

# Segmentasi Provinsi di Indonesia berdasarkan Data Runtun Waktu Produksi Padi dengan Algoritma DTW dan *K-Medoids Clustering*

## *Segmentation of Provinces in Indonesia Using Time Series Data of Rice Production with Dynamic Time Warping Algorithm and K-Medoids Clustering*

**Anne Mudya Yolanda dan Husna Savira**

Program Studi S1 Statistika Universitas Riau  
Kampus Bina Widya Jl. HR Subrantas KM. 12.5 Pekanbaru  
Email: annemudyayolanda@lecturer.unri.ac.id

Diterima: 19 April 2024

Revisi: 3 Januari 2025

Disetujui: 16 Januari 2025

### Abstrak

Sektor pertanian, khususnya tanaman pangan padi, memiliki peranan penting bagi perekonomian Indonesia. Analisis data historis produksi padi atau data runtun waktu produksi padi dapat memberikan gambaran pola produksi dan segmentasi wilayah berdasarkan karakteristik tanaman padi. Penelitian ini menggunakan algoritma DTW dan *K-Medoids Clustering* untuk melakukan segmentasi provinsi di Indonesia berdasarkan data produksi padi tahun 2013-2021. Hasil penelitian menunjukkan tiga *cluster* wilayah dengan karakteristik produksi padi yang berbeda. Setiap *cluster* menunjukkan pola produksi yang berbeda dengan anggota *cluster* lainnya, yang disebabkan oleh perbedaan jarak DTW dalam mengukur kesamaan pola produksi. *Cluster* 1 memiliki produksi tertinggi, diikuti oleh *Cluster* 2 dan *Cluster* 3, masing-masing terdiri dari 16, 7, dan 11 provinsi. Temuan ini dapat digunakan sebagai dasar pengembangan kebijakan pemerintah sesuai karakteristik masing-masing segmen.

kata kunci: padi, segmentasi, DTW, *K-Medoids*, *cluster*.

### Abstract

*The agricultural sector, particularly rice crops, plays a crucial role in Indonesia's economy. Analyzing rice production historical data, or rice production time series data, can provide insights into production patterns and regional segmentation based on rice crop characteristics. This study employed the DTW algorithm and K-Medoids Clustering to segment provinces in Indonesia based on rice production data from 2013 to 2021. The results of the study indicated three clusters of regions with distinct rice production characteristics. Some regions exhibited different production patterns from other cluster members, attributed to variations in DTW distances used to measure pattern similarity. Cluster 1 had the highest production, followed by Cluster 2 and Cluster 3, with 16, 7, and 11 provinces respectively. These findings can serve as a basis for government policy development tailored to the characteristics of each segment.*

keywords: rice, segmentation, DTW, *K-Medoids*, clusters.

## I. PENDAHULUAN

Sektor pertanian memiliki kontribusi yang besar dalam pencapaian tujuan kedua dari pembangunan berkelanjutan/SDGs (*Sustainable development goals*) adalah untuk menghapus kelaparan, mencapai ketahanan pangan, meningkatkan kualitas nutrisi, serta mempromosikan praktik pertanian yang berkelanjutan. Sebagai negara agraris dengan banyak sumber daya alam, terutama tanah

yang subur, mayoritas masyarakat Indonesia berprofesi di sektor pertanian. Sektor ini memegang peranan vital bagi pertumbuhan ekonomi nasional dengan kontribusi sebagai penyumbang terbesar ketiga terhadap Produk Domestik Bruto (Badan Pusat Statistik Provinsi Riau, 2023). Sektor ini juga mendorong roda perekonomian dengan memenuhi permintaan kebutuhan pangan khususnya dalam negeri (Hayati, dkk., 2017).

---

Menurut Badan Pusat Statistik (BPS), sektor pertanian, kehutanan, dan perikanan memiliki pengaruh signifikan terhadap perekonomian tahun 2022 yakni mencapai 12,40 persen (Badan Pusat Statistik Indonesia, 2023a). Meskipun selama pandemi Covid-19 sempat terjadi kontraksi pertumbuhan ekonomi, tetapi lini pertanian berhasil terus mengalami pertumbuhan yang positif. Menurut data tiga tahun terakhir, sektor ini konsisten meningkat dengan pertumbuhan sebesar 1,77, 1,87, dan 2,25 masing-masing pada 2020, 2021, dan 2022 (Badan Pusat Statistik Indonesia, 2023b).

Selain itu, sektor pertanian memiliki kapasitas untuk menarik banyak tenaga kerja, mencakup sekitar 28,61 persen tenaga kerja, menurut Survei Angkatan Kerja Nasional Agustus 2022. (Badan Pusat Statistik Indonesia, 2022). Hal ini mencerminkan peran strategis sektor pertanian dalam menciptakan lapangan kerja dan mendukung kesejahteraan masyarakat setempat. Salah satu produk utama dalam bidang pertanian adalah padi sebagai tanaman pangan. Ketersediaan data luas panen dan produksi adalah salah satu landasan utama bagi pemerintah, baik tingkat pusat maupun daerah dalam upaya untuk meningkatkan kapasitas produksi tanaman padi untuk mencukupi kebutuhan beras.

