International Journal of Electrical and Computer Engineering (IJECE)
Vol. 15, No. 1, February 2025, pp. 46~55
ISSN: 2088-8708, DOI: 10.11591/ijece.v15i1.pp46-55

Determination of biomass energy potential based on regional characteristics using adaptive clustering method

Ginas Alvianingsih[1,3], Haslenda Hashim[1], Jasrul Jamani Jamian[2], Adri Senen[2,3]
[1]Department of Chemical Engineering, Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
[2]Department of Electrical Power Engineering, Faculty of Electrical Engineering, Universiti Teknologi Malaysia, Johor Bahru, Malaysia
[3]Department of Electrical Engineering, Institut Teknologi PLN, Jakarta, Indonesia

Article Info

Article history:

Received Mar 6, 2024
Revised Sep 5, 2024
Accepted Oct 1, 2024

Keywords:

Biomass potential
Clustering
Energy
Fuzzy C-means
Regional characteristic

ABSTRACT

Determining the energy potential of biomass is the first step in selecting the most suitable and efficient energy conversion technology based on regional characteristics. The approach to estimating and determining biomass potential generally uses geospatial technology related to collecting and processing data about mapping an area. Unfortunately, this method is inadequate for simulating the interaction between variables, nor can it provide accurate predictions for the biomass supply chain. As a result, the results obtained from this method tend to be biased and macro, particularly in regions experiencing rapid land-use development. In this paper, the author has developed a clustering methodology with a fuzzy c-means (FCM) algorithm to determine biomass energy potential based on regional characteristics to produce data clusters with high accuracy. Grouping the characteristics of clustering-based areas involves grouping physical or abstract objects into classes or similar objects.

Corresponding Author:

Ginas Alvianingsih
Faculty of Chemical and Energy Engineering, Universiti Teknologi Malaysia
Johor Bahru, Malaysia
Email: alvianingsih@graduate.utm.my

1. INTRODUCTION

Fuel delivery becomes a problem when providing electricity in remote locations. The electricity system on islands or areas with relatively low population density is usually an isolated system supplied by diesel power plants that use marine fuel oil (MFO) and high-speed diesel (HSD) [1]. Diesel fuel is delivered to an isolated diesel power plant by tanker truck from the city center. This case results in fuel shortages and high fuel prices, so some rural areas do not receive electricity throughout the day to minimize the cost [2], [3]. On the other hand, diesel power plants are a fossil energy source that contributes to ozone depletion, acid rain, and global warming [4]. Therefore, diversification of energy sources is very important, especially for renewable resources.

Biomass energy can reduce dependence on limited amounts of fossil fuels and contribute to energy supply security, especially in remote and rural areas. Biomass energy can be produced responsibly from forestry, plantation, and agricultural waste. The use of biomass energy can help sustainable development for communities in remote locations by ensuring a stable electricity supply because remote locations often have significant biomass potential [5]. Different types of biomass potential have different physical structures, chemical compositions, and calorific values that will influence the energy conversion process and efficiency [6].

Direct combustion, thermochemical processes (pyrolysis [7], gasification, and direct liquefaction), and biochemical processes (anaerobic digestion and alcoholic fermentation) are the three main ways that biomass is used. All of these methods transform the biomass into useful products [8]. Because they are sustainable and potentially lower greenhouse gas emissions, the thermal conversion of biomass materials into dense fuels with a greater calorific value is attracting attention from around the world [9]. Some researchers have further examined plasma gasification technology [10], [11]. Many biomass power plants have not been utilized optimally [12] because the selection of biomass conversion technology did not consider the specific region's characteristics and the commodities produced. Therefore, mapping the region's features based on biomass energy potential is critical to implementing the biomass power plant and choosing the best conversion technology.

Previous studies have created the approach to estimating and determining biomass potential, which generally uses geospatial technology for collecting and processing data about mapping an area [13]. The biomass potential has been estimated by studying a few geospatial technologies [14]–[16]. Crop production statistics calculate gross biomass potential (GBP) in geographic information systems (GIS) approaches. Spatiotemporal satellite data produced geographic maps of crop residue biomass potential under current conditions. Regretfully, neither the biomass supply chain nor the interaction between variables can be accurately predicted using this method. Because of this, the outcomes of this approach are frequently skewed and broad, especially in areas where land use is developing quickly. It also takes a lot of work to identify and map areas with varied features, adding to the study's complexity and length and making it more challenging to estimate the biomass potential.

Several parameters, including electrical and non-electrical factors such as biomass calorific value, amount of biomass potential, demographic factors, and characteristics of electrical consumers, should be considered for high-quality biomass mapping and conversion technology determination. The author of this journal offers a novel method that can accommodate the best mapping of biomass potential and conversion technology determination using an area clusterization. The methodology entails identifying the area's demands, features, and patterns based on biomass potential by applying unsupervised learning techniques named clustering. Clustering is grouping related items or regions according to shared characteristics. Since the adaptive clustering algorithm doesn't require prior knowledge or labeling data about the target group, it's an excellent option for mapping broad areas with various parameters. Models are fed unlabeled data to identify underlying patterns and structures and generate precise data clusters [17]. The cluster technique reduces information loss while standardizing and simplifying data representation.

