E-ISSN : 2962-9713
P-ISSN : 2963-928X

Analisis Perbandingan Kinerja *Spectral Biclustering* Dan *Plaid Model* Dalam Mengidentifikasi Pola Ketahanan Pangan Di Indonesia

Dwi Fania Nalole[1], Fahrezal Zubedi[2], Nisky Imansyah Yahya[3]
[1]Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Gorontalo, dwifanian@e-mail.com
[2]Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Gorontalo fahrezal@ung.ac.id.com
[3]Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Gorontalo, nisky@ung.ac.id.com

Abstrak

Ketahanan pangan merupakan isu multidimensional yang dipengaruhi oleh berbagai faktor, seperti produksi pertanian, kondisi sosial ekonomi, infrastruktur, dan lingkungan. Kompleksitas hubungan antar faktor tersebut memerlukan metode yang mampu mengidentifikasi pola antarwilayah dan indikator secara simultan. Penelitian ini bertujuan membandingkan metode *Plaid Model* dan *Spectral Biclustering* dalam mengidentifikasi pola ketahanan pangan di Indonesia. Data yang digunakan merupakan data sekunder tahun 2024 yang mencakup 38 provinsi dengan tujuh indikator ketahanan pangan. Analisis dilakukan melalui tahapan *preprocessing* data, standarisasi, pembentukan *bicluster,* serta evaluasi hasil *biclustering*. Hasil penelitian menunjukkan bahwa kedua metode mampu mengidentifikasi pola ketahanan pangan berdasarkan karakteristik provinsi dan indikator yang dianalisis. *Spectral Biclustering* menghasilkan pola pengelompokan yang menggambarkan aspek produksi pertanian, sosial ekonomi, lingkungan, kesejahteraan, dan infrastruktur, sedangkan *Plaid Model* menghasilkan pola yang lebih spesifik melalui *bicluster* yang saling tumpang tindih. Berdasarkan hasil evaluasi, *Plaid Model* memberikan kualitas pengelompokan yang lebih baik sehingga lebih optimal dalam mengidentifikasi pola ketahanan pangan di Indonesia.

Kata Kunci: *Ketahanan Pangan, Biclustering, Plaid Model, Spectral Biclustering.*

Abstract

Food security is a multidimensional issue influenced by various factors, including agricultural production, socioeconomic conditions, Infrastructure, and the environment. The complexity of the relationships among these factors requires a method that can simultaneously identify patterns across regions and indicators. This research aims to compare the Plaid Model and Spectral Biclustering methods in identifying food security patterns in Indonesia. The data used consists of secondary data from 2024 covering 38 provinces and 7 food security indicators. The analysis is conducted through data preprocessing, data standardization, bicluster formation, and evaluation of biclustering results. The results show that both methods can identify food security patterns based on provincial characteristics and the indicators analyzed. Spectral Biclustering produces clustering patterns that represent aspects of agricultural production, socioeconomic conditions, the environment, welfare, and infrastructure, while the Plaid Model produces more specific patterns through overlapping biclusters. Based on the evaluation results, the Plaid Model provides better clustering quality, making it more effective in identifying food security patterns in Indonesia.

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Keywords: *FoddSecuriti, Biclustering, PlaidModel,SpectralBiclustering*

Corresponding Author:
Dwi Fania Nalole
Universitas Negeri Gorontalo ; dwifanian@e-mail.com

PENDAHULUAN

Ketahanan pangan merupakan isu global yang berkaitan dengan kesehatan masyarakat, kondisi ekonomi, dan stabilitas suatu negara. Ketahanan pangan terwujud ketika semua orang, setiap saat, memiliki akses fisik dan ekonomi terhadap pangan yang cukup, aman, dan bergizi untuk memenuhi kebutuhan gizi dan preferensi pangan mereka (Insecurity & Insecurity, 1996). Menurut laporan FAO (2023) dalam *The State of Food Security and Nutrition in the World*, sekitar 783 juta orang di dunia masih menghadapi kelaparan pada tahun 2022. Di Indonesia, ketahanan pangan juga menghadapi berbagai tantangan yang kompleks. Berdasarkan data *Global Food Security Index* (GFSI) tahun 2022, Indonesia menempati peringkat 63 dari 113 negara dengan skor 60,2.

Ketahanan pangan di Indonesia bersifat multidimensi dan mencakup ketersediaan, distribusi, konsumsi, serta status gizi masyarakat (Hestina, 2011). Kondisi antarwilayah juga menunjukkan karakteristik yang berbeda, sehingga diperlukan pendekatan analisis yang dapat menggambarkan hubungan antara wilayah dan berbagai indikator ketahanan pangan secara simultan. Penelitian sebelumnya telah menggunakan metode *K-Mean* dan *Hierarchical Clustering* untuk menganalisis dan memetakan ketahanan pangan di Indonesia. Namun, metode *clustering* konvensional umumnya mengelompokkan objek berdasarkan seluruh variabel secara bersamaan sehingga belum secara khusus mengidentifikasi keterkaitan antara kelompok wilayah dan kelompok indikator.