Pada tahun 2023, total produksi Gabah Kering Giling (GKG) mencapai 53,63 juta ton, dengan area panen seluas 10,20 juta hektar. (Badan Pusat Statistik Indonesia, 2023d). Angka tersebut setara dengan 30,90 juta ton produksi besar untuk konsumsi pangan masyarakat. Produksi padi turun sebesar 1,12 juta ton dari tahun 2022 menjadi 54,75 juta ton GKG dengan luas panen 10,45 juta ton hektar, yang setara dengan produksi beras 31,54 juta ton. Luas panen padi sendiri sebanyak 10,41 juta hektar pada tahun 2021, dengan produksi padi 54,42 juta ton GKG yang dapat dikonversi menjadi 31,36 juta ton beras (Badan Pusat Statistik Indonesia, 2023c).

Data produktivitas tanaman padi di Indonesia dapat dimanfaatkan sebagai dasar dalam perencanaan dan implementasi program pemerintah dalam meningkatkan ketahanan pangan khususnya beras agar lebih tepat sasaran. Salah satu pendekatan statistik yang

dapat diterapkan yaitu melakukan analisis segmentasi provinsi di Indonesia berdasarkan produksi tanaman pangan padi tahun 2013-2021. Pada penelitian ini, provinsi akan dikelompokkan menjadi beberapa *cluster* dengan analisis *clustering*.

Analisis *clustering* banyak diterapkan dalam melakukan segmentasi wilayah berdasarkan variabel-variabel yang menjadi ciri khusus data yang menjadi bahan kajian, di antaranya mengelompokkan wilayah menurut indikator data sosial (Imani, dkk., 2022), karakteristik status Kesehatan masyarakat (Yunitaningtyas dan Yolanda, 2022), indikator Indeks Pembangunan Manusia (Wicaksono dan Yolanda, 2021), dan sarana dan perlengkapan yang digunakan oleh fasilitas kesehatan keluarga berencana (Yolanda dan Yunitaningtyas, 2021). Pemerintah dapat menggunakan hasil analisis *clustering* untuk membuat kebijakan yang tersegmentasi berdasarkan karakteristik unik masing-masing kelompok.

Pada berbagai penelitian terdahulu di atas, analisis *clustering* diterapkan pada data *cross-section* atau hanya pada satu periode waktu tertentu dari berbagai unit observasi (Hill, dkk., 2011). Oleh karena data produktivitas padi tersedia dalam format data historis runtun waktu, maka pada penelitian ini akan menerapkan analisis *cluster time series*. Data *time series* merupakan serangkaian observasi berurutan yang dicatat dalam periode waktu tertentu untuk satu buah variabel penelitian (Montgomery, dkk., 2008).

Analisis *cluster time series* dipakai sebagai pendekatan statistika bagi data runtun waktu yang digunakan untuk mengelompokkan produksi tanaman pangan padi. Penggunaan jarak *Dynamic Time Warping* (DTW) untuk analisis *cluster* memungkinkan penggunaan *K-Medoids Clustering* untuk melihat perbandingan antar waktu dengan berbagai panjang periode karena ukuran ini dapat mengatasi objek yang memiliki nilai sangat besar yang dapat menyebabkan terjadinya penyimpangan pada distribusi data.

Metode ini sudah diimplementasikan dalam sejumlah penelitian seperti untuk mengelompokkan pola harian dari data runtun waktu mengenai sistem pompa panas sumber

tanah (Zhang, dkk., 2022). Secara keseluruhan, pengelompokan menggunakan DTW pada penelitian tersebut dapat mengurangi dimensi data dan memberikan pandangan berdasarkan data tentang bagaimana sistem memasok pendinginan dan pemanasan yang mana masing-masing diperoleh 6 dan 9 kelompok. Pada penelitian lain disebutkan bahwa *K-medoid* lebih tahan terhadap *outlier*, karena *medoid* tidak sensitif terhadap keberadaan *outlier* seperti misalnya *centroid* di *K-means* (Javed, dkk., 2020). Sementara itu, DTW memiliki hasil yang baik, tetapi membutuhkan waktu komputasi yang lebih lama terutama pada kumpulan data besar. Mengacu pada latar belakang tersebut, penulis memiliki ketertarikan untuk mengimplementasikan segmentasi provinsi menurut jarak DTW dan *K-Medoids* berdasarkan produksi tanaman pangan padi di Indonesia tahun 2013-2021.

## II. METODOLOGI

Analisis *clustering* yaitu proses mengelompokkan atau memartisi objek. Dalam analisis *clustering*, setiap anggota kelompok memiliki karakteristik atau kedekatan yang sama, atau kelompok yang sama memiliki tingkat homogenitas yang tinggi (Wei, 2019). Namun, kelompok yang dibentuk menunjukkan heterogenitas yang signifikan satu sama lain. Dalam analisis kelompok, metode pengelompokan didasarkan pada ukuran jarak dan kesamaan. Untuk mengukur seberapa mirip satu objek dengan yang lain, jarak *euclidean* dapat digunakan.

Metode pengelompokan tidak lagi relevan jika hanya mengandalkan jarak *euclidean* untuk mengukur kedekatan antar objek pada data yang fluktuatif dan berubah seiring waktu. Jarak *euclidean* lebih sesuai untuk data yang bersifat statis atau konsisten dalam jangka waktu tertentu. Namun, untuk membangun matriks kedekatan jarak antar objek pada data deret waktu, dibutuhkan pengukuran jarak yang mampu menangani sifat dinamis data tersebut. Oleh karena itu, *Dynamic Time Warping* (DTW) menjadi salah satu pendekatan yang dapat diterapkan.

