

Integration of K-Means Clustering and Elbow Method for Mapping *Baccaurea* spp. Distribution to Support Agroindustrial Development in West Sulawesi

Heri Irawan^{*1}, Muhammad Ihsan Fawzi², Syarif Hidayat³

^{1,3}Industrial Engineering, Faculty of Engineering, Universitas Jenderal Soedirman, Indonesia

²Informatics, Faculty of Engineering, Universitas Jenderal Soedirman, Indonesia

Email: ¹heri.irawan@unsoed.ac.id

Received : Apr 28, 2025; Revised : May 28, 2025; Accepted : May 28, 2025; Published : May 29, 2025

Abstract

Baccaurea spp. is a type of wild plant with potential value for sustainable agroindustrial development. This study aims to map and segment regions in West Sulawesi based on the habitat suitability for *Baccaurea* spp. using K-Means Clustering integrated with the Elbow Method. Field data were collected from 25 villages in Mamasa and Mamuju districts, involving five parameters: land area, production estimate, altitude, humidity, and average temperature. Based on the results of the exploration, 3 species of *Baccaurea* spp. have been found, namely *Baccaurea Lanceolata*, *Baccaurea Costulata*, and *Baccaurea Racemosa*. The analysis yielded three clusters, with Cluster 1 being identified as the top priority for agroindustrial development due to its high productivity and optimal land conditions. The findings provide a data-driven foundation for policymakers and industries to support the sustainable cultivation of *Baccaurea* spp. in Indonesia. This research contributes to informatics-based decision-making in agroindustry development and regional planning.

Keywords : Agroindustry, *Baccaurea* spp., Clustering, Elbow Method, K-Means, Mapping.

This work is an open access article and licensed under a Creative Commons Attribution-Non Commercial 4.0 International License



1. PENDAHULUAN

Tanaman liar merupakan tanaman yang tumbuh dengan sendirinya di alam tanpa campur tangan manusia, tanaman liar ini biasanya dapat digunakan sebagai makanan, obat-obatan, bahan bangunan, bahan bakar, hiasan, atau bahkan untuk tujuan “magis” [1]. Salah satu tanaman liar tersebut adalah genus *Baccaurea* yang berpotensi untuk pengembangan agroindustry. Menurut [2], *Baccaurea* merupakan kelompok tanaman asli tropis yang memiliki potensi tinggi untuk meningkatkan ketahanan pangan dan kesehatan masyarakat, terutama di daerah pedesaan. Tanaman liar ini perlu dipetakan untuk pengembangan agroindustry, dimana dalam pengembangan tersebut diperlukan segmentasi untuk mengidentifikasi persebaran tanaman liar yang ada.

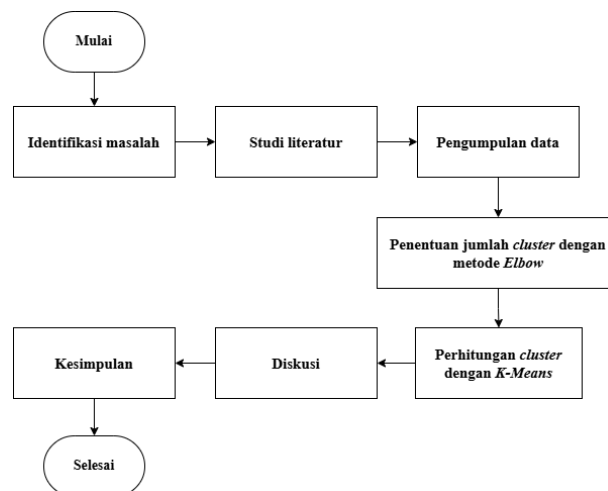
Salah satu teknik yang dapat diterapkan adalah *K-Means Clustering* yang berfungsi untuk melakukan pengelompokan suatu data dalam jumlah besar, dalam waktu yang cepat dan efektif [3][4]. Pendekatan pemodelan dengan *K-Means* memiliki tujuan yaitu untuk mengelompokkan data menjadi satu atau lebih kelompok yang memiliki kesamaan, tindakan tersebut dilakukan dengan tujuan agar data yang memiliki kesamaan dapat digabung dalam satu kelompok dan data yang tidak sama akan terpisah dalam kelompok berbeda [5][6]. Namun, *K-Means* memiliki kelemahan dalam menentukan jumlah *cluster* yang dihasilkan, dimana tidak semua *cluster* yang dihasilkan tidak sesuai dengan kebutuhan analisis [7][8]. Dengan demikian, *K-Means* perlu didukung dengan metode *Elbow* atau nilai SSE yang berfungsi untuk menentukan kekuatan *cluster* yang terbentuk [7][9]. Metode *Elbow* adalah metode yang

sering kali digunakan untuk menentukan jumlah *cluster* yang optimal, dimana hal tersebut dilakukan dengan cara mengamati persentase varians setiap *cluster* yang membentuk pola menyerupai siku pada titik tertentu [10][11]. Sedangkan menurut [12], metode *Elbow* itu sendiri merupakan suatu metode pendukung yang dapat digunakan untuk menentukan nilai K yang optimal dalam penerapan metode *K-Means*, dimana biasanya penerapan metode ini berfokus pada persentase *varians* sebagai fungsi dari jumlah *cluster*, dan dalam penerapannya untuk mencari nilai K yang paling optimal maka setiap nilai K akan dicek satu persatu untuk kemudian dicatat nilai SSE (*Sum Square Error*).