The author of this work has created a clustering methodology using the fuzzy C-means (FCM) algorithm to identify the energy potential of biomass based on geographical features and generate data clusters with a high degree of accuracy. Classifying real or abstract items or comparable objects allows grouping properties of clustering-based areas [18]. As opposed to complex clustering techniques like K-means, each piece of data is not grouped as a result. In practical scenarios, FCM performs better than other clustering techniques when classifying data. This technique may take pictures and evaluate biomass's potential by looking at a place's physical and population characteristics. Once the cluster area has been determined, a model can be created to identify the most effective technology for converting biomass energy into electrical energy. This approach will provide accurate mapping of the region's features and enhance the design of upcoming biomass power plants with higher levels of efficiency.

The structure of this paper is outlined as follows: section 2 presents the research flowchart and the fundamentals of the method that will be used to implement the idea and strategy of optimal biomass mapping. this section also elaborates on the variable that will be used as the input. Section 3 presents the data and results of our proposed method with validation to demonstrate its effectiveness. finally, we summarize our findings in section 4.

2. METHOD

This research uses micro-spatial analysis to acquire impartial results using the macro-mapping technique. Figure 1 shows the research flowchart. The methodology used in conducting this study can generally be divided into 5 parts: the preliminary work, area clusterization, cluster validation, cluster result mapping, and regional characteristic analysis. The preliminary work contains the literature study, variable decision, and data collection.

The potential biomass of 164 subdistricts in a province is calculated in this article. To choose the best conversion technology, several factors include electrical and non-electric attributes unique to each subdistrict. The variables considered for this investigation are listed in Table 1. The subdistrict's population, plantations, agriculture, and annual production are examples of non-electrical characteristics. The electrical variables are the installed capacity, electrical energy sold, the number of electrical consumers, and the capacity and running hours of the isolated diesel power plant. The crop types considered commodities in the province include palm oil, coconut, rubber, coffee, cocoa, areca nut, sago, and rice.

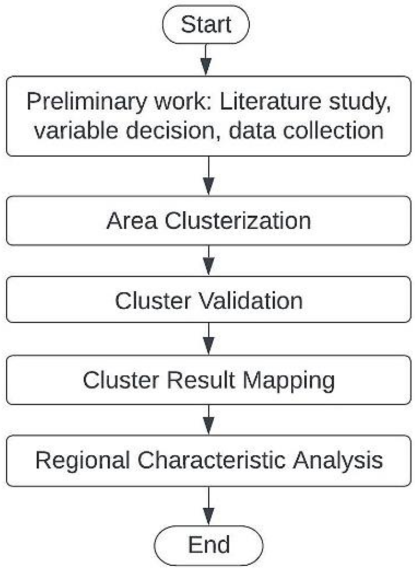


Figure 1. Research flowchart

Table 1. Electrical and non-electrical variables

Population and demography	Plantation and agriculture	Electricity
Wide of the area (km^2)	Wide of plantation (ha)	Installed capacity (VA)
Resident population (thousand people)	Annual production (tons)	Electrical energy sold (kWh)
		Number of electrical consumers

2.1. Adaptive clustering

Based on land use simulation modeling, micro-spatial analysis is an area analysis that starts by segmenting the region into smaller areas [19], [20]. There are two types of clustering, namely hard clustering and soft clustering. Hard clustering and soft clustering are the two different forms of clustering. Hard clustering entails counting the number of clusters sequentially, beginning with the smallest number. A cluster validation approach is then used to assess the clustering performance manually. Conversely, soft clustering, which is often referred to as adaptive clustering, is an algorithm that, by analyzing cluster outcomes, automatically generates data grouping with the maximum number of clusters. Soft clustering is a superior method since it more appropriately captures the condition of the data and the correlations between variables.

This study is new because it uses adaptive clustering techniques. This technique is chosen because, based on the analysis of cluster results, this algorithm can automatically build data groups with the ideal number of clusters. This is not the case with the traditional clustering technique, which clusters data sequentially or randomly, starting with the fewest possible clusters. Subsequently, the cluster validation approach is employed to analyze its performance personally. Conventional cluster techniques are more extended and less efficient. The use of unclear C-means often allows for soft clustering. FCM. The FCM algorithm is as [21]:

a. Data X needs to be entered as a matrix with a dimension of $n \times p$, where p is each data sample's characteristic and n is the total number of data samples.

   X_{kj} = data sample k^{-th} ($k = 1,2, \ldots, n$), data attribute j^{-th} ($j = 1,2,3, \ldots, m$).

b. Establish the goal function, initial iteration, maximum iteration, expected error, weighting power, and number of clusters.
c. As elements of the first partition matrix U, generate a random number ($\mu_{ik}, i = 1,2 \ldots, c;\ k = 1,2, \ldots, n$).

$$U_0 = \begin{bmatrix} \mu_{11}(x_1) & \mu_{12}(x_2) & \cdots & \mu_{1c}(x_c) \\ \vdots & \vdots & 0 & \vdots \\ \mu_{11}(x_1) & \mu_{12}(x_2) & \cdots & \mu_{1c}(x_c) \end{bmatrix} \tag{1}$$