DTW merupakan salah satu pengukuran ketidakmiripan menggunakan pendekatan tanpa

model. Algoritma ini membandingkan dua seri data dan menghitung lintasan optimum di antara keduanya. Ini memungkinkan perbandingan antar waktu dengan berbagai panjang periode (Munthe, 2019).

DTW menemukan seluruh kemungkinan jalur dan kemudian memilih jalur yang memberikan jarak terpendek antara dua deret waktu dengan menggunakan metode pemrograman dinamis. Nilai minimal dari tiga tetangga terdekat setiap elemen matriks digunakan untuk menghitung jarak kumulatif dari setiap elemen. DTW adalah generalisasi dari algoritma klasik yang melakukan perbandingan antara urutan diskrit dan urutan nilai kontinu.

Misalkan terdapat dua deret waktu, yaitu  $A = a_1, a_2, \dots, a_i, \dots, a_n$  dan  $B = b_1, b_2, \dots, b_j, \dots, b_m$ . Langkah pertama adalah membangun matriks dengan ukuran  $n \times m$ . Setiap elemen  $(i, j)$  dalam matriks ini merepresentasikan jarak kumulatif yang dihitung berdasarkan jarak antara elemen  $(i, j)$  dan nilai minimum dari tiga elemen yang berdekatan dengan elemen tersebut. Dengan kondisi  $0 < i \leq n$  dan  $0 < j \leq m$ , elemen  $(i, j)$  didefinisikan sebagaimana ditunjukkan dalam persamaan 1.

$$e_{ij} = d_{ij} + \min\{e_{(i-1)(j-1)}, e_{(i-1)j}, e_{i(j-1)}\} \dots \dots \dots (1)$$

dengan :  
 $d_{ij} = (a_i - b_j)^2$   
 $e_{ij}$  = elemen  $(i, j)$

Dalam menentukan *optimal path*, terlebih dahulu dihitung *path* yang dapat memberikan jarak kumulatif terkecil pada  $(n, m)$ . Jarak kumulatif terkecil pada  $(n, m)$  didefinisikan pada persamaan 2.

$$D_{DTW}(A, B) = \min_{w \in W} \left[ \sqrt{\sum_{l=1}^L e_{w_l}} \right] \dots \dots \dots (2)$$

dengan :  
 $w$  = elemen dari  $W$  yang ada pada matriks  $E = 1, 2, \dots, W$   
 $W$  = keseluruhan jalur *warping* yang mungkin  
 $L$  = *length* atau panjang dari jalur *warping*  
 $l$  = elemen dari  $L$  yang ada pada matriks  $E$

Berikut ini langkah-langkah yang dilaksanakan dalam analisis *clustering* dengan DTW:

**Pertama**, Menentukan matriks *cost*. Pada tahapan awal dihitung matriks *cost*  $(n \times m)$

yang merupakan selisih dari himpunan X dan Y. Setiap elemen matriks ditentukan menggunakan persamaan 3.

$$C_p = \begin{bmatrix} c_{1,1} & c_{1,2} & c_{1,3} & \dots & c_{1,j} \\ c_{2,1} & c_{2,2} & c_{2,3} & \dots & c_{2,j} \\ c_{3,1} & c_{3,2} & c_{3,3} & \dots & c_{3,j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ c_{i,1} & c_{i,2} & c_{i,3} & \dots & c_{i,j} \end{bmatrix} \dots \dots \dots (3)$$

dengan :

$$C_p \in R^{n \times m} : c_{i,j} = \|x_i - y_j\| \dots \dots \dots (4)$$

$i \in [1:n], j \in [1:m]$

Cost dari setiap matriks diperlukan untuk menentukan path dengan cost minimum dari kolom awal sampai terakhir.

**Kedua**, Menentukan jalur warping optimal.

Algoritma ini akan mengidentifikasi jalur warping dengan biaya terendah setelah matriks biaya terbentuk.  $p = \{p_1, p_2, \dots, p_k\}$   $p = \{p_1, p_2, \dots, p_k\}$  dengan  $p = \{p_i, p_j\} [1:n] \times [1:m]$ . Jalur warping yang dihitung harus memenuhi ketentuan berikut:

a. *Monotonicity*

Proses mengikuti urutan menurut series waktu  $n_1 \leq n_2 \leq \dots \leq n_k$  dan  $m_1 \leq m_2 \leq \dots \leq m_k$  agar proses dapat dijalankan straight forward sehingga tidak terjadi pengulangan pengolahan data yang sama berulang kali.

b. *Boundary Condition*

Merupakan awal dan akhir sekuen atau rangkaian data. Ini harus dilakukan.  $p_1 = (1, 1)$  dan  $p_k = (m, n)$ . Dengan demikian, data yang kita miliki diproses secara keseluruhan.

c. *Continuity*

Dengan kata lain, indeks data yang akan diproses selanjutnya tidak selalu bergerak jauh dari indeks sebelumnya. Jarak maksimum antara indeks awal dan indeks berikutnya hanya sebesar 1. Indeks data 2 – indeks data 1  $\leq 1$

Tujuannya adalah untuk menghindari data yang tidak diproses.

**Ketiga**, Menentukan matriks DTW

a. Baris pertama:

$$D(1, j) = \sum_{k=1}^j c(x_1, y_k), j \in [1, M]$$

b. Kolom pertama:

$$D(i, 1) = \sum_{k=1}^i c(x_k, y_1), i \in [1, N]$$

c. Semua elemen matriks:

$$D(i, j) = c(x_i, y_j) + \min \{D(i-1, j-1),$$

$$D(i-1, j), D(i, j-1)\}$$

$i \in 1, N, j \in 1, M$

**Keempat**, Menentukan cost optimal warping path. Setelah menghitung matriks DTW, gunakan persamaan 6 untuk menghitung total cost jarak yang melewati jalur warping:

$$\text{cost} = \frac{1}{m} \{ \min [DTW(1, m) \dots \dots \dots (5)$$

$$DTW(2, m), \dots, DTW(n, m) \}$$

Dengan cost warping path yang ideal, cost yang paling rendah terlihat pada kolom terakhir.

