen only through the dendrogram, according to the subjectivity of the researcher [13]. The steps of data analysis carried out in this study are as follows:

2.2.1 Do data exploration.

Data exploration, commonly known as data exploratory analysis, is a process of analysis and initial understanding of the data to identify patterns and relationships that may occur in the data. The main goal of data exploration is to gain a better understanding of the existing data before carrying out further analysis and building predictive models [14].

2.2.2 Clustering the time series group analysis method using Dynamic Time Warping

(DTW)

Dynamic time warping distance (DTW) is an algorithm that compares two data series and performs calculations to find the optimum path between the two data series [15][16]. The distance DTW is a generalization of the classical algorithm that compares a sequence of discrete values with a sequence of continuous values. If there are two time series, namely Q and R , where $Q = q_1, q_2, \dots, q_i, \dots, q_n$ and $R = r_1, r_2, \dots, r_j, \dots, r_m$ so the use of the DTW distance aligns the data in the two-time series so that the difference in the distance between the two is minimal [17]. An illustration of the application of the DTW distance is shown in Figure 3.

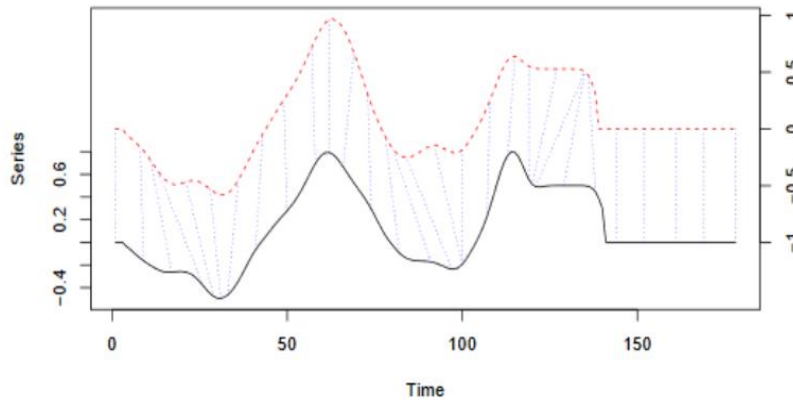


Figure 3. Illustration of Dynamic Time Warping (DTW) distance measurement
Source: Sarda-Espinosa, 2017

The application of the DTW distance measure uses computation through an $n \times m$ matrix where elements (i,j) from the matrix there is a distance (q_i, r_j) . Euclidean distance is used to calculate warping paths with $W = w_1, w_2, \dots, w_k, \dots, w_K$ where $\max(m, n) \leq K \leq m + n - 1$ is a set of matrix elements that satisfy three constraints, namely boundary conditions, continuity, and monotonicity. Boundary condition constraints require warping paths to start and finish computations diagonally opposite cell corners of the matrix, with $W_1 = (1, 1)$ and $W_K = (m, n)$. The formula for a warping path that has a minimum distance between two time series is as follows:

$$d_{DTW} = \min \frac{\sum_{k=1}^K w_k}{K}$$

with,

d_{DTW} : Dynamic Time Warping (DTW) Distance

w_k : K-th warping path

The calculation of the distance to the DTW is carried out using 3 linkage methods, namely "single linkage", "complete linkage", and "average linkage"[18]. Illustration of the three linkages as shown in Figure 4.

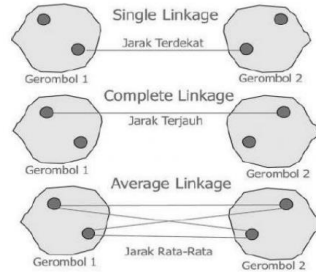


Figure 4. Illustration of Linkage

2.2.3 Perform Cophenetic correlation calculations.

The Cophenetic correlation coefficient is the correlation coefficient between the original elements of the dissimilarity matrix (Euclidean distance matrix) and the elements produced by the dendrogram (Cophenetic matrix based on distance measures and the connectedness method used) [19][20]. The formula for calculating the Cophenetic correlation coefficient is as follows:

$$r_{Coph} = \frac{\sum_{i < k} (d_{ik} - \bar{d})(d_{Cik} - \bar{d}_C)}{\sqrt{[\sum_{i < k} (d_{ik} - \bar{d})^2][\sum_{i < k} (d_{Cik} - \bar{d}_C)^2]}}$$

with,

r_{Coph} : Cophenetic correlation coefficient

d_{ik} : Distance of the i-th and k-th euclidean objects

\bar{d} : Average d_{ik}

d_{Cik} : The distance of the i-th and k-th cophenetic objects

\bar{d}_C : Average d_{Cik}

The value of r_{Coph} ranges between -1 and 1. The closer the value is to 1, the better the solution resulting from the clustering process. The distance measure of DTW with the connectedness method that produces the largest cophenetic value is the best solution for the hierarchical clustering method.