In fuzzy clustering, the partition matrix needs to fulfill the subsequent requirements:

$$\mu_{ik} = [0,1];\ (1 \le i \le c;\ 1 \le k \le n)$$
$$\sum_{i=1}^{n} \mu_{ik} = 1; 1 \le i \le c$$
$$the\ \ 0 < \sum_{i=1}^{c} \mu_{ik} < c\ ; 1 \le k \le n \tag{2}$$

Determine and ascertain each column's (attribute's) quantity:

$$Q_j = \sum_{i=1}^{c} \mu_{ik} \tag{3}$$

where j = 1,2,3, …, m, then determine each data's degree of membership

$$\mu_{ik} = \frac{\mu_{ik}}{Q_j} \tag{4}$$

d. Determine the k-cluster's cluster centroid: V_{ij}, where $i = 1,2,3, \dots, c$ and $j = 1,2,3, \dots, m$

$$V_{ij} = \frac{\sum_{k=1}^{n}((\mu_{ik})^m * X_{kj})}{\sum_{k=1}^{n}(\mu_{ik})^m} \tag{5}$$

$$V = \begin{bmatrix} v_{11} & \cdots & v_{1m} \\ \vdots & \ddots & \vdots \\ v_{c1} & \cdots & v_{cm} \end{bmatrix}$$

e. Calculate the objective function, P_t

$$P_t = \sum_{k=1}^{n} \sum_{i=1}^{c} \left(\left[\sum_{j=1}^{m} \left(X_{kj} - V_{ij}\right)^2 \right] (\mu_{ik})^m \right) \tag{6}$$

f. Calculate the partition matrix change:

$$\mu_{ik} = \frac{\left[\sum_{j=1}^{p}\left(X_{kj} - V_{ij}\right)^2\right]^{\frac{-1}{p-1}}}{\sum_{i=1}^{c}\left[\sum_{j=1}^{p}\left(X_{kj} - V_{ij}\right)^2\right]^{\frac{-1}{p-1}}} \tag{7}$$

g. Verify the stop condition:
   - It is finished if ($|Pt - Pt - 1| < \xi$) or ($t < maximum\ iteration$)
   - If not, go back to step d with $t = t + 1$.

2.2. Index of cluster validation

The cluster validation index (CVI) is one tool for evaluating a cluster's quality and strength [22]. The optimum cluster outcome is influenced by the cluster method, data set properties, data size, number of clusters used, and data structure. As a result, CVI is crucial for assessing the cluster's quality. There are many techniques for CVI, including the Dunn index [23], [24], the Davies-Bouldin score, the Calinski-Harabasz score, and the Silhouette index. Two CVI are mentioned in a few literary works [25], [26].

In this case, the validity technique used is the Silhouette index, also known as the Kellogg-Silhouette index. Combining the cohesion method, which involves analyzing data within a cluster, with the separation method, which determines the relationship between the outcomes of those clusters, results in the silhouette coefficient method. The average between an object and every other object in the same cluster and objects in different clusters can be determined by applying the Silhouette algorithm [27]. Every group's silhouette is plotted to compare the quality of contributions based on the silhouettes' area (or light length). The silhouette coefficient can be calculated using the equation that follows:

$$s(i) = \frac{b(i) - a(i)}{\max(a(i), b(i))} \tag{8}$$

where
$s(i)$ = value of silhouette coefficient
(i) = average distance of i-data
$b(i)$ = average distance of i-data with all members
The silhouette width index is used to interpret silhouette value, as indicated in Table 2 [21].

Table 2. Silhouette width index [21]

Silhouette coefficient	Interpretation between objects and groups formed
0.7 < SC < 1	Strong structure
0.5 < SC < 0.7	Medium structure
0.25 < SC < 0.5	Weak structure
SC < 0.25	No relation

2.3. Clustering in biomass mapping

When a vast quantity and variety of geographic data is available, sampling a region with numerous frequently connected parameters is done to help. The application process is carried out after a cluster is constructed to connect data with different locations, after which it is connected and analyzed, and finally, the results are reported. Spatial data must be processed to produce spatially and temporally oriented data that can be used as a reference system.

Data properties for spatial data and plotting, such as topography and land zones on maps, can be customized with biomass mapping. Additionally, it is possible to mix spatial data with other spatial data to create a complementary data layer. Technical standards such as map projection systems and types of layers are required to ensure the reliability of geographic information systems. These speeds up the need for biomass mapping and biomass power plant planning.

2.4. Regional characterization

Regional characterization determination for each cluster will be done using correlation techniques. Each cluster member will have the exact regional characterization. The correlation coefficient between the width of the plantation and the total area can analyze the most biomass potential in a cluster. This approach will make the data easier to understand and provide users with more relevant information.

The correlation coefficient, denoted by r, measures the strength of the straight-line or linear relationship between two variables [28]. The following points are the accepted guidelines for interpreting the correlation coefficient:

- 0 indicates no linear relationship.
- +1 indicates a perfect positive linear relationship – as one variable increases in its values, the other increases through an exact linear rule.
- −1 indicates a perfect negative linear relationship – as one variable increases in its values, the other decreases through an exact linear rule.
- Values between 0 and 0.3 (0 and −0.3) indicate a weak positive (negative) linear relationship through a shaky linear rule.
- Values between 0.3 and 0.7 (0.3 and −0.7) indicate a moderate positive (negative) linear relationship through a fuzzy-firm linear rule.
- Values between 0.7 and 1.0 (−0.7 and −1.0) indicate a strong positive (negative) linear relationship through a firm linear rule.