Selain menggunakan analisis clustering dengan DTW, penelitian ini menerapkan pendekatan lain berupa analisis clustering yaitu K-Medoid. K-Medoids Clustering merupakan metode pengelompokan non-hierarki, juga dikenal sebagai metode PAM (Partitioning Around Medoid). Metode ini menggunakan suatu objek representatif sebagai pusat kelompok (medoid). Dalam pengelompokan, medoid adalah objek di setiap cluster yang meminimalkan jumlah jarak dari medoid ke objek lain dalam cluster (Dietrich, dkk., 2015).

Keuntungan menggunakan PAM adalah “pusat” setiap cluster adalah objek sebenarnya dalam dataset. Dengan kata lain, suatu medoid adalah objek dalam kelompok yang memiliki jarak rata-rata ketidakmiripan yang paling kecil atau paling sedikit dari objek lain dalam kelompok tersebut. Metode ini menghasilkan k-cluster dari set data yang terdiri dari n objek dengan peneliti terlebih dahulu menentukan jumlah cluster yang terbentuk. Kelebihan metode k-medoids adalah dapat menangani masalah outlier dari distribusi data.

Dalam metode pengelompokan non-hierarkis, nilai Pseudo F dimanfaatkan untuk menentukan jumlah cluster. Nilai Pseudo F yang tertinggi menunjukkan jumlah cluster yang akan digunakan, di mana keragaman dalam setiap

kelompok cenderung homogen, sementara keragaman antar kelompok bersifat heterogen. Persamaan 6 menunjukkan rumus untuk menghitung nilai *Pseudo F*.

$$\text{Pseudo } F = \frac{(R^2/(k-1))}{\left(\frac{1-R^2}{n-k}\right)} \dots \dots \dots (6)$$

dengan :

$$R^2 = \frac{\text{SST}-\text{SSW}}{\text{SST}}$$

$$\text{SST} = \sum_{i=1}^{nk} \sum_{j=1}^k \sum_{v=1}^p (x_{ijv} - \bar{x}_v)^2$$

$$\text{SSW} = \sum_{i=1}^{nk} \sum_{j=1}^k \sum_{v=1}^p (x_{ijv} - \bar{x}_{jv})^2$$

Keterangan :

*SST* = total jumlah kuadrat total dari kuadrat jarak terhadap rata-rata

*SSW* = total jumlah kuadrat dari kuadrat jarak observasi terhadap rata-rata

*n* = banyak observasi

*k* = banyak kelompok

*p* = banyak variabel

*x<sub>ijv</sub>* = observasi ke-*j* pada kelompok ke-*k* variabel ke-*v*

$\bar{x}_v$  = rata-rata seluruh observasi variabel ke-*v*

$\bar{x}_{jv}$  = rata-rata observasi kelompok ke-*j* variabel ke-*v*

Data yang digunakan pada laporan kerja praktik menggunakan data sekunder berupa data tanaman pangan padi tahun 2013-2021. Data tanaman pangan dalam laporan ini berasal dari publikasi Badan Pusat Statistik (BPS) Indonesia tentang produksi tanaman pangan padi dalam ton. Data ini dikumpulkan melalui kegiatan survei pertanian (SP) dan survei ubinan.

### III. HASIL DAN PEMBAHASAN

Matriks DTW produksi tanaman pangan padi tahun 2013-2021 di Indonesia disajikan sebagai berikut:

$$D_{34 \times 34} = \begin{bmatrix} 0 & \dots & 28107874 \\ \vdots & \ddots & \vdots \\ 28107874 & \dots & 0 \end{bmatrix}$$

Berdasarkan produksi tanaman pangan padi, makin banyak kesamaan atau kemiripan karakteristik yang ada di antara dua provinsi menunjukkan bahwa mereka berada dalam kelompok yang sama. Sebaliknya, makin sedikit kesamaan atau kemiripan antar dua provinsi berdasarkan produksi tanaman pangan padinya

menunjukkan bahwa provinsi tersebut berada pada *cluster* yang berbeda.

Banyak *cluster* yang dibentuk mengacu pada nilai optimum *Pseudo F* yang disajikan pada Tabel 1 yang mana tampak bahwa nilai tertingginya berada pada *cluster* dengan tiga kelompok yang dijadikan dasar penentuan *cluster* dalam *K-Medoids Clustering*.

**Tabel 1.** Jumlah *Cluster* dengan *Pseudo F*

Jumlah <i>Cluster</i>	<i>Pseudo F</i>
2	23,07727
<b>3</b>	<b>27,00661</b>
4	23,85239
5	17,40134
6	13,48367
7	10,85188
8	9,04624
9	7,531907
10	6,428138

Setelah mendapatkan kelompok awal, yang terdiri dari tiga kelompok atau *cluster*, dilakukan pengulangan atau iterasi dengan menentukan pusat *cluster* baru secara acak dari setiap *cluster*. Iterasi dihentikan saat mendapatkan kelompok akhir, yang memiliki jarak paling sedikit antara *non-medoids* dan pusat *cluster* keseluruhan, dan pusat *cluster (medoid)* tidak berubah. Analisis menunjukkan bahwa karena hasil iterasi kedua tidak berubah, iterasi pertama dihentikan, sehingga hasil pengelompokan berhasil didapatkan.