Salah satu pendekatan yang dapat digunakan untuk mengatasi keterbatasan tersebut adalah *biclustering*, yaitu metode yang mengelompokkan baris dan kolom matriks data secara simultan (Castanho et al., 2024). Pendekatan ini memungkinkan identifikasi submatriks yang memiliki pola tertentu pada kelompok objek dan variabel secara bersamaan (Patowary et al., 2020). Dengan demikian, *biclustering* dapat memberikan gambaran yang lebih spesifik mengenai keterkaitan antarprovinsi dan antarindikator ketahanan pangan.

Dua metode *biclustering* yang memiliki karakteristik berbeda adalah *Plaid Model* dan *Spectral Biclustering*. *Plaid Model* yang dikembangkan oleh Lazzeroni dan Owen (2002) membangun bicluster secara berlapis sehingga memungkinkan terbentuknya pola yang saling tumpang tindih (*overlapping*) (Majd et al., 2016). Sementara itu, *Spectral Biclustering* yang diperkenalkan oleh Kluger et al. (2003) menggunakan pendekatan spektral dan *Singular Value Decomposition* (SVD) untuk menemukan pola hubungan antara baris dan kolom dalam data (Kluger et al., 2003). Penelitian Oghabian et al. (2014) menunjukkan bahwa *Plaid Model* mampu mengidentifikasi pola lokal dan tumpang tindih, sedangkan Totti dan Setiyawati (2022) menunjukkan bahwa *Spectral Biclustering* mampu menghasilkan pengelompokan yang baik pada data yang dianalisis.

Berdasarkan permasalahan tersebut, penelitian ini bertujuan membandingkan *Plaid Model* dan *Spectral Biclustering* dalam mengidentifikasi pola ketahanan pangan

E-ISSN : 2962-9713
P-ISSN : 2963-928X

di Indonesia. Analisis dilakukan dengan membentuk *bicluster* berdasarkan 38 provinsi dan tujuh indikator ketahanan pangan tahun 2024, kemudian mengevaluasi kualitas *bicluster* yang dihasilkan oleh kedua metode. Hasil penelitian diharapkan dapat memberikan gambaran mengenai pola keterkaitan antara provinsi dan indikator ketahanan pangan serta menjadi informasi pendukung dalam memahami karakteristik ketahanan pangan antarwilayah di Indonesia.

METODE

Penilitian ini menggunakan pendekatan kuantitaitf dengan data sekunder yang bersumber dari publikasi Badan Pusat Statistik Republik Indonesia tahun 2025, data tahun 2024, serta Badan Pangan Nasional (Bapanas) 2024. Mencakup 38 provinsi sebagai objek penilitian dan tujuh indikator ketahanan pangan sebagai variabel, yaitu luas panen padi, produksi padi, produksi jagung, presentasi penduduk miskin, angka harapan hidup, rumah tangga dengan akses utama dari listrik, dan curah hujan rata-rata tahunan. Seluruh provinsi dijadikan sampel melalui teknik sampling jenuh. Analisis dilakukan menggunakan bahasa pemograman Phyton untuk menerapkan metode *Spectral Biclustering* dan *Plaid Model*, dan QGIS untuk memvisualisasikan pengelompokkan wilayah.

Prepocesing Data

Preprocessing data bertujuan untuk menyiapkan data yang biasanya tidak sempurna. Penangan missing value dilakukan menggunakan metode *mean imputation*, yaitu nilai yang hilang ditangani dengan mengganti data numerik diisi dengan rata-rata dari variabel terkait (Alwateer2024MissingDataImputation). Proses penangan missing value melibatkan persamaan sebagai berikut:

$$x_{ij} = \sum_{i:x_{ij\epsilon}C_k} \frac{x_{ij}}{n_k} \quad (1)$$

Keterangan :

x_{ij} : Nilai hasil imputasi untuk observasi ke- i pada variabel ke- j.
x_{ij} : Nilai asli observasi ke-i pada variabel ke- j.
C_k : Kelompok atau cluster ke- k.
$\{i: x_{ij} \in C_k\}$: Himpunan indeks i yang termasuk dalam cluster C_k.
$\sum_{i:x_{ij\epsilon}C_k}$: Penjumlahan seluruh nilai xij yang berada pada cluster C_k.
n_k : Jumlah anggota (observasi) dalam cluster C_k.

Standarisasi Data

Standarisasi data merupakan salah satu tahap penting dalam proses *preprocessing* yang bertujuan untuk menyetarakan skala antar variabel sehingga tidak ada satu atribut yang mendominasi hasil analisis. *Z-score standardization* digunakan agar setiap variabel memiliki kontribusi yang seimbang terhadap hasil analisis. Proses standarisasi dapat menggunakan persamaan berikut:

$$Z = \frac{x-\mu}{\mu} \quad (2)$$

Keterangan :

Z : Nilai asli data.
$x - \mu$: Nilai rata-rata.
μ : Simpangan baku.