2.2.4 Perform non-hierarchical clustering using the k-means algorithm.

Non-hierarchical clustering using the k-means algorithm is a method commonly used in data analysis to group objects into adjacent groups based on feature similarities [21]. K-means is a non-

hierarchical clustering method that divides data into one or more groups [22]. This method divide data into several groups so that data with characteristics are grouped into one group. Data with properties are placed in the following groupings.

2.2.5 Perform silhouette coefficient calculations to see the accuracy of the k-means clustering.

The following is the silhouette coefficient calculation formula [23]:

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

Where,

$a(i)$: The average distance between members in the cluster

$b(i)$: The smallest distance between cluster members and the nearest neighbor

Following are the criteria for the accuracy and quality of the clustering results based on the silhouette coefficient values [23][24]:

Table 1. Category Clustering based on Silhouette Coefficient

Coeffisient Silhouette Value	Clustering Quality Category
0.71 – 1.00	Strong Classification
0.51 – 0.70	Good Classification
0.26 – 0.50	Weak Classification
0.00 – 0.25	Bad Classification

Source: Kauffman dan Rousseuw (1990)

3. MAIN RESULTS

3.1 DATA DESCRIPTION

Before carrying out the process of calculating data, the first step that must be taken is to collect data. The data used in this study is secondary data obtained from the Central Statistics Agency (BPS) website, namely regarding the amount of caught fish production in Indonesia in 2017-2019 with the object of observation being 34 provinces in Indonesia. The data on the amount of fishery production used is the total production of several types of marine fish such as skipjack, cob, tuna, shrimp, and others. Total production of capture fisheries at sea in Indonesia as a whole for 2017-

MAPPING INDONESIAN POTENTIAL FISHING ZONE

2019 is presented in Figure 5.