Tabel 2 menunjukkan hasil pengelompokan provinsi di Indonesia berdasarkan produksi tanaman pangan padi tahun 2013-2021. Hasil analisis menunjukkan bahwa pusat kelompok 1 berada di Provinsi Riau, pusat kelompok 2 berada di Provinsi Jawa Timur, dan pusat kelompok 3 berada di Provinsi Sumatera Barat.

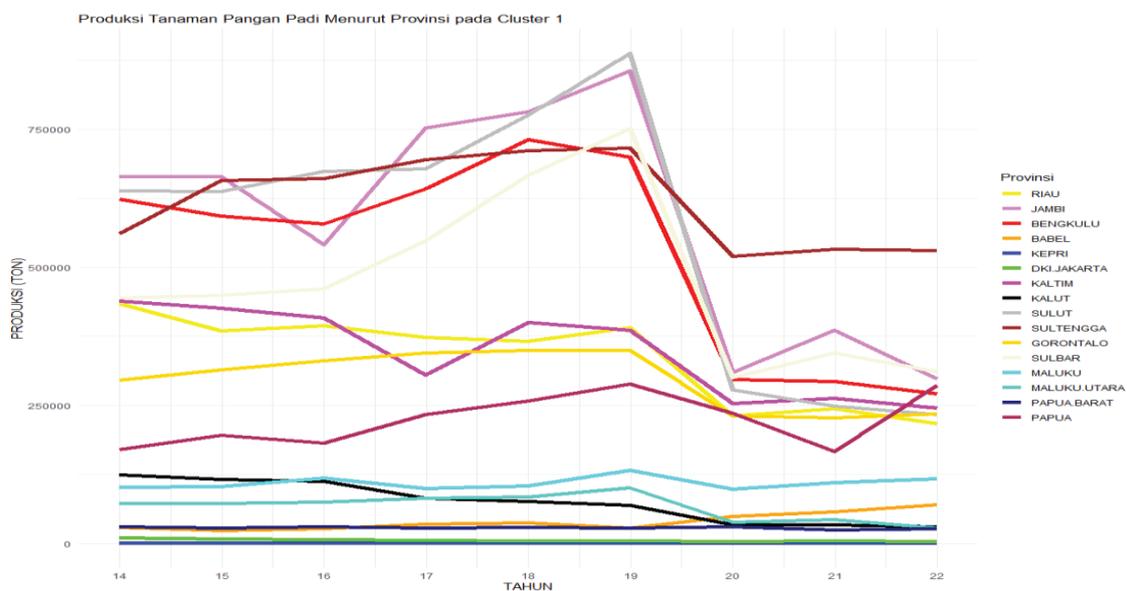
Dalam rangka menjamin keberlanjutan produksi padi dan stabilitas ekonomi sektor pertanian, implementasi kebijakan berbasis data seperti ini sangat penting dilakukan yaitu dengan adanya kebijakan sesuai kriteria masing-masing wilayah berdasarkan hasil *clustering*. Masing-masing kelompok digambarkan dalam plot pada Gambar 1 yang menunjukkan plot pergerakan

**Tabel 2.** Hasil *clustering* dengan 3 *cluster*

Kelompok	Provinsi
1	“Riau, Jambi, Bengkulu, Kepulauan Bangka Belitung, Kepulauan Riau, DKI Jakarta, Kalimantan Timur, Kalimantan Utara, Sulawesi Utara, Sulawesi Tenggara, Gorontalo, Sulawesi Barat, Maluku, Maluku Utara, Papua dan Papua Barat”
2	“Sumatera Utara, Sumatera Selatan, Lampung, Jawa Barat, Jawa Tengah, Jawa Timur dan Sulawesi Selatan”
3	“Aceh, Sumatera Barat, DI Yogyakarta, Banten, Bali, Nusa Tenggara Barat, Nusa Tenggara Timur, Kalimantan Barat, Kalimantan Tengah, Kalimantan Selatan dan Sulawesi Tengah”

produksi tanaman pangan padi dalam ton di 16 provinsi Indonesia yang termasuk dalam kelompok 1. Ke-16 provinsi ini digabungkan ke dalam satu kelompok karena memiliki karakteristik yang mirip atau serupa.

seperti perubahan iklim atau bencana alam. Adanya ketergantungan akan cuaca dan peluang terjadinya anomali iklim pada wilayah-wilayah di *cluster* 1 menjadi landasan kuat perlunya menyusun strategi mitigasi risiko