Penelitian mengenai integrasi *K-Means* dengan metode *Elbow* telah dilakukan dalam beberapa penelitian. Diantaranya penelitian [13] yang menggunakan integrasi metode *Elbow* untuk mengimprovisasi *K-Means*. Penelitian [14] juga menerapkan metode *Elbow* untuk menentukan nilai K yang paling optimal. Hal yang sama juga dilakukan pada penelitian [15] yang menggunakan *Sum Squared Error* (SSE) untuk meningkatkan performa metode *Elbow*. Dengan demikian, metode *Elbow* merupakan metode yang cukup efektif diterapkan untuk diintegrasikan dengan *K-Means* agar dapat menghasilkan *cluster* yang lebih optimal [16]. Perbedaan penelitian ini adalah belum adanya pemetaan berbasis integrasi *K-Means* dan *Elbow* terhadap tanaman *Baccaurea Spp.* di Sulawesi Barat menggunakan pendekatan spasial dan agroindustri. Dengan Temuan pada penelitian ini diharapkan dapat memberikan landasan berbasis data bagi pembuat kebijakan dan industri untuk mendukung budidaya *Baccaurea spp.* di Indonesia. Penelitian ini memberikan kontribusi terhadap pengambilan keputusan berbasis informatika dalam pengembangan agroindustri dan perencanaan wilayah.

2. METODE PENELITIAN

Gambar 1 berikut ini menggambarkan proses penelitian clusterisasi dan pemetaan *Baccaurea spp.* untuk mendukung pengembangan agroindustri di Sulawesi Barat, yang menguraikan tahapan dari identifikasi masalah, studi literatur, pengumpulan data, hingga perhitungan data untuk menyelesaikan permasalahan yang ada.



Gambar 1. Flowchart penelitian

2.1. Identifikasi Masalah

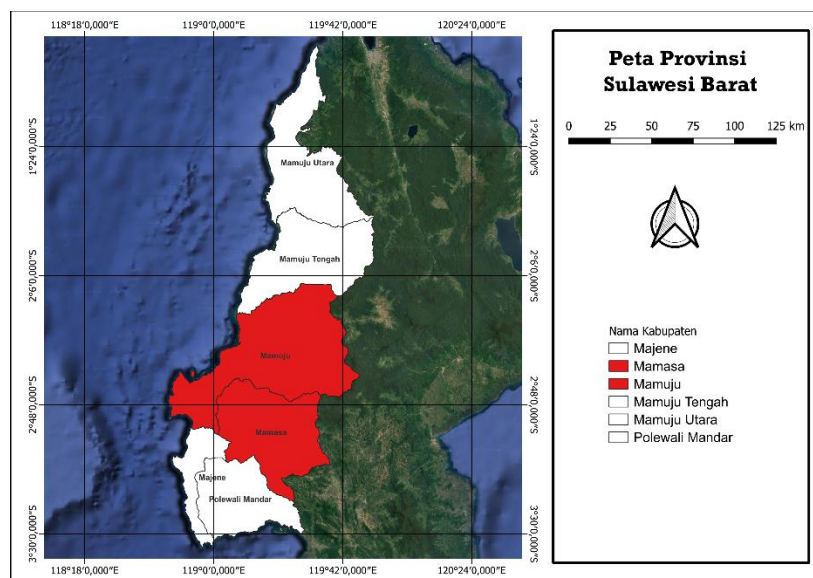
Identifikasi masalah dalam penelitian ini adalah bagaimana melakukan pengelompokkan dan pemetaan potensi tanaman *Baccaurea spp.* di Sulawesi Barat, dengan tujuan untuk dapat memberikan rekomendasi dan pertimbangan bagi pihak pembuat kebijakan dan industri untuk dapat mengetahui segmentasi daerah di Sulawesi Barat yang memiliki potensi terdapat tanaman *Baccaurea spp.* agar kedepannya dapat dikembangkan menjadi agroindustri.

2.2. Studi Literatur

Penelitian yang diusulkan adalah mengenai pengelompokkan dan pemetaan potensi tanaman *Baccaurea spp.* di Sulawesi Barat, dengan menerapkan algoritma *K-Means* dengan disertai perhitungan optimasi *cluster* dengan menggunakan metode *Elbow*. Oleh karena itu, peneliti melakukan studi literatur dengan menelaah berbagai tutorial mengenai penerapan data mining, serta membaca artikel jurnal dan hasil penelitian terkait metode *clustering*, guna memperkaya referensi dan memperkuat landasan teoritis dalam penelitian ini.

2.3. Pengumpulan Data

Pengumpulan data pada penelitian ini dilakukan dengan melakukan eksplorasi pada desa/kelurahan yang memiliki potensi terdapat tanaman *Baccaurea spp.* di Kabupaten Mamasa dan Mamuju, Provinsi Sulawesi Barat. Dimana lokasi penelitian ini dipilih berdasarkan pertimbangan bahwa kondisi lingkungan di dua kabupaten tersebut sesuai dengan habitat tanaman *Baccaurea spp.* Pada pengumpulan data ini terdapat 25 desa/kelurahan di Kabupaten Mamasa dan Mamuju yang menjadi lokasi ditemukannya tanaman *Baccaurea spp.*, dimana pada proses pengumpulan data ini juga dilakukan pengukuran variabel suhu rata-rata dan kelembaban dengan menggunakan alat *environmental meter*, dan pengukuran variabel ketinggian dengan menggunakan alat *altimeter*, proses pengukuran dilakukan dengan durasi 2-5 menit untuk memastikan data sesuai dengan kondisi yang ada. Sedangkan proses pengumpulan data untuk variabel kapasitas lahan dan estimasi produksi dilakukan melalui diskusi dan *brainstorming* dengan pakar biologi dan masyarakat setempat untuk mengetahui perkiraan terkait data variabel tersebut. Gambar 2 berikut ini menunjukkan peta lokasi penelitian, dimana Kabupaten Mamasa dan Mamuju yang merupakan lokasi penelitian ditandai dengan warna merah.