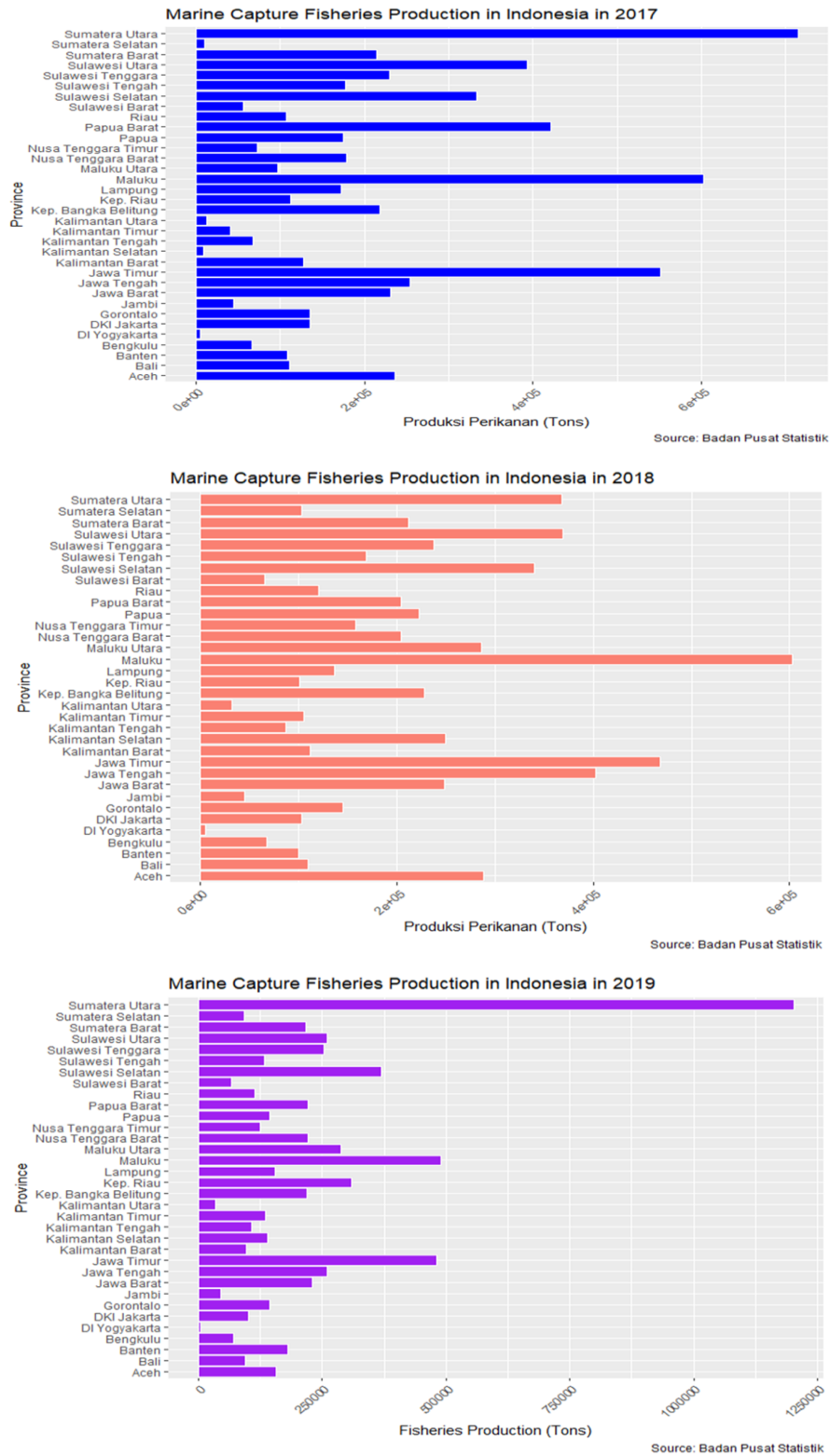


Figure 5. Plot of marine capture fisheries production (tons) by province in Indonesia in 2017-2019

3.2 Time Series Clustering

The time series cluster analysis in this study aims to classify provinces in Indonesia based on the amount of fish caught in the sea (tons). The first step taken in this analysis is to create a Dynamic Time Warping (DTW) distance measurement matrix, as shown in Table 2.

The shortest distance based on the Dynamic Time Warping (DTW) distance is between the Province of Kep. Bangka Belitung and West Sumatra Provinces with a distance value of 22191. Therefore, the first time classifying the application of DTW distances was carried out by clustering the Kep Provinces. Bangka Belitung and West Sumatra Province.

Table 2. Dynamic Time Warping (DTW) distance measurement matrix

PROVINSI	ACEH	SUMUT	SUMBAR	RIAU	...
ACEH	0				
SUMUT	1899260	0			
SUMBAR	196208	1948860	0		
...	0	
LAMPUNG	215933	2231815	269713	161198	
BABEL	174587	1911709	22191	519495	
...					

Source: Results of data processing

3.3 Hierarchical Clustering

The next step is to describe a hierarchical time series cluster analysis dendrogram with Dynamic Time Warping (DTW) distances and the application of various linkage methods ("single linkage", "complete linkage", and "average linkage"). The results of the delineation and clustering of the dendrogram are as follows.