3. RESULTS AND DISCUSSION

3.1. Data

Tables 3 to 5 show data that include the variables used. The 164 data used in this research show only five subdistricts. The data needs to be normalized before the clustering process can begin. Normalization is essential to arrange the qualities of different entities into a helpful connection structure (without redundancy or data repetition). This process will remove the majority of the ambiguity.

The data plot comparison in Figure 2 demonstrates the substantial impact of the normalization procedure on the data. Since the raw data are not dispersed equally at some point, it is clear from the plot of the actual data in Figure 2(a) that there is frequent uncertainty in the results. However, after normalizing the data displayed in Figure 2(b), the distribution of the normalized data forms a straightforward, non-redundant entity, guaranteeing that the data is of high quality for the subsequent stage.

Table 3. Population and demography variables

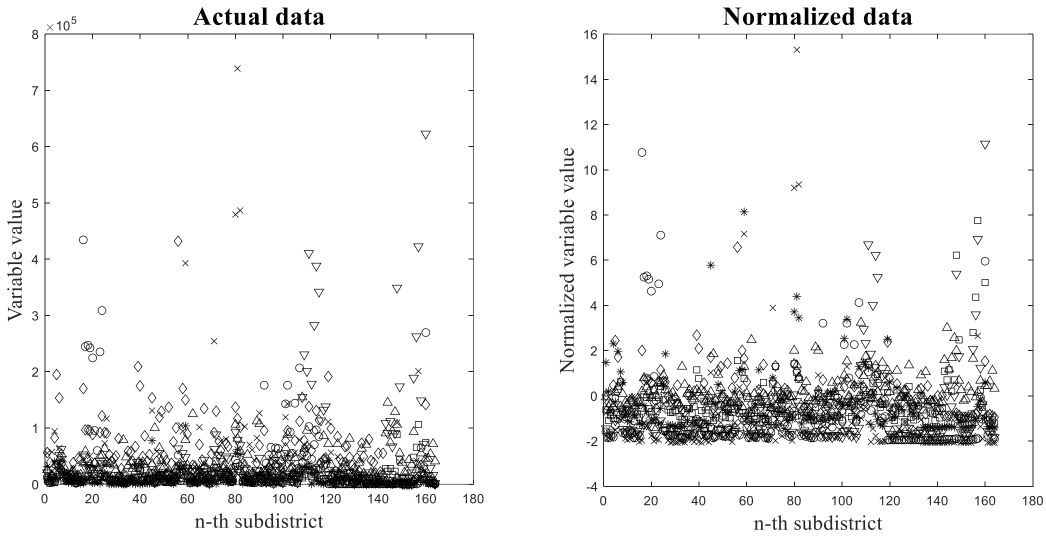
Sub-district (Grid)	Wide of the area (km^2)	Resident population (thousand people)
Kuantan Mudik	564.28	25.83
Hulu Kuantan	384.40	9.55
Gunung Toar	165,25	14.20
Pucuk Rantau	821.64	10.54
Singingi	1,953.66	35.43

Table 4. Plantation field and production parameters

Sub-district (Grid)	Palm Oil		Rice	
	Wide of plantation (ha)	Annual production (tons)	Wide of plantation (ha)	Annual production (tons)
Kuantan Mudik	31,083.36	55,446.26	355.70	1,574.39
Hulu Kuantan	17,291.65	10,615.75	242.31	1,072.51
Gunung Toar	8,029.49	57,274.63	104.17	461.06
Pucuk Rantau	39,853.90	94,233.82	517.92	2,292.45
Singingi	11,896.60	24,145.97	1,231.50	5,450.88

Table 5. Electrical parameters

Sub-district (Grid)	Electrical energy sold (kWh)	Number of electrical customers
Kuantan Mudik	20,700,968	7,215
Hulu Kuantan	4,875,170	1,699
Gunung Toar	12,075,069	4,208
Pucuk Rantau	7,050,774	2,457
Singingi	37,762,728	13,161



(a) (b)

Figure 2. Compares the data plots for (a) actual data and (b) normalized data

3.2. Clustering result and validation

To determine the ideal number of clusters, use adaptive clustering based on the silhouette algorithm and FCM techniques. This procedure requires input, the final iteration, the maximum error, the lowest and highest possible number of clusters, and the normalized data. The clustering algorithms organize the data into a minimal number of groups using adaptive clustering and continue until the maximum number is established in the first steps. Next, silhouette approaches are used to assess the outcome of each clustering algorithm's performance. The optimum silhouette index is selected to acquire the optimal quantity of clusters.