Gambar 2. Peta Lokasi penelitian

2.4. Penentuan Jumlah Cluster Dengan Metode Elbow

Metode *Elbow* merupakan metode yang biasanya digunakan untuk menentukan jumlah *cluster* yang ideal, setelah secara berurutan menganalisis kumpulan data yang berasal dari data primer [17]. Metode *Elbow* adalah salah satu metode yang sering digunakan untuk menentukan jumlah *cluster* yang optimal, dimana hal tersebut dilakukan dengan cara mengamati persentase varians setiap *cluster* yang membentuk pola menyerupai siku pada titik tertentu [10]. Sedangkan menurut [7], metode *Elbow* itu sendiri merupakan suatu metode pendukung yang dapat menentukan nilai *K* yang optimal dalam penerapan metode *K-Means*, dimana biasanya penerapan metode ini berfokus pada persentase *variens*

sebagai fungsi dari jumlah *cluster*, dan dalam penerapannya untuk mencari nilai K yang paling optimal maka setiap nilai K akan dicek satu persatu untuk kemudian dicatat nilai SSE (*Sum Squared Error*).

Adapun persamaan yang umum digunakan dalam metode *Elbow* adalah perhitungan nilai WSS (*Within Cluster Sum of Squares*), atau biasanya disebut dengan SSE (*Sum Squared Error*) dengan rumus sebagai berikut [10]:

$$SSE = \sum_{k=1}^k \sum_{x_i \in C_k} (X_i - Y_k)^2 \quad (1)$$

Keterangan:

C_k = K *cluster* yang terbentuk

k = banyak *cluster*

X_i = data x pada fitur ke-i

Y_k = rata-rata K *cluster* pada nilai k ($k = 1, 2, 3, \dots, K$)

2.5. Perhitungan Cluster Dengan Metode K-Means

Proses analisis data juga dilakukan dengan menggunakan Metode *K-Means Clustering*. Dimana *K-Means Clustering* adalah suatu metode *clustering* yang mampu mengelompokkan data dalam jumlah yang cukup besar dengan waktu komputasi yang cepat dan efisien [18] [19]. Metode ini dapat digunakan untuk melakukan pengelompokan data dalam satu *cluster*/kelompok atau lebih [20].

Berikut merupakan tahapan yang dilakukan ketika menggunakan metode *K-Means Clustering* [21]:

- Tentukan jumlah k sebagai jumlah *cluster* yang akan dibentuk.
- Tentukan secara acak k *centroid* (titik pusat *cluster*).

$$v = \frac{\sum_{i=1}^n x_i}{n}; i = 1, 2, 3, \dots, n \quad (2)$$

Dimana :

v = *centroid* pada *cluster*;

x_i = objek ke-I;

n = banyaknya objek/jumlah objek yang menjadi anggota *cluster*.

- Lakukan perhitungan jarak setiap data yang ada terhadap seluruh *centroid* dari masing-masing *cluster*. Perhitungan dapat dilakukan dengan menggunakan rumus berikut ini:

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2}; i = 1, 2, 3, \dots, n \quad (3)$$

Dimana :

x_i = objek x ke-I;

y_i = daya y ke-I;

n = banyaknya objek

- Setiap objek diklasifikasikan ke dalam kelompok berdasarkan kedekatannya dengan pusat *cluster* (*centroid*) terdekat.
- Proses dilanjutkan dengan iterasi untuk memperbarui posisi *centroid* menggunakan rumus yang ada.
- Ulangilah langkah ke 3 apabila posisi *centroid* baru tidak sama.

3. RESULT

3.1. Data Penelitian

Penelitian *K-Means* dilakukan dengan mengambil data melalui eksplorasi yang dilakukan di desa-desa di Kabupaten Mamasa dan Mamuju yang bersinggungan langsung dengan Gunung Gandang Dewata dan memiliki karakteristik yang sesuai dengan habitat tanaman *Baccaurea spp.* Tabel 1 berikut ini adalah data hasil eksplorasi tumbuhan *Baccaurea spp.* di Kabupaten Mamasa dan Kabupaten Mamuju

Tabel 1. Data hasil penelitian tanaman *Baccaurea spp.* di Kabupaten Mamasa dan Kabupaten Mamuju

No	Nama Desa	Kapasitas Lahan (Ha)	Estimasi Produksi (Kg/Bulan)	Ketinggian (Mdpl)	Kelembaban (%)	Suhu Rata-Rata ($^{\circ}$ C)
1	Malatiro	303	1000	600	53	24
2	Lakahang	575	345	1200	54	28
3	Pangandaran	416	865	800	50	24
4	Burana	568	373	1000	53	24
5	Tampak Kurra	181	285	900	54	25
6	Salutambun Timur	141	549	800	50	24
7	Salutambun	123	448	700	51	25
8	Aralle Timur	264	185	800	53	25
9	Kebanga	158	336	800	54	24
10	Aralle Selatan	79	228	700	53	25
11	Tapalinna	136	499	700	52	24
12	Talipukki	218	479	700	53	25
13	Bujung Manurung	78	329	600	53	25
14	Sondong Layuk	96	389	700	53	25
15	Sendana	85	368	700	53	25
16	Uhaimatte	249	585	1100	50	28
17	Keang	459	685	1000	54	25
18	Pammulukang	261	479	800	55	26
19	Guliling	409	735	800	54	25
20	Sondoang	252	429	800	53	26
21	Salutiwo	503	1250	700	52	24
22	Bone Hau	1535	885	700	53	25
23	Tamalea	215	385	800	52	25
24	Hinua	296	499	700	53	24
25	Banuada	1129	769	700	53	24