**Gambar 1.** Grafik *cluster* 1 pada Produksi Tanaman Pangan Padi di Indonesia

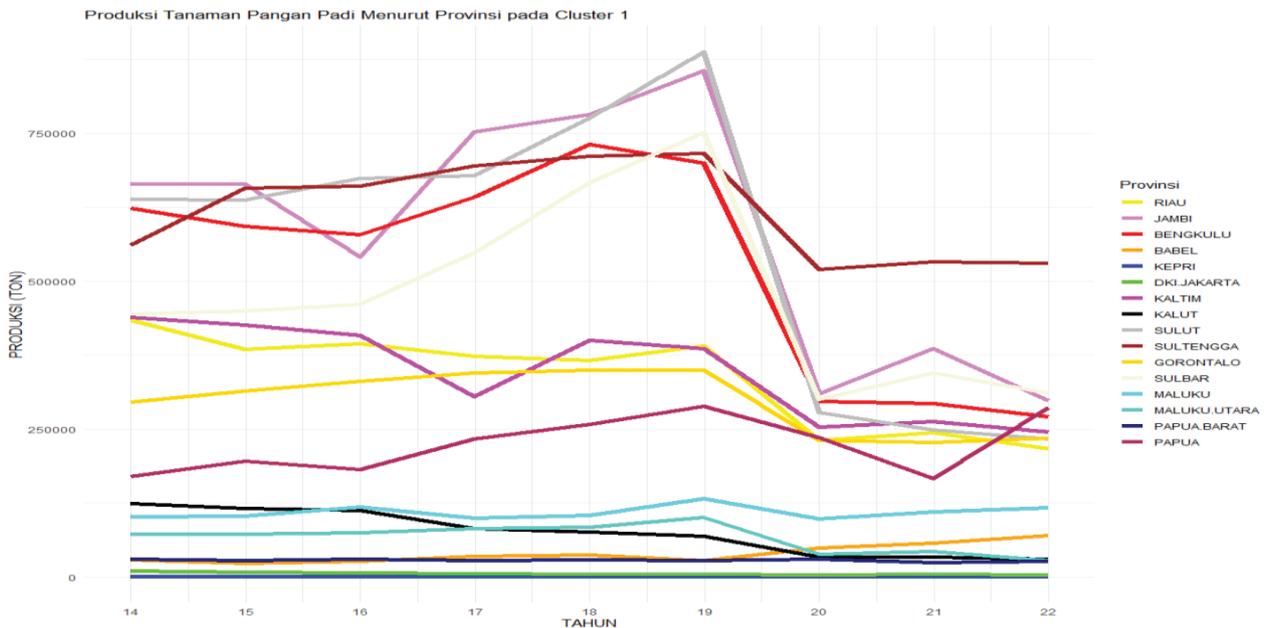
*Cluster* 1 menunjukkan variasi produksi tanaman pangan padi yang signifikan, mengindikasikan provinsi-provinsi ini memiliki potensi tinggi dalam produksi padi. Meskipun demikian, wilayah dalam kelompok ini tetap memerlukan perhatian dari pemerintah untuk mengoptimalkan hasil produksi padi. Berdasarkan data sempat terjadi penurunan ekstrem di provinsi Sulawesi Utara, Bengkulu dan Jambi pada tahun 2018-2019 menunjukkan perlunya intervensi kebijakan yang adaptif. Penurunan produksi padi pada wilayah tersebut dimungkinkan pula terjadi akibat faktor eksternal

untuk mengatasi fluktuasi produksi padi di masa depan.

Inovasi dari sisi teknologi seperti sistem *Climate-Smart Agriculture* (Konfo, dkk., 2024) dapat diimplementasikan pada *cluster* 1 untuk memastikan produktivitas meskipun terjadi perubahan iklim dan tuntutan pasar yang beragam, serta tentunya untuk memastikan ketahanan pangan. Menurut laporan BULOG dan data SDGS (Kementerian PPN/Bappenas, 2021; Perum BULOG, 2020) ketahanan pangan adalah salah satu tujuan utama dalam pembangunan berkelanjutan. Dengan demikian,

segmentasi wilayah dapat membantu dalam pencapaian target SDGs terkait ketahanan pangan di Indonesia.

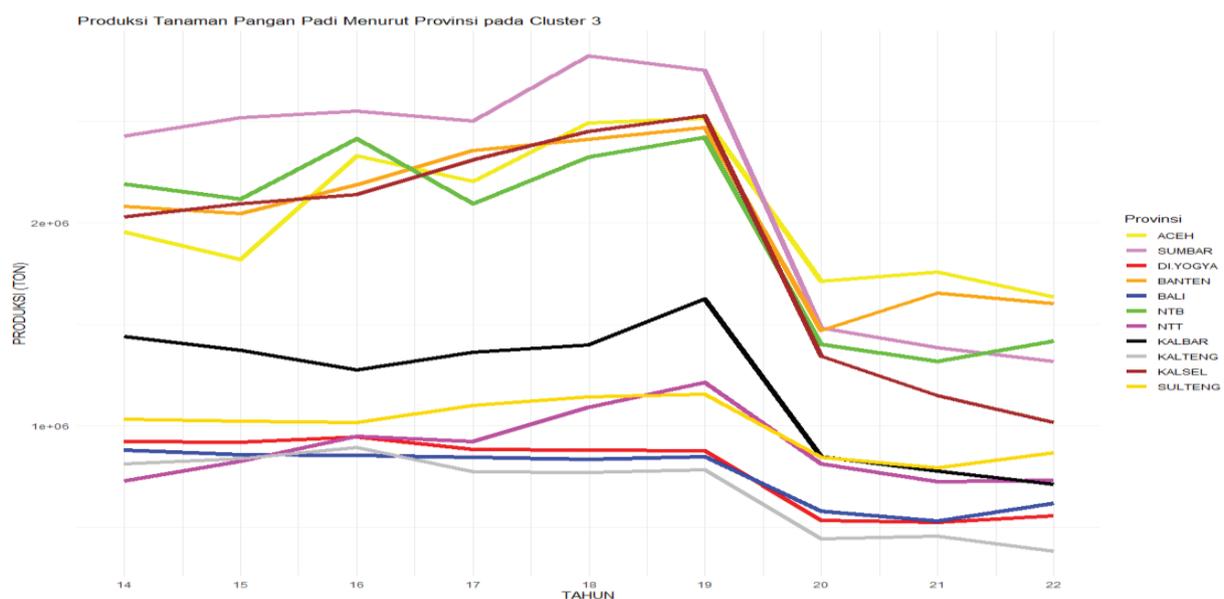
dalam negeri sebelum mempertimbangkan ekspor. Kebijakan seperti irigasi yang efektif dan teknologi pertanian modern yang dapat