Figure 3 displays the outcomes of clustering results and validation. Each number of the cluster's silhouette index is displayed. From the 4 experiments conducted, it was found that the best Silhouette index was at a cluster number of 4. The value shows how the created clusters and sub-districts are interpreted. The subdistrict is separated from other groups produced by the excellent silhouette coefficient value near +1. The grid is at or reasonably near the decision border of the grid if it’s worth is 0. A grid should be in a separate cluster if its value is more harmful, which suggests that the grid overlaps. As shown in Figure 3, the silhouette coefficient has a mean value of 0.7. Table 2 indicates that the silhouette value is higher than 0.7, indicating a solid structure developed between the items. Table 6 shows the total number of members of each cluster. Cluster 1, with 136 sub-districts, or 82.9% of all the sub-districts examined, has the most significant number of members. Clusters 2, 3, and 4 consist of 4, 9, and 12 sub-districts, respectively.

To verify that four clusters are the ideal number to build, validation using the silhouette approach is done using (8); Figure 4 displays the outcomes of cluster validation. Each sub-district's silhouette index is displayed. The value shows how the created groupings and sub-districts are interpreted. The grid is separated from other groups produced by the excellent silhouette coefficient value near +1. The grid is at or reasonably near the decision border of the grid if it's worth is 0. Let's say that the value is almost +1. A grid should be in a separate cluster if its value is more harmful, which suggests that the grid overlaps. As shown in Figure 4, the silhouette coefficient has a mean value of 0.7. Table 2 indicates that the silhouette value is higher than 0.7, indicating a solid structure developed between the items.

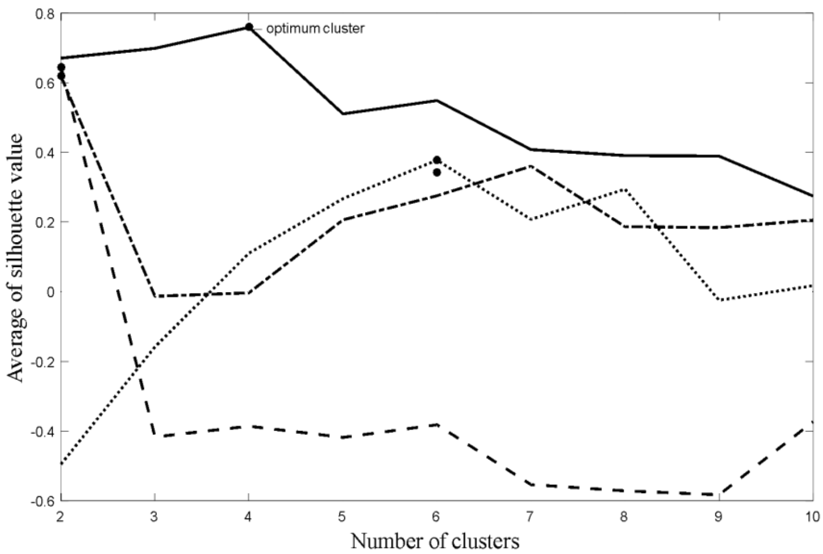


Figure 3. The optimal silhouette index derived using cluster validity procedures

Table 6. Total number of participants in each cluster

kth cluster	Number of members
1	136
2	4
3	9
4	12

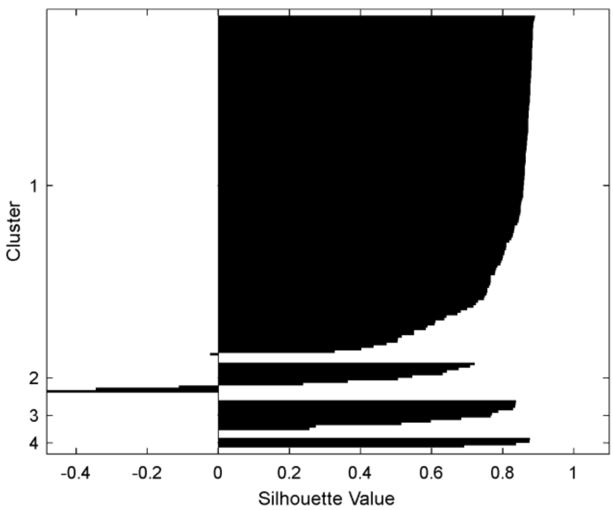


Figure 4. Silhouette value of optimum cluster

3.3. Clustering result visualization

To facilitate characteristic determination from each area, the distribution and positioning of each subdistrict's area inside each cluster are then displayed by plotting and mapping the clustering findings. Figure 5 displays the area's mapping results. Geographically speaking, clusters 2 and 3 are somewhat close together. However, cluster 4 is split between the province's west and south. In future research, this mapping can be the basis for determining the optimum location of a biomass power plant to be built by combining it with geographical conditions and land use maps as the additional clustering variables.