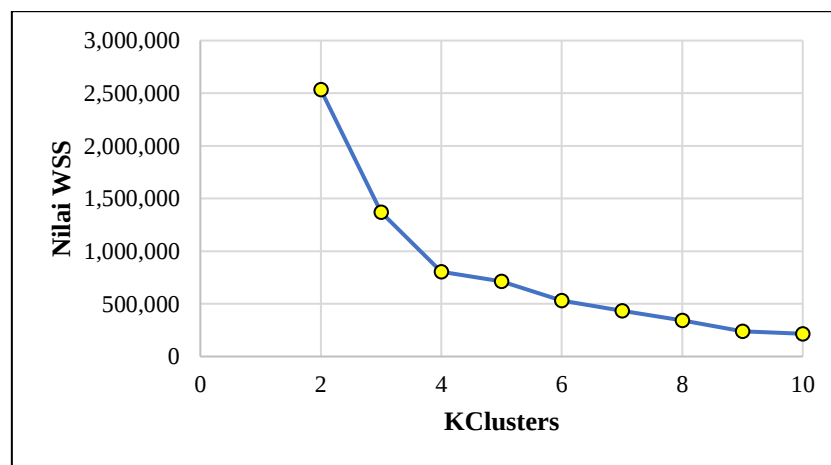
3.2. Metode Elbow

Langkah pertama dalam melakukan *clustering* dengan menggunakan metode *K-Means Clustering* adalah dengan menentukan jumlah *k* terlebih dahulu [13] [22], dimana kali ini penentuan jumlah *k* dilakukan dengan menggunakan metode *Elbow* untuk menentukan jumlah *k* yang paling optimal [23]. Gambar 3 menggambarkan grafik hasil perhitungan WSS (*Within Cluster Sum of Square*) yang berbentuk siku atau yang biasa disebut dengan titik siku [24]. Tabel 2 berikut ini merupakan data hasil perhitungan *Elbow* dengan jumlah *cluster* percobaan 2 - 10 *cluster*.

Tabel 2. Hasil perhitungan *elbow*

Cluster	WSS	Selisih
2	2.533.594,141	
3	1.370.875,000	1.162.719,141
4	803.551,287	567.323,713
5	714.404,800	89.146,487
6	530.645,262	183.759,538
7	433.455,768	97.189,494
8	341.851,283	91.604,485
9	238.900,320	102.950,963
10	215.966,820	22.933,500

Keterangan :

WSS : *Within Cluster Sum of Square*Gambar 3. Grafik Hasil Perhitungan *Elbow*

Berdasarkan grafik hasil perhitungan *Elbow* pada Gambar 3 dapat disimpulkan bahwa titik optimal (titik siku) berada pada $K=3$, sehingga jumlah *cluster* yang optimal adalah 3 *cluster*. Alasan pemilihan 3 *cluster* karena berdasarkan hasil perhitungan *elbow* terjadi penurunan Within SS yang cukup signifikan dari $K=2$ ke $K=3$, selain itu setelah $K=3$ penambahan *cluster* tidak memberikan penurunan Within SS yang signifikan.

3.3. Metode *K-Means Clustering*

Setelah menentukan jumlah K dengan menggunakan metode *elbow*, langkah selanjutnya adalah perhitungan *K-Means* yaitu menentukan pusat *cluster* awal/*centroid* [25] [26]. Tabel 3 berikut ini merepresentasikan hasil penentuan pusat *cluster* awal/*centroid*.

Table 3. Hasil Perhitungan Pusat *Cluster* Awal/*Centroid*

Variable	Pusat <i>Cluster</i> awal/ <i>Centroid</i>		
	Cluster		
	1	2	3
Kapasitas lahan	503.00	1535.00	79.00
Estimasi produksi	1250.00	885.00	228.00
Ketinggian	700	700	600
Kelembaban	52.00	53.00	53.00
Suhu rata-rata	24.00	25.00	25.00

Berdasarkan hasil penentuan pusat *cluster* awal dengan bantuan software SPSS pada tabel diatas, diperoleh 3 pusat *cluster*, dimana untuk *cluster* pertama yang menjadi pusat *cluster* adalah data ke-10. Kemudian untuk *cluster* kedua, pusat *cluster* nya adalah data ke-21. Sedangkan untuk *cluster* 3, pusat *cluster* nya adalah data ke-22. Setelah dilakukan penentuan *centroid* secara acak, langkah selanjutnya adalah melakukan perhitungan jarak antara setiap data dengan *centroid* menggunakan *Euclidean* sebagaimana tercantum pada persamaan 3 [27] [28]. Berikut terdapat data hasil perhitungan jarak pada iterasi ke-1 dan penentuan jarak terdekat, dapat dilihat pada Tabel 4 di bawah ini.

Tabel 4. Hasil Perhitungan Jarak Pada Iterasi Ke-1

Data ke - i	Jarak ke <i>centroid</i>			Jarak Terdekat	Cluster Yang Diikuti
	C1	C2	C3		
1	336	1242	804	336	1
2	1036	1210	787	787	3
3	407	1123	748	407	1
4	929	1135	648	648	3
5	1037	1494	322	322	3
6	795	1438	383	383	3
7	887	1478	246	246	3
8	1096	1455	276	276	3
9	982	1486	241	241	3
10	1106	1597	100	100	3
11	836	1451	294	294	3
12	822	1378	304	304	3
13	1019	1563	101	101	3
14	952	1522	190	190	3
15	976	1539	172	172	3
16	817	1380	637	637	3
17	641	1135	716	641	1
18	814	1341	369	369	3
19	533	1140	637	533	1
20	864	1365	332	332	3
21	0	1095	1111	0	1
22	1094	0	1600	0	2
23	917	1415	288	288	3
24	779	1297	362	362	3
25	789	422	1185	422	2

Setelah melakukan perhitungan iterasi 1, selanjutnya dilakukan perhitungan iterasi 2 dengan tujuan untuk mengecek apakah data pada setiap kelompok mengalami perubahan atau tetap seperti hasil perhitungan iterasi sebelumnya. Tabel 5 berikut ini adalah hasil perhitungan pusat cluster awal untuk perhitungan iterasi 2.