MAPPING INDONESIAN POTENTIAL FISHING ZONE

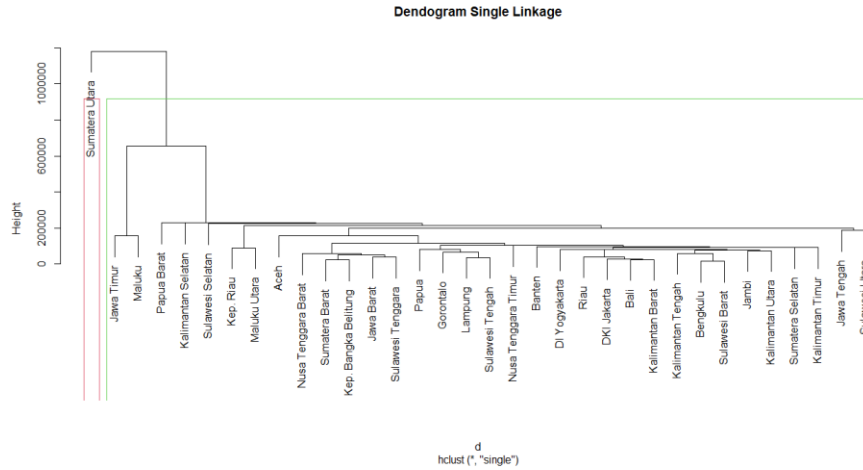


Figure 6. Dendrogram of cluster analysis using the single linkage method to measure the distance between DTWs

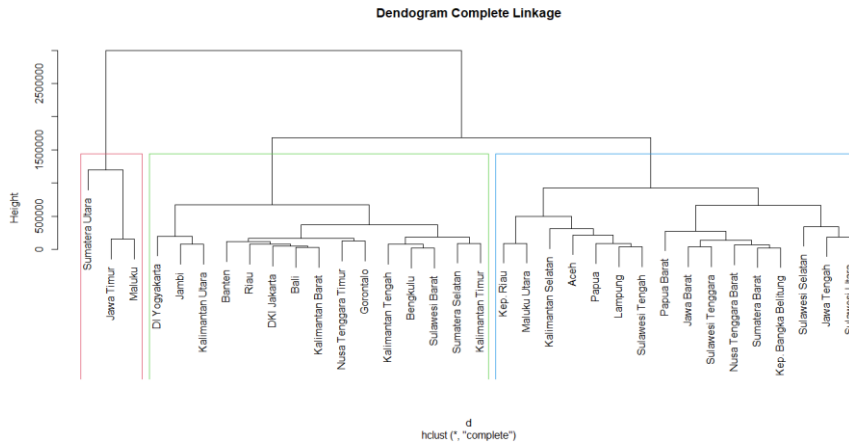


Figure 7. Cluster analysis dendrogram using the complete linkage method to measure the distance between DTWs

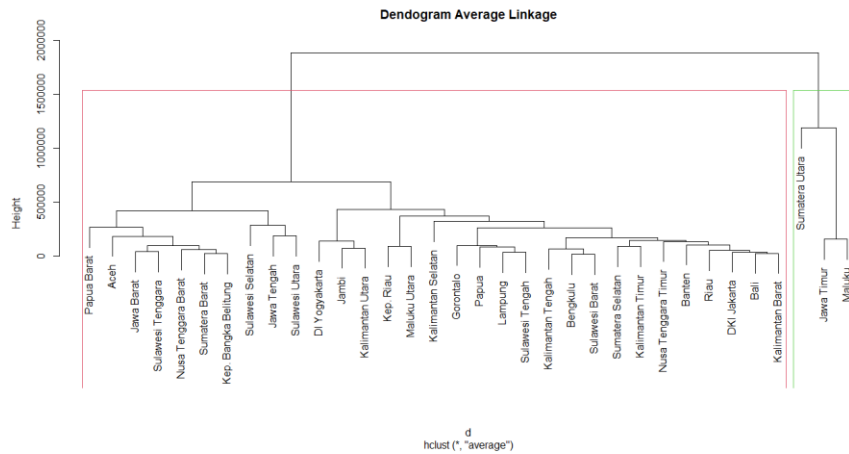


Figure 8. Dendrogram of cluster analysis with the average linkage method of measuring the distance between DTWs

Based on the cluster analysis dendogram clustering with the distance measurement of the DTW "single linkage" and "average linkage" linkage methods, the results of the clusters formed are as many as 2 clusters (Figure 6 and Figure 8). Meanwhile, the "complete linkage" method produces 3 clusters (Figure 7).

The next step in this study is to calculate the cophenetic correlation values as shown in Table 3.