Plaid Model

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Plaid Model adalah salah satu pendekatan *biclustering* yang diperkenalkan oleh Lazzeroni dan Owen pada tahun 2002 untuk menganalisis data dalam bentuk matriks, khususnya data ekspresi gen, dengan tujuan untuk mengidentifikasi pola-pola lokal yang tidak terlihat. Menurut Siswantining, Aminanto, Sarwinda, dan Swasti (2021), *Plaid Model* merupakan salah satu pendekatan dalam analisis *biclustering* yang bersifat *probabilistik* dan memungkinkan terjadinya *overlapping bicluster*, yaitu kondisi di mana baris dan kolom data dapat tergabung ke dalam lebih dari satu kelompok secara bersamaan (Siswantining & Swasti, 2021).

$$Y_{ij} = \mu_0 + \sum_{k=1}^{K}(\mu_k + \alpha_{ik} + \beta_{jk})\rho_{ik}K_{jk} \quad (3)$$

Keterangan :

Y_{ij} : Menyatakan nilai data pada baris ke-i dan kolom ke- j dalam matriks data.

$\sum_{k=1}^{K}$:Menunjukkan penjumlahan kontribusi dari bicluster ke – 1 hingga ke – K, dengan K sebagai jumlah total *bicluster* yang di bentuk.

μ_k : Merupakan nilai rata-rata keseluruhan (*Grand Mean*) dari seluruh data.

α_{ik} : Merupakan efek baris pada observasi ke - i dalam *bicluster* ke- k.

β_{jk} : Merupakan efek kolom pada variabel ke- j dalam bicluster ke- k.

ρ_{ik} : Merupakan indikator keanggotaan baris ke- i pada bicluster ke- k.

K_{jk} : Merupakan indikator keanggotaan kolom ke- j pada bicluster ke- k, di mana keduanya bernilai 1 jika termasuk dalam bicluster dan 0 jika tidak.

Menurut (Siswantining & Swasti, 2021), penerapan metode *Plaid Model* dalam analisis *biclustering* dilakukan melalui beberapa tahapan sistematis berikut:

1. Pembentukan Model Awal dan Matriks Residual
   Menyusun model awal yang memuat efek rata-rata keseluruhan, efek baris, dan efek kolom, kemudian menghitung matriks residual sebagai selisih antara data asli dengan model awal tersebut.
2. Penentuan Kandidat Awal Bicluster
   Menentukan keanggotaan awal baris dan kolom untuk bicluster, yang dapat diperoleh secara acak atau melalui metode *two-way clustering* seperti *k-means.*
3. Estimasi Parameter Model
   Mengestimasi parameter efek *bicluster* (seperti rata-rata dan efek baris/kolom) dengan cara meminimalkan selisih kuadrat antara data dan hasil model.
4. Estimasi Keanggotaan *Bicluster*
   Memperbarui keanggotaan baris dan kolom menggunakan metode *binary least squares* (0 atau 1), untuk menentukan apakah suatu elemen masih termasuk dalam bicluster.
5. Pruning (Pemangkasan)
   Menghapus baris atau kolom yang tidak sesuai dengan pola bicluster berdasarkan nilai ambang batas (*threshold*) tertentu agar pola yang terbentuk lebih homogen.
6. Pembaruan Residual dan Iterasi Lanjut
   Melakukan *back-fitting* dengan mengurangi pengaruh bicluster yang telah ditemukan dari data, kemudian mengulangi proses pencarian bicluster baru pada residual yang diperbarui.
7. Evaluasi Hasil *Biclustering*
   Mengevaluasi hasil bicluster untuk menilai keseragaman serta kestabilan pola yang terbentuk.
8. Interpretasi Hasil

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Menafsirkan setiap *bicluster* sebagai kelompok baris dan kolom yang memiliki pola serupa, sesuai konteks data yang dianalisis.

Spectral Biclustering

Spectral Biclustering merupakan salah satu pendekatan dalam *biclustering* yang memanfaatkan teknik aljabar linier dan analisis spektral (eigenvalue/eigenvector) untuk menemukan struktur submatriks *bicluster* dalam suatu matriks baris dan kolom. Konsep utama dari *Spectral Biclustering* adalah bahwa jika sebuah matriks menunjukkan struktur *checkerboard-like* (yaitu blok-blok bicluster yang tersembunyi) setelah dilakukan pengacakan pada baris serta kolom, maka pola tersebut dapat diidentifikasi melalui dekomposisi *spektral* dari matriks yang telah dinormalisasi. *Spectral Biclustering* termasuk ke dalam kategori *simultaneous clustering* karena secara bersamaan menemukan kelompok baris dan kolom dengan menggunakan analisis spektral untuk mengidentifikasi struktur korelasi lokal dalam data (Henriques & Madeira, 2016). Menurut Klager et al., (2003), penerapan metode *Spectral Biclustering* dalam analisis *biclustering* dilakukan melalui beberapa tahapan sistematis berikut :