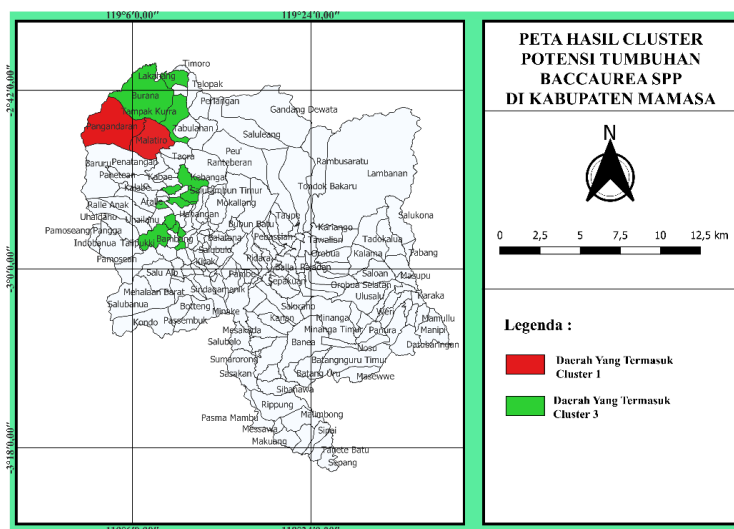
Tabel 5. Hasil Perhitungan Pusat Cluster Awal/Centroid Untuk Iterasi-2

Variable	Pusat Cluster awal/Centroid		
	Cluster		
	1	2	3
Kapasitas lahan	418.00	1332.00	221.00
Estimasi produksi	907.00	827.00	339.00
Ketinggian	780	700	806
Kelembaban	53.00	53.00	53.00
Suhu rata-rata	24.00	25.00	25.00

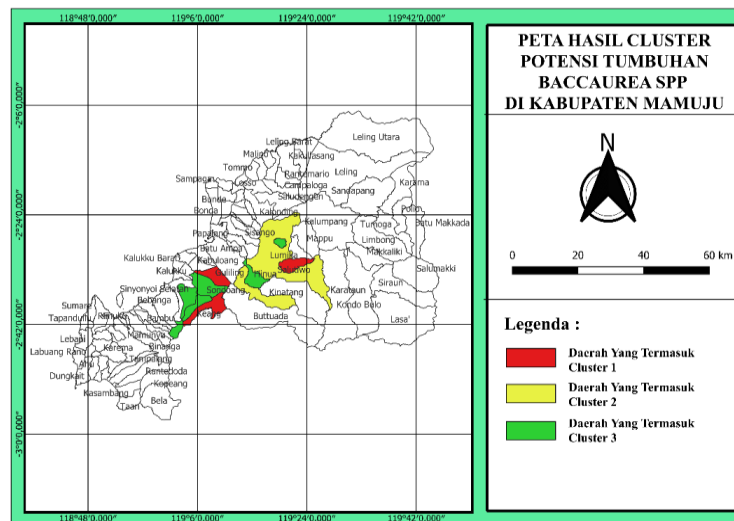
Berdasarkan hasil penentuan pusat *cluster* awal/*centroid*, telah didapatkan 3 pusat *cluster* baru, setelah didapatkan pusat *cluster* awal untuk masing-masing *cluster*, selanjutnya dilakukan perhitungan iterasi 2, dimana hasil perhitungan ini didapatkan data yang sama dengan hasil iterasi 1. Dengan cara yang sama seperti pada saat perhitungan iterasi ke-1, perhitungan yang dilakukan pada iterasi ke-2 menghasilkan nilai yang dapat dilihat pada Tabel 6 di bawah ini.

Tabel 6. Hasil perhitungan jarak pada iterasi ke-1

No	Jarak ke centroid			Jarak Terdekat	Cluster Yang Diikuti
	C1	C2	C3		
1	233	1048	640	233	1
2	719	1027	533	533	3
3	47	922	505	47	1
4	597	938	398	398	3
5	676	1287	154	154	3
6	453	1227	170	170	3
7	551	1267	152	152	3
8	739	1250	219	219	3
9	628	1276	89	89	3
10	763	1389	246	246	3
11	502	1240	168	168	3
12	479	1167	132	132	3
13	694	1353	260	260	3
14	615	1311	164	164	3
15	639	1329	175	175	3
16	484	1179	349	349	3
17	315	934	420	315	1
18	456	1130	89	89	3
19	173	933	385	173	1
20	506	1155	44	44	3
21	362	931	902	362	1
22	1120	211	1405	211	2
23	560	1205	17	17	3
24	433	1086	164	164	3
25	729	211	986	211	2

Gambar 4. Peta hasil cluster tanaman *Baccaurea Spp.* pada Kabupaten Mamasa

Hasil perhitungan iterasi 2 menunjukkan bahwa tidak terdapat perbedaan yang signifikan antara data yang didapatkan pada iterasi 1 dan iterasi 2. Hal ini menunjukkan bahwa tidak terdapat perbedaan yang signifikan antara data yang didapatkan pada iterasi 1 dan iterasi 2, sehingga perhitungan tidak dilanjutkan ke tahap selanjutnya [29][30]. Berikut adalah hasil akhir dari 3 jenis cluster yang digambarkan pada peta hasil cluster tanaman *Baccaurea* di bawah ini, dimana Gambar 4 merepresentasikan peta hasil cluster tanaman *Baccaurea* spp. pada Kabupaten Mamasa dan Gambar 5 merepresentasikan peta hasil cluster tanaman *Baccaurea* spp. pada Kabupaten Mamuju.