Table 3. Cophenetic correlation values for the distance between DTWs based on the linkage

method		
No.	Linkage Methode	Cophenetic Correlation
1	Single Linkage	0.8377439
2	Complete Linkage	0.8123367
3	Average Linkage	0.8836320

From Table 3, the highest cophenetic correlation value (0.8836320) is obtained using the average linkage distance measure. Thus, based on the analysis of the clustering time series (clustering time series) hierarchical method, the optimal cluster formed is 2 clusters. The distribution of provinces in the 2 optimal clusters of the hierarchical method is:

1. Cluster 1 consists of 3 provinces, namely North Sumatra, East Java and Maluku.
2. Cluster 2 consists of 31 provinces, namely West Papua, Aceh, West Java, Southeast Sulawesi, West Nusa Tenggara, West Sumatra, Kep. Bangka Belitung, South Sulawesi, Central Java, North Sulawesi, DI Yogyakarta, Jambi, North Kalimantan, Kep. Riau, North Maluku, South Kalimantan, Gorontalo, Papua, Lampung, Central Sulawesi, Central Kalimantan, Bengkulu, West Sulawesi, South Sumatra, East Kalimantan, East Nusa Tenggara, Banten, Riau, DKI Jakarta, Bali and West Kalimantan.

MAPPING INDONESIAN POTENTIAL FISHING ZONE

The figure below (Figure 9) is the result of clustering if displayed in a mapping landscape in Indonesia.

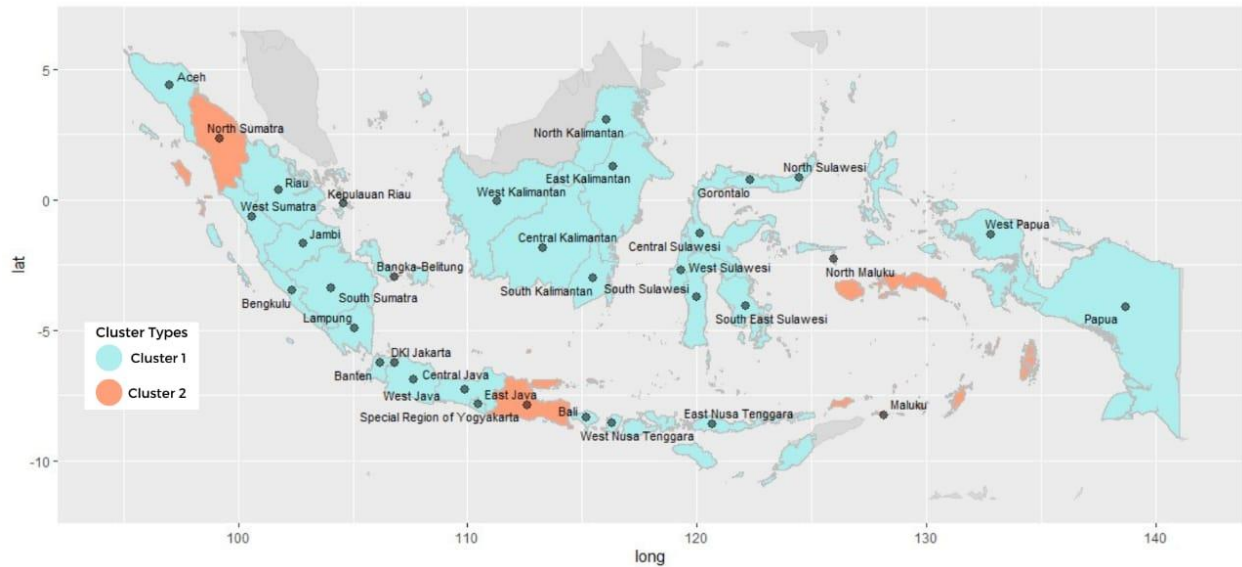


Figure 9. Cluster Landscaping Mapping is formed based on the results of fish production in the territory of Indonesia

3.4 Non-Hierarchical Clustering

In non-hierarchical clustering, one way to estimate the quality of the clusters formed by the k-means algorithm can be seen from the average silhouette value approach.