**Gambar 2.** Grafik *cluster 2* pada Produksi Tanaman Pangan Padi di Indonesia

Pada Gambar 2, terlihat tujuh provinsi yang memiliki karakteristik yang mirip terkait dengan produksi tanaman pangan padi di daerahnya. *Cluster 2* adalah *cluster* yang mengalami kemajuan dalam produksi tanaman pangan padi, dan perlu mendapat perhatian khusus dari pemerintah, terutama dalam hal penyediaan fasilitas untuk meningkatkan produksi padi sehingga dapat memenuhi kebutuhan di

meningkatkan produktivitas pertanian secara signifikan merupakan salah satu kebijakan yang direkomendasikan berdasarkan hasil segmentasi wilayah. Kebijakan yang tepat untuk meningkatkan produktivitas tanaman pangan padi pada *cluster 2* tidak hanya berdampak pada dukungan provinsi dalam ketahanan pangan nasional, tetapi juga turut



**Gambar 3.** Grafik *Cluster 3* pada Produksi Tanaman Pangan Padi di Indonesia

mendorong perbaikan posisi ekonomi Indonesia di pasar global. Penerapan teknologi produksi pangan berkelanjutan dapat pula menjadi solusi dengan menerapkan *digital agriculture* karena memberikan pengamatan yang akurat dan *real time* kepada produsen pangan mengenai berbagai kondisi yang memengaruhi produktivitas, di antaranya kesehatan tanaman, kualitas tanah, kondisi cuaca, dan tekanan hama penyakit (Abiri, dkk., 2023).

Sebelas provinsi lainnya dikelompokkan ke dalam satu *cluster*, yaitu *cluster 3*. Pada Gambar 3, ditampilkan plot pergerakan produksi tanaman pangan padi di provinsi-provinsi yang termasuk dalam *cluster 3*. Secara visual, pola pergerakan produksi terlihat cukup seirama, yang mana seluruh provinsi menunjukkan kenaikan dan penurunan produksi tanaman pangan padi dalam rentang waktu yang relatif berdekatan. Pada tahun 2018 menuju 2019, produksi tanaman pangan padi mengalami penurunan pada semua provinsi di Indonesia. Hal ini bisa disebabkan karena pada tahun tersebut negara Indonesia dilanda oleh Covid-19 yang menyebabkan para petani di Indonesia mengharuskan untuk berhenti terlebih dahulu dalam memproduksi tanaman pangan padi. Namun, mengalami kenaikan pada tahun-tahun selanjutnya walaupun secara perlahan.

Oleh karena *cluster 3* merupakan provinsi-provinsi dengan produksi paling rendah dibandingkan dengan dua *cluster* lainnya, maka wilayah ini memerlukan perhatian yang khusus dari pemerintah baik di tingkat pusat maupun daerah. Dukungan finansial dan subsidi untuk benih, pupuk, dan peralatan pertanian dapat membantu akselerasi pemulihan produksi pasca pandemi dengan lebih cepat. Kerja sama dengan berbagai pihak, baik industri maupun akademisi untuk menyediakan program pelatihan dan penyuluhan bagi petani juga membantu peningkatan efisiensi produksi dan adaptasi terhadap perubahan kondisi lingkungan.

Pada hasil *cluster*, mayoritas produksi padi terbanyak terdapat pada daerah pulau Jawa. Hal ini bisa terjadi karena pulau Jawa memiliki letak geografis yang lebih memadai, lembab dan cukup sinar matahari.

#### IV. KESIMPULAN

Hasil analisis kluster menggunakan metode DTW menghasilkan tiga kluster dengan nilai *Pseudo F* tertinggi. Masing-masing kluster menunjukkan karakteristik yang berbeda dalam produksi tanaman pangan padi, namun terdapat kesamaan karakteristik di antara anggota dalam kluster yang sama. Pada metode DTW, perhitungan dilakukan dengan membandingkan data dari awal hingga akhir periode, sehingga memungkinkan identifikasi perbedaan pola yang signifikan, tetapi tetap mempertahankan jarak terkecil yang mungkin. Kluster 1, 2, dan 3 dibentuk berdasarkan variasi fluktuasi produksi tanaman pangan padi.

#### UCAPAN TERIMA KASIH

Kami menyampaikan apresiasi kepada Badan Pusat Statistik sebagai penyedia data, serta kepada Fungsi Statistik Sosial BPS Provinsi Riau atas dukungan yang diberikan selama proses diskusi dalam penelitian ini.