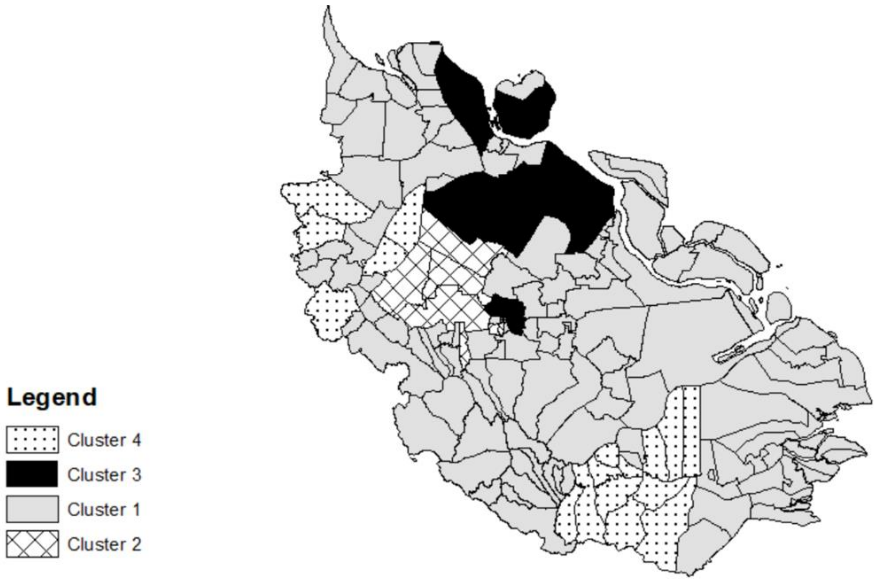


Figure 5. Mapping result of 4 clusters

3.4. Regional characterization analysis

The regional characterization is determined by grouping all areas using descriptive statistical techniques for each cluster. Correlation analyses are developed using SPSS software. Table 7 shows that cluster 1 has a good biomass potential for palm oil and rice with a moderate linear relationship. In cluster 2, the biomass potential is highest for palm oil, coconut, areca nut, and rice, with a strong linear relationship, and cocoa, with a moderate linear relationship. In cluster 3, the biomass potential is highest for palm oil, coconut, rubber, and cocoa, with a moderate linear relationship, and rice has a strong linear relationship. In cluster 4, the most biomass potential is coffee and rice with a moderate linear relationship. We can clearly conclude that all of the cluster has rice as a potential commodity.

From the analysis results, biomass conversion technology selection can be carried out more effectively based on the calorific value of the commodities in the cluster. Clusters that have biomass with relatively small calorific values must use conversion technology with higher efficiency values. In addition, production capacity can also be a constraint. Incineration technology can be implemented for biomass power plants with large capacity, while gasification technology can be implemented for biomass power plants with small capacity [29], [30]. From the mapping result, we can conclude that in this province most areas have the same characteristics, so the same type of technology can be applied as well.

As a result, area clusterization can be applied to map the biomass potential in a province by clustering each sub-district based on the biomass potential owned and other variables. The cluster results are then analyzed for calorific value and capacity so that the appropriate conversion technology can be determined. This will result in more effective biomass power plant planning in terms of energy conversion and also faster technology determination because there is no need for repeated analyses for several areas when several biomass power plants will be built. To determine the optimal biomass power plant location point, sub clustering can be carried out in an area with different variable constraints, for example, the cheapest cost of transporting biomass to the plant or lower electrical power losses.

Table 7. Correlation coefficients between the width of the plantation and the total area of each cluster

Cluster Number	Correlation between the width of the plantation and the total area							
	Palm Oil	Coconut	Rubber	Coffee	Cocoa	Areca Nurt	Sago	Rice
1	0.383	0.278	-0.014	0.02	0.02	0.146	0.13	0.311
2	0.747	0.057	0	0.623	0.623	0.733	0	0.737
3	0.514	0.599	0.264	0.303	0.303	0	0	0.781
4	-0.352	-0.322	0.4	0.111	0.111	0	0	0.448

4. CONCLUSION

From the research, we can conclude that adaptive clustering can be used to arrange the sub-districts into clusters that contain cluster members with similar profiles. From the characteristics of each cluster, we can select appropriate biomass conversion technology following the type of biomass potential to achieve optimal power generation output. Area clusterization methods and Silhouette validation will accommodate the best mapping of biomass potential. For each cluster, the results of the clustering process were thus examined using correlation analysis to determine its area characteristics. The correlation coefficient between the width of the plantation and the total area can analyze the significant biomass potential in a cluster. The same procedure can be implemented in other areas. For future research, conversion technology determination can be made based on the range of calorific values of the commodity in each cluster. Sub-clustering can also be used to determine the optimal point of biomass power plants with different constraint variables, such as geographical landmarks and land use.

REFERENCES

[1] G. Alvianingsih and I. Garniwa, "A design of palm oil and diesel oil fuel mixture heater system for small scale diesel power plant," in *2017 International Conference on High Voltage Engineering and Power Systems (ICHVEPS)*, Oct. 2017, pp. 156–164, doi: 10.1109/ICHVEPS.2017.8225934.

[2] G. Alvianingsih, V. Antono, and I. Garniwa, "Financial and technical forecast analysis of a hybrid biomass-diesel power plant: Case study in Tinggi Island, South Bangka," in *2021 3rd International Conference on High Voltage Engineering and Power Systems, ICHVEPS 2021*, Oct. 2021, vol. 11, pp. 305–308, doi: 10.1109/ICHVEPS53178.2021.9601102.

[3] I. Garniwa, G. Alvianingsih, and V. Antono, "Waste to energy implementation using gasification technology in Tinggi Island," *E3S Web of Conferences*, vol. 211, 2020, doi: 10.1051/e3sconf/202021103005.