1. Representasi Data
   Data dinyatakan dalam bentuk matriks, dengan baris sebagai gen dan kolom sebagai kondisi (sampel).
2. Normalisasi Data
   Dilakukan normalisasi baris dan kolom secara iteratif agar setiap baris dan kolom memiliki rata-rata yang sama, sehingga bias ekspresi dapat dihilangkan.
3. Transformasi Spektral (SVD)
   Matriks hasil normalisasi didekomposisi menggunakan *Singular Value Decomposition* (SVD) untuk memperoleh eigenvector gen dan kondisi yang merepresentasikan pola utama.
4. Identifikasi Pola *Checkerboard*
   Pola *bicluster* muncul sebagai struktur *checkerboard* dari hasil perkalian antara *eigenvector* baris dan kolom.
5. Pemilihan dan Clustering
   Eigenvector dengan nilai eigen terbesar digunakan untuk membentuk klaster baris dan kolom melalui metode seperti *k-means* atau *normalized cut*.
6. Evaluasi dan Interpretasi
   Hasil bicluster dievaluasi berdasarkan keseragaman nilai dalam blok, untuk menilai seberapa homogen atau mirip nilai-nilai data di dalam setiap blok *bicluster*.

Evaluasi Kinerja Metode

Kinerja algoritma *biclustering* dapat di evaluasi menggunakan *Means Squared Residu* (MSR), MSR berfungsi untuk mengukur tingkat keseragaman internal suatu *bicluster* dengan menghitung variasi antar elemen di dalam submatriks, di mana nilai MSR yang rendah menunjukkan pola yang lebih seragam dan berkualitas tinggi (pontes2015biclustering). Persamaan MSR dapat di definisikan pada persamaan berikut :

$$MSR = (\beta)\ \frac{1}{|I| . |J|} \sum_{i=1}^{|I|} \sum_{i=1}^{|J|} (\ b_{ij} - b_{iJ} - b_{Ij} + b_{IJ})^2 \qquad (4)$$

Keterangan :

b_{ij} : Nilai elemen pada baris i dan kolom j.
b_{iJ} : Rata - rata baris ke i.
b_{Ij} : Rata – rata kolom ke j.
b_{IJ} : Rata-rata seluruh elemen dalam *bicluster*.

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Selain menggunakan *Mean Squared Residue* (MSR), kualitas hasil *biclustering* dapat dievaluasi menggunakan *Average Residue* (ASR). ASR digunakan untuk mengukur kualitas keseluruhan hasil *biclustering* berdasarkan rata-rata nilai MSR dari seluruh *bicluster* yang terbentuk. (Lee., 2009) Rumus ASR ditunjukkan pada persaman berikut :

$$ASR = \frac{1}{K} \sum_{k=1}^{K} \text{MSR } (\beta_k) \quad (5)$$

Keterangan :

ASR : *Average Residue Score.*
K : Jumlah total *bicluster* yang dihasilkan.
MSR (β_k) : nilai *Mean Squared Residue* pada bicluster ke- K.
β_k : *bicluster* ke- K yang terbentuk.
$\sum_{k=1}^{K}$ MSR (β_k) : Jumlah nilai MSR dari seluruh *bicluster* .

Evaluasi kualitas hasil juga dilengkapi dengan ukuran *variance* (VAR) , yang berfungsi untuk mengukur tingkat penyebaran nilai elemen di dalam suatu *bicluster*. Semakin homogen suatu *bicluster*, maka pola hubungan antara objek dan variabel yang terbentuk akan semakin baik dan konsisten. Salah satu metrik yang umum digunakan untuk mengukur homogenitas bicluster adalah *Variance* (VAR) (Castanho et al., 2022). Secara matematis, nilai VAR dirumuskan sebagai berikut :

$$VAR\ (B)\ \frac{1}{|I|\ |J|} \sum_{i=1}^{|I|}\ \sum_{i=1}^{|J|} (\quad b_{ij} \quad - \quad b_{IJ})^2 \quad (6)$$

Keterangan :

$VAR\ (B)$: Nilai *variance* pada *bicluster.*
$|I|$: Jumlah baris pada *bicluster.*
$|J|$: Jumlah kolom pada *bicluster.*
b_{ij} : Nilai elemen pada baris ke-i dan kolom ke-j.
b_{IJ} : Rata-rata seluruh elemen dalam *bicluster*.