Gambar 5. Peta hasil *cluster* tanaman *Baccaurea* Spp. pada Kabupaten Mamuju

4. DISKUSI

Secara keseluruhan penelitian ini menginformasikan potensi pengembangan tanaman *Baccaurea* di Sulawesi sebagai agroindustri. Hal ini dapat dilihat dari ditemukannya 3 jenis tanaman *Baccaurea* tersebut dengan populasi yang cukup melimpah, yang membuka peluang besar pengembangan untuk budidaya dan pemanfaatan secara komersial. Salah satu alasan pentingnya pengembangan tanaman *Baccaurea* spp. ini salah satunya adalah karena pasar bahan baku sediaan obat di Indonesia pada saat ini masih sangat tergantung dari bahan impor, sehingga potensi ini sangat sayang apabila tidak dimanfaatkan dan diprioritaskan. Pengembangan tanaman *Baccaurea* spp. dapat menjadi alternatif untuk mengurangi ketergantungan impor dan mendorong kemandirian farmasi nasional. Penelitian ini berpotensi menjadi landasan ilmiah bagi pengambil kebijakan dan pelaku industri untuk lebih memperhatikan dan mengoptimalkan pengembangan tanaman *Baccaurea* spp. di Sulawesi.

Hasil clustering yang dilakukan dalam penelitian ini menggunakan metode *K-Means* menunjukkan adanya pengelompokan wilayah berdasarkan tingkat potensi keberadaan *Baccaurea* spp. Metode ini tidak hanya mampu mengidentifikasi wilayah dengan potensi tinggi, sedang, dan rendah, tetapi juga memberikan visualisasi spasial yang membantu dalam menyusun strategi pengembangan wilayah. Studi serupa seperti yang dilakukan oleh [31] dalam pemetaan pengelompokan lahan produksi jagung di Kabupaten Pasuruan juga menggunakan metode *K-Means* untuk mengelompokkan wilayah berdasarkan luas panen dan hasil produksi jagung di setiap kecamatan yang ada.

Di sisi lain penelitian [32] dengan menggunakan metode *K-Means* berhasil melakukan pemetaan dan pengelompokan lahan produksi tandan buah segar yang dapat memberikan informasi yang berharga bagi perusahaan di lokasi penelitian. Sementara itu, penelitian [33] melakukan clusterisasi dengan menggunakan metode *K-Means* dengan berdasarkan potensi pertanian di Jawa Tengah, dimana pada penelitian ini berhasil mendapatkan data *cluster* kabupaten di Jawa Tengah yang memiliki potensi pertanian. Oleh karena itu, berdasarkan penelitian yang ada menunjukan bahwa metode *K-Means* sangat

efektif dalam menyederhanakan data kompleks menjadi kelompok yang bermakna untuk pengambilan keputusan. Dalam konteks ini, peran Ilmu Informatika sangat vital dalam mendukung keputusan agroindustri berbasis data. Dengan penerapan teknologi data mining, dan *geographic information systems* (GIS), analisis terhadap data spasial dan non-spasial dapat dilakukan secara efisien untuk menggali pola, potensi, dan tren yang tersembunyi. Informatika menyediakan kerangka kerja untuk melakukan prediksi dan simulasi untuk skenario pengembangan agroindustri. Dengan pendekatan ini, pengambilan keputusan menjadi lebih objektif, transparan, dan berbasis bukti (*evidence-based*), sehingga mampu meningkatkan daya saing sektor pertanian dan agroindustri lokal.

5. KESIMPULAN

Penelitian ini berhasil mengidentifikasi dan memetakan potensi tanaman *Baccaurea spp.* di Kabupaten Mamasa dan Mamuju, Sulawesi Barat. Tiga spesies utama yang ditemukan adalah *Baccaurea Lanceolata*, *Baccaurea Costulata*, dan *Baccaurea racemosa*. Analisis menggunakan integrasi *K-Means Clustering* dan metode *Elbow* menghasilkan tiga segmentasi kelompok desa berdasarkan kapasitas lahan, estimasi produksi, ketinggian, kelembaban, dan suhu rata-rata. Di mana untuk *cluster* yang terpilih sebagai prioritas pengembangan adalah *cluster* 1. Berdasarkan hasil eksplorasi dan klusterisasi, *cluster* 1 merupakan daerah segmentasi yang menunjukkan potensi paling besar untuk dikembangkan menjadi agroindustri, di mana untuk pengembangan secara berkelanjutan nantinya dapat dilakukan dengan budidaya pada lahan pekarangan atau kebun-kebun milik warga masyarakat sekitar daerah yang termasuk dalam *cluster* 1. Sedangkan untuk *cluster* 2 dan 3 nantinya dapat menjadi daerah segmentasi pengembangan tahap selanjutnya. Kontribusi lain dari penelitian ini adalah memperkaya penerapan ilmu komputasi dan spasial dalam analisis pengelompokan wilayah berbasis karakteristik lingkungan, yang dapat digunakan sebagai acuan metodologis dalam studi pemetaan potensi sumber daya hayati lainnya. Sedangkan saran untuk penelitian di masa mendatang adalah melakukan analisis spasial yang lebih mendalam dengan memanfaatkan pengelompokan berbasis GIS atau penerapan model *machine learning* untuk meningkatkan akurasi dan ketepatan segmentasi wilayah.