[4] I. Dincer, "Renewable energy and sustainable development: a crucial review," *Renewable & sustainable energy reviews*, vol. 4, no. 2, pp. 157–175, Jun. 2000, doi: 10.1016/S1364-0321(99)00011-8.

[5] B. C. H. Simangunsong *et al.*, "Potential forest biomass resource as feedstock for bioenergy and its economic value in Indonesia," *Forest Policy and Economics*, vol. 81, pp. 10–17, Aug. 2017, doi: 10.1016/j.forpol.2017.03.022.

[6] L. Conrad and I. Prasetyaning, *Overview of the waste-to-energy potential for grid-connected electricity generation (solid biomass and biogas) in Indonesia*. LCORE-Indo, 2014.

[7] G. Maschio, C. Koufopanos, and A. Lucchesi, "Pyrolysis, a promising route for biomass utilization," *Bioresource Technology*, vol. 42, no. 3, pp. 219–231, Jan. 1992, doi: 10.1016/0960-8524(92)90025-S.

[8] H. Qian, W. Zhu, C. Liu, X. Lu, G. M. Kontogeorgis, and R. Gani, "Exergy efficiency based design and analysis of utilization pathways of biomasses," in *Computer Aided Chemical Engineering*, vol. 43, Elsevier, 2018, pp. 857–862.

[9] F. Ronsse, R. W. Nachenius, and W. Prins, "Carbonization of Biomass," in *Recent Advances in Thermo-Chemical Conversion of Biomass*, Elsevier, 2015, pp. 293–324.

[10] P. M. Lal and C. J. Singh, "Plasma gasification: A sustainable solution for the municipal solid waste management in the state of Madhya Pradesh, India," *International Journal of Enviromental Sciences*, vol. 3, no. 1, pp. 305–319, 2012.

[11] M. Pourali, "Application of plasma gasification technology in waste to energy challenges and opportunities," in *1st IEEE-PES/IAS Conference on Sustainable Alternative Energy, SAE 2009 - Proceedings*, Sep. 2009, pp. 1–6, doi: 10.1109/SAE.2009.5534883.

[12] S. Dani and A. Wibawa, "Challenges and policy for biomass energy in Indonesia," *International Journal of Business, Economics and Law*, vol. 15, no. 5, pp. 41–47, 2018.

[13] M. Mohd Noor, N., Mohd Nor, "Geospatial technology and innovation enablement for physical planning, space and process," *International Congress of Architecture and Planning (ICONARCH IV)*, pp. 77–86, 2020.

[14] A. Chakraborty *et al.*, "Developing a spatial information system of biomass potential from crop residues over India: a decision support for planning and establishment of biofuel/biomass power plant," *Renewable and Sustainable Energy Reviews*, vol. 165, Sep. 2022, doi: 10.1016/j.rser.2022.112575.

[15] N. A. Jasim, "Spatial analysis of forest biomass in Mississippi state using GIS," in *International Conference on Advances in Sustainable Engineering and Applications, ICASEA 2018 - Proceedings*, 2018, pp. 169–173, doi: 10.1109/ICASEA.2018.8370977.

[16] M. O. Ukoba, E. O. Diemuodeke, T. A. Briggs, M. Imran, K. Owebor, and C. O. Nwachukwu, "Geographic information systems (GIS) approach for assessing the biomass energy potential and identification of appropriate biomass conversion technologies in Nigeria," *Biomass and Bioenergy*, vol. 170, Mar. 2023, doi: 10.1016/j.biombioe.2023.106726.

[17] I. Tahyudin, G. Firmansyah, A. G. Ivansyah, W. Ma'arifah, and L. Lestari, "Comparison of K-means algorithms and fuzzy C-means algorithms for clustering customers dataset," in *APICS 2022 - 2022 1st International Conference on Smart Technology, Applied Informatics, and Engineering, Proceedings*, Aug. 2022, pp. 211–216, doi: 10.1109/APICS56469.2022.9918726.

[18] Z. Zakaria, K. L. Lo, and M. H. Sohod, "Application of fuzzy clustering to determine electricity consumers' load profiles," in *First International Power and Energy Conference, (PECon 2006) Proceedings*, Nov. 2006, vol. 45, pp. 99–103, doi: 10.1109/PECON.2006.346627.

[19] N. Anuar, N. K. K. Baharin, N. H. M. Nizam, A. N. Fadzilah, S. E. M. Nazri, and N. M. Lip, "Determination of typical electricity load profile by using double clustering of fuzzy C-means and hierarchical method," in *2021 IEEE 12th Control and System Graduate Research Colloquium, ICSGRC 2021 - Proceedings*, Aug. 2021, vol. 2018 janua, pp. 277–280, doi: 10.1109/ICSGRC53186.2021.9515295.

[20] C. Widyastuti, A. Senen, and O. Handayani, "Micro-spatial electricity load forecasting using clustering technique," in *Proceeding - 2020 2nd International Conference on Industrial Electrical and Electronics, ICIEE 2020*, Oct. 2020, vol. 2018, pp. 17–21, doi: 10.1109/ICIEE49813.2020.9277274.