Selain ukuran evaluasi berbasis homogenitas seperti *Mean Squared Residue* (MSR), *Average Residue* (ASR), dan *variance*, Kualitas hasil *biclustering* juga dapat dievaluasi dari aspek representasi data menggunakan *Matrix Coverage. Matrix Coverage* menunjukkan proporsi sel dalam matriks data yang berhasil dicakup oleh seluruh *bicluster* dibandingkan dengan jumlah total sel dalam matriks. Secara konseptual, *Matrix Coverage* didefinisikan sebagai berikut (Saber & Elloumi, 2014) :

$$Matrix\ Covarage\ = \frac{\text{Jumlah sel yang dicakup oleh seluruh bicluster yang di ekstraksi}}{Jumlah\ total\ sel\ dalam\ matrix}$$

Secara matematis, *matrix covarage* dapat dirumuskan sebagai berikut :

$$MC = \frac{U_{k=1}^{K}\ B_k}{|X|} \quad (7)$$

Keterangan :

MC : Nilai *Matrix Coverage.*
B_k : Himpunan sel yang tercakup pada bicluster ke- k.
$U_{k=1}^{K}\ B_k$: Gabungan seluruh sel unik yang tercakup oleh *bicluster*.
K : Jumlah *bicluster* yang terbentuk.
$|X|$: Jumlah total sel dalam matrix data.

E-ISSN : 2962-9713
P-ISSN : 2963-928X

HASIL DAN PEMBAHASAN

Prepocesing Data

Pada tahap awal pengolahan data, dilakukan *prepocesing* data yaitu identifikasi terhadap keberadaan nilai yang hilang pada setiap variabel. Hasil pemeriksaan menunjukkan adanya nilai yang tidak tersedia pada variabel produksi jagung untuk Provinsi DKI Jakarta serta variabel curah hujan rata-rata tahunan untuk Provinsi Papua Selatan dan Papua Tengah. Untuk mengatasi permasalahan tersebut, nilai yang hilang diisi menggunakan metode *mean imputation*, sehingga seluruh data provinsi dapat digunakan dalam proses analisis selanjutnya.

Tabel 1. Penangan Missing Value

Provinsi	Variabel	Nilai Mean Imputation
DKI Jakarta	Produksi Jagung	553478.1
Papua Selatan	Curah Hujan Rata-rata tahunan	2561.421
Papua Tengah	Curah Hujan Rata-rata tahunan	2561.421

Berdasarkan **tabel 1**, terdapat 3 provinsi yang memiliki missing dengan variabel Prouksi jagung dan Curah hujan rata-rata tahunan. nilai yang hilang tersebut ditangani menggunakan metode *mean imputation*, yaitu dengan menggantinya menggunakan nilai rata-rata pada variabel yang bersangkutan.

Standarisasi Data

Sebelum dilakukan proses biclustering, data terlebih dahulu di standarisasi mengguanakan metode Z-Score. Proses standarisasi data bertujuan untuk menyetarakan skala antar variabel sehingga masing-masing memiliki peran yang seimbang dalam analisis. Data yang digunakan terdiri atas 38 Provinsi (Baris) dan 7 variabel Penilitian sebagai unit pengamatan (Kolom). Berikut adalah hasil standarisasi data dengan menggunakan metode Z-Score.

Tabel 2. Hasil Standarisasi Data

Index	X1	X2	X3	X4	X5	X6	X7
1	0.087	-0.025	-0.416	0.463	2.299	0.337	-0.243
2	0.367	0.144	1.125	-0.474	-0.433	0.322	-0.230
3	0.073	0.120	0.127	-0.777	-0.433	0.282	-0.201
---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---
---	---	---	---	---	---	---	---
36	-0.514	-0.475	-0.475	0.945	-0.433	-1.940	4.242
37	-0.622	-0.541	-0.471	2.796	-0.433	-4.842	4.242
38	-0.625	-0.543	-0.477	3.278	-0.432	-2.762	-0.280

Berdasarkan **tabel 2,** hasil stndarisasi data menggunakan metode Z- Score menunjukkan bahwa nilai setiap variabel telah berada pada skala yang relatif sebanding dengan rata-rata mendekati 0. Nilai positif menunjukkan bahwa nilai suatu provinsi berada di atas rata-rata variabel, sedangkan nilai negatif menunjukkan nilai berada di bawah rata-rata. Proses ini dilakukan untuk menyetarakan skala antarvariabel, sehingga perbedaan satuan tidak memengaruhi hasil analisis dan proses clustering dapat dilakukan secara lebih objektif.

Analisis *Plaid Model*

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Analisis biclustering menggunakan Plaid Model dilakukan untuk mengidentifikasi pola hubungan secara simultan antara provinsi dan variabel yang diamati. Tahap awal dilakukan dengan membentuk model berdasarkan rata-rata global, efek baris, dan efek kolom, kemudian menghitung matriks residual dari selisih nilai aktual dan hasil model. Selanjutnya, bicluster dibentuk secara iteratif melalui estimasi parameter, pembaruan keanggotaan baris dan kolom, serta proses *backfitting* hingga model mencapai kondisi konvergen. Setelah model awal *Plaid Model* terbentuk tahap selanjutnya, dilakukan *update membership* untuk menentukan keanggotaan provinsi dan variabel pada masing-masing *bicluster*. Keanggotaan ditentukan menggunakan *threshold* 0,5, sehingga elemen yang memenuhi batas tersebut dipertahankan dalam *bicluster*.