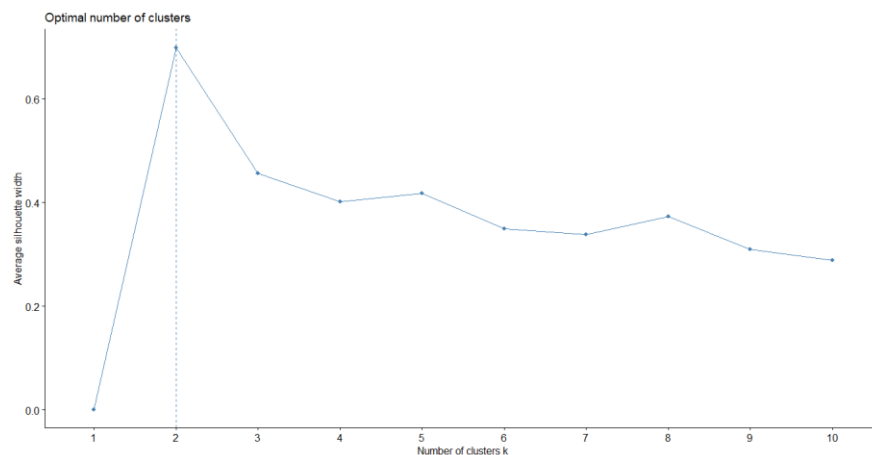


Figure 10. The optimal cluster generated by the k-means algorithm

Based on Figure 10, the number of optimal clusters formed is at $k=2$, while the number of

optimal clusters for the second is $k=3$. The results of this non-hierarchical cluster method were compared with the results of the hierarchical method, so what was used and further analyzed was the number of 3 clusters ($k=3$).

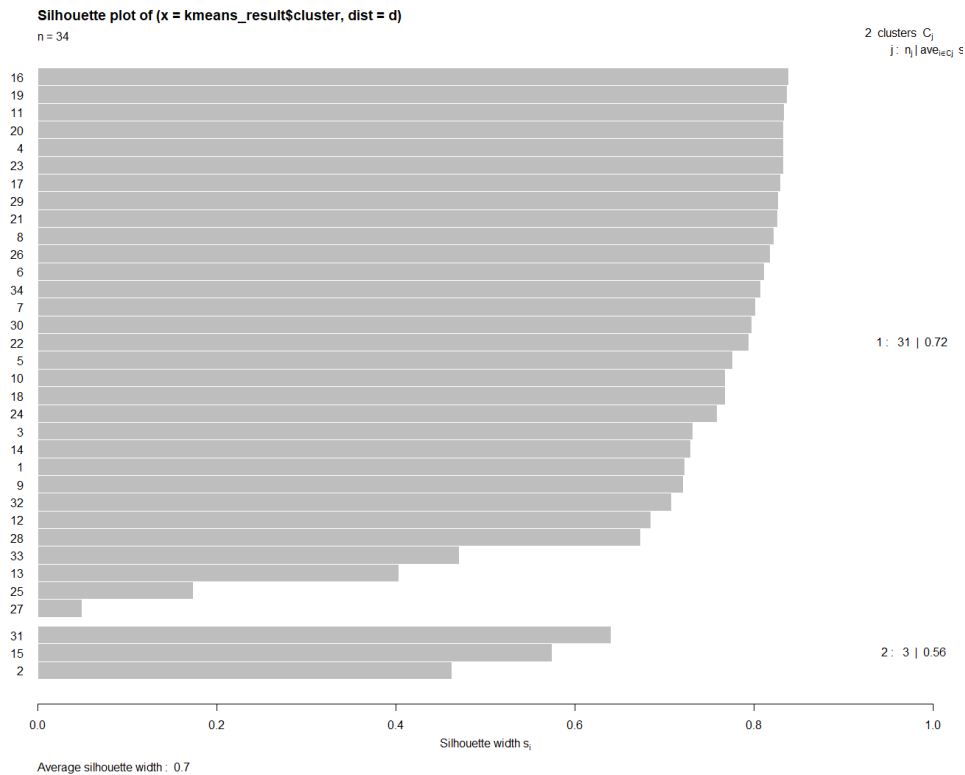


Figure 11. Silhouette plot of Indonesian fish production data for the 2017-2019 period

Calculations using the k-means algorithm and the optimal number of clusters of 2 produce the same cluster distribution as clustering using the hierarchical method. Clustering distribution and silhouette coefficient calculation results to see the accuracy of clustering as shown in Figure 8. Based on the clustering criteria with silhouette coefficient values, cluster 1 with a value of 0.72 is categorized in the Strong Classification and cluster 2 with a value of 0.56 is categorized in the Good Classification. On average from the 2 clusters formed, the resulting silhouette coefficient value is 0.64. This means that the clustering that has been done is categorized in the Good Classification.