[21] H. Bian, Y. Zhong, J. Sun, and F. Shi, "Study on power consumption load forecast based on K-means clustering and FCM–BP model," *Energy Reports*, vol. 6, pp. 693–700, Dec. 2020, doi: 10.1016/j.egyr.2020.11.148.

[22] J. Hämäläinen, S. Jauhiainen, and T. Kärkkäinen, "Comparison of internal clustering validation indices for prototype-based clustering," *Algorithms*, vol. 10, no. 3, Sep. 2017, doi: 10.3390/a10030105.

[23] R. Al-Otaibi, N. Jin, T. Wilcox, and P. Flach, "Feature construction and calibration for clustering daily load curves from smart-meter data," *IEEE Transactions on Industrial Informatics*, vol. 12, no. 2, pp. 645–654, Apr. 2016, doi: 10.1109/TII.2016.2528819.

[24] H. Bulut, Y. Öner, and Ç. Sözen, "Clustering of member and candidate countries of the European Union," *International Journal of Sciences: Basic and Applied Research (IJSBAR)*, vol. 36, no. 7, pp. 18–25, 2017.
[25] T. Thinsungnoen, N. Kaoungku, P. Durongdumronchai, K. Kerdprasop, and N. Kerdprasop, "The clustering validity with silhouette and sum of squared errors," in *The Proceedings of the 2nd International Conference on Industrial Application Engineering 2015*, 2015, pp. 44–51, doi: 10.12792/iciae2015.012.
[26] X. Wang and Y. Xu, "An improved index for clustering validation based on Silhouette index and Calinski-Harabasz index," *IOP Conference Series: Materials Science and Engineering*, vol. 569, no. 5, Jul. 2019, doi: 10.1088/1757-899X/569/5/052024.
[27] N. Thamrin and A. W. Wijayanto, "Comparison of soft and hard clustering: a case study on welfare level in cities on Java Island," *Indonesian Journal of Statistics and Its Applications*, vol. 5, no. 1, pp. 141–160, Mar. 2021, doi: 10.29244/ijsa.v5i1p141-160.
[28] B. Ratner, "The correlation coefficient: Its values range between 1/1, or do they," *Journal of Targeting, Measurement and Analysis for Marketing*, vol. 17, no. 2, pp. 139–142, May 2009, doi: 10.1057/jt.2009.5.
[29] S. Capareda, *Introduction to biomass energy conversions*. CRC Press, 2013.
[30] T. G. Chuah, A. G. K. Wan Azlina, Y. Robiah, and R. Omar, "Biomass as the renewable energy sources in Malaysia: An overview," *International Journal of Green Energy*, vol. 3, no. 3, pp. 323–346, Sep. 2006, doi: 10.1080/01971520600704779.

BIOGRAPHIES OF AUTHORS

Ginas Alvianingsih received a bachelor's and master's degree in electrical engineering from Universitas Indonesia, Indonesia, in 2016 and 2017. She is pursuing a Ph.D. in the Faculty of Chemical and Energy Engineering at Universiti Teknologi Malaysia. Her research interests mainly focus on renewable energy and electrical power generation. She can be contacted at email: ginas@itpln.ac.id.

Haslenda Hashim is a chemical and energy engineering Professor and Head of the Green Energy and Environmental Planning (GREEN) research group at the University of Technology, Malaysia (UTM). She first obtained her B.Eng. (Hons) and MSc. in chemical engineering at UTM. She also holds a PhD in Carbon Capture Utilizations and Storage (CCUS) from the University of Waterloo, Canada. She is a registered professional engineer to the Board of Engineers Malaysia (BEM) and is one of the AEMAS Certified Trainers for the Energy Manager Training Course (EMTC), which focuses on relay coordination projects and off-grid solar PV design. She is the author and co-author of more than 80 publications in international journals and proceedings in power systems and energy. Her research interests include network reconfiguration, optimization techniques, and renewable energy. She can be contacted at email: haslenda@utm.my.

Jasrul Jamani Jamian received a Bachelor of Engineering (B.Eng. (Hons)) degree, Master of Engineering (M.Eng.), and Ph.D degree in electrical (power) engineering from Universiti Teknologi Malaysia in 2008, 2010, and 2013, respectively. He is currently the Director of the Power Engineering Division at the School of Electrical Engineering, Universiti Teknologi Malaysia. He is actively involved in research as a principal investigator and leader in consultancy projects with several companies, such as Petronas and Tenaga Nasional Berhad, which focuses on relay coordination projects and off-grid solar PV design. He is the author and co-author of more than 80 publications in international journals and proceedings in power systems and energy. His research interests include network reconfiguration, optimization techniques, and renewable energy. He can be contacted at email: jasrul@fke.utm.my.

Adri Senen received a bachelor's degree in electrical engineering from Andalas University, Indonesia, in 2004 and a Master's degree in electrical power engineering from Bandung Institute of Technology (ITB), Indonesia, in 2008. He is a PhD student in the Faculty of Electrical Engineering at Universiti Teknologi Malaysia. His research interests concern load forecasting, management energy, electrical planning, renewable energy, and power systems. He can be contacted at email: adrisenen@itpln.ac.id.