Tabel 3. Keanggotan *Bicluster* Hasil *Update Membership*

Bicluster	Jumlah Wilayah	Variabel yang terlibat
BC 1	25	Luas Panen Padi, Produksi Padi, Produksi Jagung, Angka Harapan Hidup
BC 2	22	Rumah Tangga dengan Akses Utama dari Listrik, Curah Hujan Rata-rata Tahunan
BC 3	1	Persentase Penduduk Miskin
BC 4	3	Rumah Tangga dengan Akses Utama dari Listrik, Curah Hujan Rata-rata Tahunan
BC 5	2	Luas Panen Padi, Produksi Padi, Produksi Jagung, Persentase Penduduk Miskin, Angka Harapan Hidup
BC 6	0	-

Berdasarkan tabel 3, hasil *update membership*, diperoleh beberapa *bicluster* dengan jumlah wilayah dan kombinasi variabel yang berbeda. BC1 mencakup 25 wilayah dengan variabel pertanian dan angka harapan hidup (X1, X2, X3, dan X5), sedangkan BC2 dan BC4 masing-masing terdiri atas 22 dan 3 wilayah dengan karakteristik yang berkaitan dengan akses listrik dan curah hujan (X6 dan X7). BC3 dan BC5 memiliki jumlah anggota yang relatif kecil dengan pola yang masing-masing berkaitan dengan kemiskinan serta kombinasi indikator pertanian dan sosial. Sementara itu, BC6 tidak memiliki anggota setelah proses *thresholding*. Hasil tersebut menunjukkan bahwa pola wilayah dalam data dapat dikelompokkan berdasarkan keterkaitan indikator pertanian-sosial serta aspek infrastruktur dan kondisi iklim.

Keanggotaan yang diperoleh selanjutnya diseleksi melalui tahap *pruning* berdasarkan ukuran minimum *bicluster*. Tahap ini dilakukan untuk mengeliminasi kelompok yang memiliki jumlah anggota terlalu kecil sehingga diperoleh *bicluster* yang lebih representatif.

Tabel 4. Hasil Pruning Bicluster

Keterangan	Jumlah
Jumlah Bicluster Hasil Update Membership	5
Jumlah Bicluster yang Lolos Tahap Pruning	3
Jumlah Bicluster yang Dieliminasi	2

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Berdasarkan **Tabel 4,** proses *pruning* mempertahankan BC1, BC2, dan BC4, sedangkan BC3 dan BC5 dieliminasi karena tidak memenuhi ukuran minimum. Ketiga *bicluster* tersebut kemudian digunakan pada tahap *backfitting*. *Bicluster* yang lolos tahap *pruning* selanjutnya melalui proses *backfitting* untuk memperbarui estimasi efek baris, kolom, dan interaksi secara iteratif. Kestabilan model kemudian dievaluasi berdasarkan nilai *mean residual* dan variansi residual. Hasil *back-fitting* menunjukkan mean residual yang sangat dekat dengan nol pada BC1, BC2, dan BC4, masing-masing sekitar $-2{,}998\times10^{-17}$, $3{,}533\times10^{-17}$, dan $1{,}850\times10^{-17}$. Variansi residual masing-masing sebesar 0,01981; 0,07132; dan 0,04458. Hal ini menunjukkan bahwa model telah mencapai kondisi konvergen dan bicluster final memiliki pola yang relatif stabil.

Berdasarkan seluruh tahapan tersebut, diperoleh *bicluster* final yang telah memenuhi proses seleksi dan menunjukkan kestabilan model. Hasil akhir *biclustering* menggunakan *Plaid Model* disajikan pada Tabel berikut.

Tabel 5. Hasil *Bicluster* Metode *Plaid Model*

Bicluster	Jumlah Provinsi	Variabel yang Terlibat
BC 1	25	Luas Panen Padi, Produksi Padi, Produksi Jagung, Angka Harapan Hidup
BC 2	22	Rumah Tangga dengan Akses Utama dari Listrik, Curah Hujan Rata-rata Tahunan
BC 4	3	Rumah Tangga dengan Akses Utama dari Listrik, Curah Hujan Rata-rata Tahunan

Berdasarkan **tabel 5,** diperoleh tiga *bicluster* akhir, yaitu BC1, BC2, dan BC4, masing-masing mencakup 25, 22, dan 3 provinsi. BC1 menunjukkan pola yang berkaitan dengan indikator pertanian dan sosial, meliputi luas panen, produksi padi, produksi jagung, serta angka harapan hidup. Sementara itu, BC2 dan BC4 memiliki pola pada akses listrik dan curah hujan. Hasil ini menunjukkan adanya perbedaan karakteristik ketahanan pangan antarwilayah yang terutama tercermin melalui aspek pertanian-sosial dan infrastruktur-iklim.