UCAPAN TERIMA KASIH

Terimakasih kami ucapkan kepada pihak DRTPM yang telah memberikan pendanaan Riset Katalis 2024, dan tak lupa kami ucapkan terimakasih kepada semua tim penelitian yang terlibat.

DAFTAR PUSTAKA

- [1] Á. Vári, I. Arany, Á. Kalóczkai, K. Kelemen, J. Papp, and B. Czúcz, "Berries, greens, and medicinal herbs - Mapping and assessing wild plants as an ecosystem service in Transylvania (Romania)," *J. Ethnobiol. Ethnomed.*, vol. 16, no. 1, pp. 1–14, 2020.
- [2] Z. I. Navia, A. B. Suwardi, and T. Harmawan, "The diversity of wild Tampoi (*Baccaurea*, *Phyllantaceae*) and their potential to improve the livelihoods of local people in Aceh, Indonesia," *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1114, no. 1, 2022, doi: 10.1088/1755-1315/1114/1/012088.
- [3] H. Mahulae, "Pengelompokan Potensi Produksi Buah-Buahan di Provinsi Sumatera Utara dengan Menerapkan K-Clustering (Studi Kasus : Dinas Tanaman Pangan dan Holtikultura)," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 2, p. 312, 2020, doi: 10.30865/jurikom.v7i2.2122.
- [4] D. F. Pasaribu, I. S. Damanik, E. Irawan, Suhada, and H. S. Tambunan, "Memanfaatkan Algoritma K-Means Dalam Memetakan Potensi Hasil Produksi Kelapa Sawit PTPN IV Marihat," *BIOS J. Teknol. Inf. dan Rekayasa Komput.*, vol. 2, no. 1, pp. 11–20, 2021, doi: 10.37148/bios.v2i1.17.
- [5] I. Sanela, A. Nazir, F. Syafria, E. Haerani, and L. Oktavia, "Penerapan Metode Clustering Dengan K-Means Untuk Memetakan Potensi Tanaman Padi di Sumatera," *J. Comput. Syst.*

-
- Informatics*, vol. 5, no. 1, pp. 82–92, 2023, doi: 10.47065/josyc.v5i1.4523.
- [6] E. Muningsih, I. Maryani, and V. R. Handayani, “Penerapan Metode K-Means dan Optimasi Jumlah Cluster dengan Index Davies Bouldin untuk Clustering Propinsi Berdasarkan Potensi Desa,” *J. Sains dan Manaj.*, vol. 9, no. 1, p. 96, 2021, [Online]. Available: www.bps.go.id
 - [7] B. Wang, G. Wang, Y. Wang, Z. Lou, S. Hu, and Y. Ye, “A K-means clustering method with feature learning for unbalanced vehicle fault diagnosis,” *Smart Resilient Transp.*, vol. 3, no. 2, pp. 162–176, 2021, doi: 10.1108/srt-01-2021-0003.
 - [8] A. A. Vernanda, A. Faisol, and N. Vendyansyah, “Penerapan Metode K-Means Clustering Untuk Pemetaan Daerah Rawan Kecelakaan Lalu Lintas Di Kota Malang Berbasis Website,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 5, no. 2, pp. 836–844, 2021, doi: 10.36040/jati.v5i2.3791.
 - [9] A. Rajsya and A. Rachman, “Rancang Bangun Penerapan Metode Elbow Pada K-Means Untuk Clustering Data Persediaan Barang,” *LINIER (Literatur Inform. Komputer)*, vol. 1, no. 4, pp. 395–403, 2024, doi: 10.33096/linier.vxix.xxxx.
 - [10] N. A. Maori and E. Evanita, “Metode Elbow dalam Optimasi Jumlah Cluster pada K-Means Clustering,” *Simetris J. Tek. Mesin, Elektro dan Ilmu Komput.*, vol. 14, no. 2, pp. 277–288, 2023, doi: 10.24176/simet.v14i2.9630.
 - [11] N. T. Hartanti, “Metode Elbow dan K-Means Guna Mengukur Kesiapan Siswa SMK Dalam Ujian Nasional,” *J. Nas. Teknol. dan Sist. Inf.*, vol. 06, no. 02, pp. 82–89, 2020, doi: <https://doi.org/10.25077/TEKNOSI.v6i2.2020.82-89>.
 - [12] A. B. H. Kiat, Y. Azhar, and V. Rahmayanti, “Penerapan Metode K-Means Dengan Metode Elbow Untuk Segmentasi Pelanggan Menggunakan Model RFM (Recency, Frequency & Monetary),” *J. Repos.*, vol. 2, no. 7, pp. 945–952, 2020, doi: 10.22219/repositor.v2i7.973.
 - [13] M. A. Syakur, B. K. Khotimah, E. M. S. Rochman, and B. D. Satoto, “Integration K-Means Clustering Method and Elbow Method for Identification of the Best Customer Profile Cluster,” *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 336, no. 1, 2018, doi: 10.1088/1757-899X/336/1/012017.
 - [14] F. Marisa, A. R. Wardhani, W. Purnomowati, A. V. Vitianingsih, A. L. Maukar, and W. Puspitarini, “Potential Customer Analysis Using K-Means With Elbow Method,” vol. 7, no. 2, pp. 307–312, 2023.
 - [15] P. M. Hasugian, B. Sinaga, J. Manurung, and S. A. Al Hashim, “Best Cluster Optimization with Combination of K-Means Algorithm And Elbow Method Towards Rice Production Status Determination,” *Int. J. Artif. Intell. Res.*, vol. 5, no. 1, pp. 102–110, 2021, doi: 10.29099/ijair.v6i1.232.
 - [16] S. S. Febrian and A. Mutasowifin, “Selection of agricultural industry stocks by application of K-means algorithm with Elbow method,” *BIO Web Conf.*, vol. 04003, 2025.
 - [17] T. Santoso, A. Darmawan, N. Sari, M. A. F. Syadza, E. C. B. Himawan, and W. A. Rahman, “Clusterization of Agroforestry Farmers using K-Means Cluster Algorithm and Elbow Method,” *J. Sylva Lestari*, vol. 11, no. 1, pp. 107–122, 2023, doi: 10.23960/jsl.v11i1.646.
 - [18] A. M. Ikotun, A. E. Ezugwu, L. Abualigah, B. Abuhaija, and J. Heming, “K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data,” *J. Inf. Sci.*, vol. 622, pp. 178–210, 2023, doi: 10.1016/j.ins.2022.11.139.
 - [19] E. Irwansyah and M. Faisal, *Advanced Clustering Teori dan Aplikasi*. DEEPUBLISH, 2012.
 - [20] A. A. Abdunnassar and L. R. Nair, “Performance analysis of Kmeans with modified initial centroid selection algorithms and developed Kmeans9+ model,” *Meas. Sensors*, vol. 25, no. 100666, 2023, doi: 10.1016/j.measen.2023.100666.
 - [21] J. Pérez-Ortega, N. Nely Almanza-Ortega, A. Vega-Villalobos, R. Pazos-Rangel, C. Zavala-Díaz, and A. Martínez-Rebollar, *The K -Means Algorithm Evolution*. 2019. doi: 10.5772/intechopen.85447.
 - [22] N. Solikin, B. Hartono, Sugiono, and Linawati, “Farming in Kediri Indonesia: Analysis of cluster k-means,” *IOP Conf. Ser. Earth Environ. Sci.*, vol. 1041, no. 1, 2022, doi: 10.1088/1755-1315/1041/1/012015.
 - [23] M. Al Ghifari and W. T. Harsanti Putri, “Clustering Courses Based On Student Grades Using K-Means Algorithm With Elbow Method For Centroid Determination,” *Inf. J. Ilm. Bid. Teknol. Inf. dan Komun.*, vol. 8, no. 1, pp. 42–46, 2023, doi: 10.25139/inform.v8i1.4519.
 - [24] D. Anggreani and D. Setiawan, “Optimization of K-Means Clustering Method by Using Elbow
-