4. DISCUSSION

The background for conducting this study is that Indonesia has extraordinary potential in marine resources such as fish production. However, it can be seen that fish production in Indonesia sometimes increases or decreases. This can be influenced by several factors, namely natural influences such as climate, sea waves, inadequate management of marine resources, and the lack of an equal distribution of facilities to support increased fish production in Indonesia. Therefore, this study was conducted with the main objective of knowing the clustering in each province in Indonesia, based on the results of captured fish production in 2017-2019. Besides that, it is to see how clusters are formed based on the amount of fish production in each province. This research can be said to research with a breakthrough because previous research used data on the value of rice production for clustering provinces in Indonesia with the result that the clusters formed were 3 clusters [25]. In this study, using hierarchical and non-hierarchical cluster analysis, the results showed that there were 2 clusters formed where cluster 1 showed 3 provinces that had a high amount of fish production, while cluster 2 showed 31 provinces that had a lower amount of fish production than cluster 1. This indicates that Cluster 1 can be used as an example for the 31 provinces in Cluster 2, including finding out how to increase fish production in Cluster 1 or exchanging information about techniques or tools used to increase fishing in the sea. With the formation of these 2 clusters, it is hoped that the government or parties related to fishing activities will pay more attention to the 31 provinces in Cluster 2 with a lower number of fish production than the 3 provinces in Cluster 1. Actions that can be taken are to provide and distribute facilities that support fishing activities, provide insight or direction to the 31 provinces in Cluster 2 based on what is implemented in the 3 provinces in Cluster 1, and also make more targeted policies to increase fish production in the 31 provinces in Cluster 2.

5. CONCLUSIONS

Analysis of time series clustering, also known as Clustering Time Series, in this study aims to cluster fish production (tons) per year in all provinces in Indonesia for the 2017-2019 period using

the K-Means algorithm so that clusters are formed according to The best marine aquaculture opportunities are seen in the amount of fish production per year. The analytical method used for data on marine capture fisheries production by main commodity (tons) in Indonesia in 2017-2019 uses the Non-Hierarchical and Hierarchical Clustering methods. Based on the results of the analysis, the highest cophenetic correlation value was obtained, namely on the Dynamic Time Warping (DTW) distance measure of average linkage, with a value of 0.8836320. Dendrogram using the DTW average linkage hierarchical method shows the formation of two optimal clusters. Non-hierarchical clustering with two clusters produces the same distribution of province members as group members in hierarchical clustering. Cluster 1 (3 provinces) is an area with a high fish production category. While cluster 2 (31 provinces) is an area with a low fish production category. The results of the two clusters formed on average obtained a silhouette coefficient value of 0.64, which means that the clustering is categorized as Good Classification.

FUNDING

This research is supported by the Department of Statistics, Padjadjaran University and the Rector of Padjadjaran University.

CONFLICT OF INTERESTS

The authors declare that there is no conflict of interests.

REFERENCES

- [1] Pemerintah Indonesia, Undang-Undang (UU) Nomor 31 Tahun 2004 tentang Perikanan, LL Sekretariat Negara No. 5587, Jakarta, (2004).
- [2] Nurlina, Analisis Keterkaitan Sub Sektor Perikanan Dengan Sektor Lain Pada Perekonomian Di Provinsi Aceh, *Jurnal Samudra Ekonomika*, 2 (2018), 20-29.
- [3] P.K. Thornton, P.J. Ericksen, M. Herrero, et al. Climate variability and vulnerability to climate change: a review, *Glob. Change. Biol.* 20 (2014), 3313-3328. <https://doi.org/10.1111/gcb.12581>.