Analisis *Spectral Biclustering*

Spectral Biclustering diterapkan untuk mengidentifikasi kelompok provinsi dan variabel berdasarkan pola yang terbentuk secara simultan. Tahapan analisis meliputi dekomposisi SVD, pengelompokan baris dan kolom, serta penyusunan ulang matriks untuk memperoleh struktur *checkerboard*. Hasil SVD menunjukkan bahwa tiga komponen awal memiliki nilai singular terbesar, yaitu 10,6742; 8,7724; dan 5,7311, sehingga ketiganya digunakan sebagai representasi utama dalam proses pengelompokan. Hasil dekomposisi ini selanjutnya menjadi dasar untuk menentukan struktur kelompok pada dimensi baris dan kolom.

Penentuan jumlah kelompok dilakukan menggunakan evaluasi *Silhouette Score* pada hasil representasi spektral. Berikut merupakan jumlah pengelompokkan berdasarkan nilai *Silhouette Score.*

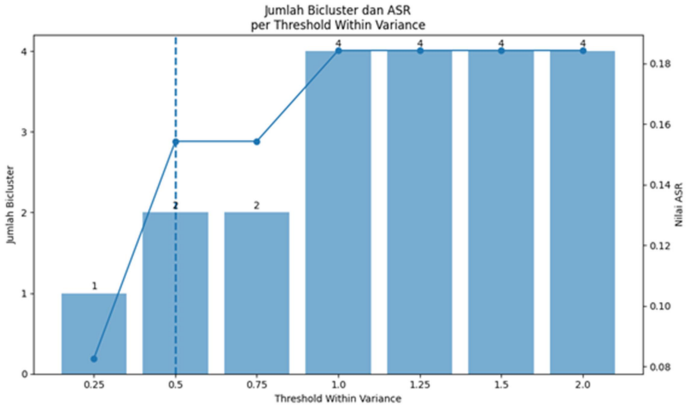
Tabel 6. Nilai *Silhouette Score* Tertinggi

Parameter	Nilai Optimal
K_{Row}	3

E-ISSN : 2962-9713
P-ISSN : 2963-928X

Silhouette Score Row	0.621192
Kcol	2
Silhouette Score Col	0.365542

Berdasarkan **tabel 6.** Hasil pengujian menunjukkan bahwa dimensi baris membentuk 3 kelompok dengan nilai Silhouette sebesar 0,621192, sedangkan dimensi kolom membentuk 2 kelompok dengan nilai sebesar 0,365542. Jumlah kelompok tersebut kemudian digunakan dalam proses pengelompokan provinsi dan variabel**.** Setelah struktur kelompok pada baris dan kolom diperoleh, tahap berikutnya dilakukan dengan menyusun kembali matriks untuk melihat pola kelompok secara simultan, Struktur tersebut kemudian digunakan sebagai dasar untuk menentukan tingkat homogenitas *bicluster* melalui pengujian parameter *within variance*. Pengujian *within variance* dilakukan pada beberapa nilai threshold untuk memperoleh struktur bicluster yang sesuai.



Gambar 1 Nilai ASR dan Jumlah Bicluster per Threshold Within Variance

Berdasarkan **gambar 1**, menunjukkan bahwa *threshold* 0,50 menghasilkan 2 *bicluster* valid dengan nilai *Average Squared Residue* (ASR) sekitar 0,1543. Nilai *threshold* tersebut dipilih berdasarkan pertimbangan keseimbangan antara jumlah bicluster yang terbentuk dan struktur kelompok yang dihasilkan, sehingga model tetap memiliki pola yang cukup homogen tanpa menghasilkan kelompok yang terlalu kompleks. Oleh karena itu, threshold 0,50 dianggap paling sesuai untuk digunakan dalam proses evaluasi bicluster yang memiliki tingkat homogenitas yang baik. Berdasarkan hasil penentuan *threshold* tersebut, tahap selanjutnya adalah mengidentifikasi komposisi *bicluster* akhir yang terbentuk dari kombinasi kelompok provinsi dan variabel.