- Method in Predicting Blood Requirement of Pelamonia Hospital Makassar,” *Iota*, vol. 04, 2024, doi: 10.31763/iota.v4i3.755.
- [25] A. Alamsyah *et al.*, “Customer Segmentation Using the Integration of the Recency Frequency Monetary Model and the K-Means Cluster Algorithm,” *Sci. J. Informatics*, vol. 9, no. 2, pp. 189–196, 2022, doi: 10.15294/sji.v9i2.39437.
- [26] D. A. F. Fadhillah, A. Faisol, and N. Vendyansyah, “Penerapan Metode K-Means Clustering Pada Pemetaan Lahan Kopi Di Kabupaten Malang,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 5, no. 1, pp. 162–170, 2022, doi: 10.36040/jati.v6i1.4617.
- [27] D. A. Setiady and H. Leong, “Implementation of K-Means Algorithm Elbow Method and Silhouette Coefficient for Rainfall Classification,” *Proxies J. Inform.*, vol. 4, no. 1, pp. 18–25, 2024, doi: 10.24167/proxies.v4i1.12433.
- [28] R. P. Harjono and M. A. I. Pakereng, “Penerapan Metode K-Means Clustering Untuk Analisis Potensi Lahan Pangan Pada Provinsi Kalimantan Selatan,” *J-SAKTI (Jurnal Sains Komput. Inform.)*, vol. 7, no. 1, pp. 2549–7200, 2023, [Online]. Available: <http://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/view/596%0Ahttps://ejurnal.tunasbangsa.ac.id/index.php/jsakti/article/viewFile/596/574>
- [29] S. Noya, N. K. Putrianto, and W. V. Salendu, “Classification of Small and Medium Industry (SMI) in Malang Regency Using K-Means based Clustering Method,” *J. Ilm. Tek. Ind.*, vol. 20, no. 2, pp. 233–238, 2021, doi: 10.23917/jiti.v20i2.15374.
- [30] M. Rosyidi and Isnurani, “Analisis Klaster Metode K-Means Dalam Berdasarkan Faktor-Faktor Produksi Bawang Merah,” *MATH Vis. J. Mat.*, vol. 5, no. 1, pp. 16–21, 2023.
- [31] D. D. K. Saputra, K. Auliasari, and A. Faisol, “Penerapan Metode K-Means Clustering untuk Pemetaan Pengelompokan Lahan Produksi Jagung di Kabupaten Pasuruan,” *JATI (Jurnal Mhs. Tek. Inform.)*, vol. 5, no. 8, pp. 8364–8372, 2024, doi: 10.31294/inf.v10i1.15621.
- [32] A. Al Masykur, S. K. Gusti, S. Sanjaya, F. Yanto, and F. Syafria, “Penerapan Metode K-Means Clustering untuk Pemetaan Pengelompokan Lahan Produksi Tandan Buah Segar,” *J. Inform.*, vol. 10, no. 1, 2023, doi: 10.31294/inf.v10i1.15621.
- [33] A. M. Khariyani, “Clustering dengan Metode K-Means Berdasarkan Potensi Pertanian di Jawa Tengah,” in *Prosiding Seminar Nasional Pendidikan Matematika dan Matematika*, 2022. doi: doi.org/10.21831/pspmm.v6i2.246.