MAPPING INDONESIAN POTENTIAL FISHING ZONE

- [4] R.W. Sari, D. Hartama, Data Mining: Algoritma K-Means Pada Pengelompokan Wisata Asing ke Indonesia Menurut Provinsi, Seminar Nasional Sains dan Teknologi Informasi. 1 (2018), 322-326.
- [5] A.A. Mattjik, I.M. Sumertajaya, Sidik Peubah Ganda Dengan Menggunakan SAS, IPB Press, Bogor, (2011).
- [6] S.-H. Jung, H. Lee, J.H. Huh, A novel model on reinforce k-means using location division model and outlier of initial value for lowering data cost, *Entropy*. 22 (2020), 902. <https://doi.org/10.3390/e22080902>.
- [7] Y. Agusta, K-Means-Penerapan, Permasalahan dan Metode Terkait, *J. Sist. Inform.* 3 (2007), 47-60.
- [8] J.O. Ong, Implementasi algoritma K-means clustering untuk menentukan strategi marketing president university, *Jurnal Ilmiah Teknik Industri*. 12 (2013), 10-20.
- [9] Sugiyono, Metode Penelitian Kuantitatif, Alfabeta, Bandung, (2018).
- [10] BPS (Badan Pusat Statistik), Jumlah Produksi Ikan (ton) menurut provinsi di Indonesia, 2017-2019.
- [11] A. Sardá-Espinosa, Comparing time-series clustering algorithms in R using the dtwclust package, *R package vignette*, 12, (2017).
- [12] G. Gan, C. Ma, J. Wu, Data clustering: theory, algorithms, and applications, vol. 20. SIAM, Philadelphia, (2007).
- [13] D. Rachmatin, Application of Agglomerative Methods in Cluster Analysis to Air Pollution Level Data, *Jurnal Ilmiah Program Studi Matematika STKIP Siliwangi Bandung*, 3 (2014), 133-149.
- [14] A. Setiawan, Analisis Data Eksploratif, (2022).
- [15] J. Stübinger, D. Walter, Using multi-dimensional dynamic time warping to identify time-varying lead-lag relationships, *Sensors*. 22 (2022), 6884. <https://doi.org/10.3390/s22186884>.
- [16] T. Giorgino, Computing and visualizing dynamic time warping alignments in R: The dtw Package, *J. Stat. Soft.* 31 (2009), 1-24. <https://doi.org/10.18637/jss.v031.i07>.
- [17] H. Sakoe, S. Chiba, Dynamic programming algorithm optimization for spoken word recognition, *IEEE Trans. Acoust., Speech, Signal Process.* 26 (1978), 43-49. <https://doi.org/10.1109/tassp.1978.1163055>.
- [18] J. Bu, W. Liu, Z. Pan, et al. Comparative study of hydrochemical classification based on different hierarchical cluster analysis methods, *Int. J. Environ. Res. Public Health*. 17 (2020), 9515. <https://doi.org/10.3390/ijerph17249515>.
- [19] I. Yahya, G.N.A. Wibawa, L. Laome, Penggunaan Korelasi Cophenetic Untuk Pemilihan Metode Cluster Berhierarki Pada Mengelompokkan Kabupaten/Kota Berdasarkan Jenis Penyakit Di Provinsi Sulawesi Tenggara

Tahun 2020, Seminar Nasional Sains dan Terapan, 6 (2022), 1-16.

- [20] G. Haumahu, Y.W. Nanlohy, Penerapan Analisis Klaster Hierarki Untuk Pengelompokkan Kabupaten/Kota Di Provinsi Maluku Berdasarkan Konsumsi Kalori Penduduk, *VARIANCE: J. Stat. Appl.* 2 (2020), 75-79.
- [21] A.M. Ikotun, A.E. Ezugwu, Boosting k-means clustering with symbiotic organisms search for automatic clustering problems, *PLoS ONE*. 17 (2022), e0272861. <https://doi.org/10.1371/journal.pone.0272861>.
- [22] Y. Agusta, Penerapan, permasalahan dan metode terkait K-Means, *J. Sist. Inform.* 3 (2007), 47-60.
- [23] L. Kaufman, P.J. Rousseuw, Finding group in data: an introduction to cluster analysis, John Wiley & Sons, Hoboken, (1990).
- [24] R.O. Sinnott, H. Duan, Y. Sun, A case study in big data analytics, in: *Big Data*, Elsevier, 2016: pp. 357-388. <https://doi.org/10.1016/B978-0-12-805394-2.00015-5>.
- [25] A.D. Munthe, Penerapan clustering time series untuk menggerombolkan provinsi di Indonesia berdasarkan nilai produksi padi, *Jurnal Litbang Sukowati: Media Penelitian dan Pengembangan*, 2 (2019), 11. <https://doi.org/10.32630/sukowati.v2i2.61>.