#### DAFTAR PUSTAKA

- Abiri, R., N. Rizan, S. K. Balasundram, A. B. Shahbazi, and H. Abdul-Hamid. 2023. Application of Digital Technologies for Ensuring Agricultural Productivity. *Heliyon*, 9 (12):e22601. <https://doi.org/10.1016/j.heliyon.2023.e22601>.
- Badan Pusat Statistik Indonesia. 2022. *Keadaan Ketenagakerjaan Indonesia Agustus 2022*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistik Indonesia. 2023a. *Luas Panen Dan Produksi Padi Di Indonesia 2022*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistik Indonesia. 2023b. *Pertumbuhan Ekonomi Indonesia Triwulan IV 2022*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistik Indonesia. 2023c. *Ringkasan Eksekutif Luas Panen Dan Produksi Padi Di Indonesia 2022*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistik Indonesia. 2023d. *Ringkasan Eksekutif Luas Panen Dan Produksi Padi Di Indonesia 2023 (Angka Sementara)*. Jakarta: Badan Pusat Statistik Indonesia.
- Badan Pusat Statistik Provinsi Riau. 2023. *Luas Panen Dan Produksi Padi Di Provinsi Riau 2022*. Pekanbaru: Badan Pusat Statistik Provinsi Riau. <https://lampung.bps.go.id/indicator/53/568/1/padi-produks-.html>.
- Dietrich, D., B. Heller, and B. Yang. 2015. *Data Science and Big Data Analytics*. Canada: John Wiley & Sons, Inc.

- Hayati, M., Elfiana, dan Martina. 2017. Peranan Sektor Pertanian Dalam Pembangunan Wilayah Kabupaten Bireuen Provinsi Aceh. *Jurnal Sains Pertanian*, 1 (3): 213–22.
- Hill, R.C., W.E. Griffiths, and G.C. Lim. 2011. *Principles of Econometrics*. 4th ed. United States of America: John Wiley & Sons, Inc.
- Imani, N., A.I. Alfassa, dan A.M. Yolanda. 2022. Self Organizing Map (Som) *Clustering* Untuk Analisis Data Indikator Sosial Di Provinsi Nusa Tenggara Timur. *Jurnal Gaussian*, 11 (3):458–67. <https://doi.org/10.14710/J.GAUSS.11.3.458-467>.
- Javed, A., B.S. Lee, and D.M. Rizzo. 2020. A Benchmark Study on Time Series *Clustering*. *Machine Learning with Applications* 1: 100001. <https://doi.org/10.1016/j.mlwa.2020.100001>.
- Kementerian PPN/Bappenas. 2021. Laporan Pelaksanaan Pencapaian TPB/SDGs Indonesia Tahun 2021. *Laporan Kementerian PPN/Bappenas*. Jakarta. <https://sdgs.bappenas.go.id/sekilas-sdgs/>.
- Konfo, T.R,C, A.B.P Chabi, A.A. Gero, C. Lagnika, F. Avlessi, G. Biaou, and C.K.D. Sohounhloue. 2024. Recent Climate-Smart Innovations in Agrifood to Enhance Producer Incomes through Sustainable Solutions. *Journal of Agriculture and Food Research*, 15(January). <https://doi.org/10.1016/j.jafr.2024.100985>.
- Montgomery, D.C., C.L. Jennings, and M. Kulahci. 2008. *Introduction to Time Series Analysis and Forecasting*. New Jersey (US): y John Wiley & Sons. Inc.
- Munthe, A.D. 2019. Penerapan *Clustering* Time Series Untuk Menggerombolkan Provinsi Di Indonesia Berdasarkan Nilai Produksi Padi. *Litbang Sukowati*, 2 (2): 1–11.
- Perum BULOG. 2020. *Memperkuat Ketahanan Pangan Di Masa Penuh Tantangan (Annual Report 2020)*. [www.bulog.co.id](http://www.bulog.co.id).
- Wei, W.W.S. 2019. *Multivariate Time Series Analysis and Applications*. United States of America: John Wiley & Sons, Ltd.
- Wicaksono, A.S, dan A.M Yolanda. 2021. Pengelompokan Kabupaten / Kota Di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Indeks Pembangunan Manusia Menggunakan K-Medoids *Clustering*. *JSTAR: Jurnal Statistika Terapan*, 1 (1): 79–90. <https://doi.org/10.5300/jstar.v1i1.7>.
- Yolanda, A.M, dan K. Yunitaningtyas. 2021. Segmentasi Provinsi Berdasarkan Sarana Dan Perlengkapan Faskes Keluarga Berencana Tahun 2021. *Jurnal Keluarga Berencana*, 6 (1): 20–30. <https://doi.org/10.37306/kkb.v6i1.70>.
- Yunitaningtyas, K., dan A.M. Yolanda. 2022. Klasifikasi Kabupaten/Kota Di Provinsi Nusa Tenggara Timur Berdasarkan Indikator Status Kesehatan Masyarakat. *JSTAR: Jurnal Statistika Terapan*, 2 (1): 1–18. <https://doi.org/10.5300/JSTAR.V2I01.10>.
- Zhang, S. L. Zhang, and X. Zhang. 2022. *Clustering* Based on Dynamic Time Warping to Extract Typical Daily Patterns from Long-Term Operation Data of a Ground Source Heat Pump System. *Energy*, 249 (123767). <https://doi.org/https://doi.org/10.1016/j.energy.2022.123767>.

#### BIODATA PENULIS :

**Anne Mudya Yolanda** dilahirkan di Padang, 23 Juni 1995. Penulis menyelesaikan pendidikan S1 Statiska di Universitas Islam Indonesia tahun 2016 dan S2 Statiska Terapan Institut Pertanian Bogor tahun 2019.

**Husna Savira Yolanda** dilahirkan di Tanjung Balai Karimun, 19 Februari 2001. Penulis menyelesaikan pendidikan S1 d Program Studi Statistika Universitas Riau tahun 2023.