Tabel 7. Hasil Seleksi *Bicluster Metode Spectral Biclustering*

Bicluster	Ukuran	Variabel
3	(28,5)	Luas Panen Padi, Produksi Padi, Produksi Jagung, Persentase Penduduk Miskin, Curah Hujan Rata-Rata Tahunan
4	(28,2)	Angka Harapan Hidup, Rumah Tangga dengan Akses Utama dari Listrik

Berdasarkan **tabel 7**, hasil *Spectral Biclustering* menunjukkan bahwa dari enam kombinasi bicluster yang terbentuk melalui 3 cluster baris dan 2 cluster kolom, terdapat dua *bicluster* yang dipertahankan berdasarkan kriteria *within variance*, yaitu BC3 dan

E-ISSN : 2962-9713
P-ISSN : 2963-928X

BC4. BC3 mencakup 28 provinsi dan lima variabel, yaitu luas panen padi (X1), produksi padi (X2), produksi jagung (X3), persentase penduduk miskin (X4), serta curah hujan rata-rata tahunan (X7). Kelompok ini menggambarkan keterkaitan antara kapasitas produksi pangan, kondisi sosial ekonomi, dan faktor lingkungan yang secara bersama-sama membentuk karakteristik ketahanan pangan wilayah. Sementara itu, BC4 terdiri atas 28 provinsi dengan dua variabel, yaitu angka harapan hidup (X5) dan akses utama rumah tangga terhadap listrik (X6). Kedua indikator tersebut merepresentasikan aspek kesejahteraan masyarakat dan ketersediaan infrastruktur dasar. Pola ini menunjukkan bahwa karakteristik ketahanan pangan tidak hanya berkaitan dengan produksi pertanian, tetapi juga memiliki keterkaitan dengan kondisi kehidupan masyarakat dan dukungan infrastruktur di suatu wilayah.

Secara keseluruhan, kedua *bicluster* menunjukkan dua pola utama dalam data, yaitu keterkaitan indikator pertanian, sosial ekonomi, dan lingkungan pada BC3 serta keterkaitan kesejahteraan dan infrastruktur pada BC4. Hasil tersebut memperlihatkan *bahwa Spectral Biclustering* dapat mengidentifikasi pola hubungan antara kelompok provinsi dan variabel secara simultan, sehingga memberikan gambaran yang lebih terstruktur mengenai karakteristik ketahanan pangan antarwilayah.

Evaluasi Kinerja Metode

Evaluasi ini bertujuan untuk menilai sejauh mana pola hubungan antara wilayah dan variabel dalam setiap bicluster bersifat homogen dan stabil. Proses evaluasi utama dilakukan menggunakan ukuran Mean Squared Residue (MSR) dan Average Residue (ASR). Selain itu, evaluasi kualitas hasil juga dapat dilengkapi dengan ukuran berbasis statistik deskriptif yaitu evaluasi variance (VAR) . Selanjutnya, kualitas hasil biclustering juga dapat dievaluasi dari aspek representasi data menggunakan Matrix Coverage.

Tabel 8. Nilai MSR pada setiap *Bicluster*

Metode	Bicluster	MSR
Plaid Model	BC 1	0.019
	BC 2	0.071
	BC 4	0.044
Sepectral Biclustering	BC 3	0.225
	BC 4	0.082

Setelah diperoleh nilai Mean Squared Residue (MSR) pada setiap bicluster, langkah selanjutnya adalah menentukan *Average Residue* (ASR) global.

Tabel 9. Perbandingan Nilai ASR Antar Metode

Metode	ASR
Plaid Model	0.0452
Spectral Biclustering	0.1543

Berdasarkan **tabel 9**, nilai ASR, Plaid Model menghasilkan nilai yang lebih rendah, yaitu 0,0452, dibandingkan Spectral Biclustering sebesar 0,1543. Nilai tersebut menunjukkan bahwa Plaid Model menghasilkan pola bicluster dengan residual yang lebih rendah, sehingga struktur kelompok yang terbentuk cenderung lebih homogen berdasarkan kriteria ASR. Dengan demikian, Plaid Model menunjukkan nilai ASR yang lebih rendah dibandingkan Spectral Biclustering pada data penelitian ini.

Tabel 10. Hasil *Variance* (*Plaid Model*)

Bicluster	Rows	Cols	Variance
1	25	4	0.054851
2	22	2	0.157241

E-ISSN : 2962-9713
P-ISSN : 2963-928X

4	3	2	0.116501

Berdasarkan Tabel 10, nilai *variance* pada *bicluster* Plaid Model menunjukkan tingkat penyebaran data yang berbeda. BC 1 memiliki variance terendah sebesar 0,054851, yang menunjukkan pola data paling homogen dibandingkan bicluster lainnya. Sementara itu, BC2 dan BC4 memiliki variance yang lebih tinggi sehingga menunjukkan keragaman data yang relatif lebih besar.

Tabel 11. Hasil *Variance* (*Spectral Biclustering*)

Bicluster	Rows	Cols	Variance
3	28	5	0.309131
4	28	2	0.248636

Berdasarkan **Tabel 9 dan Tabel 10**, nilai variance pada kedua metode menunjukkan tingkat penyebaran data yang berbeda. *Plaid Model* memiliki rentang variance 0,054851–0,157241, 0, 116501, sedangkan *Spectral Biclustering* sebesar 0,248636–0,309131. Nilai yang lebih rendah pada *Plaid Model* menunjukkan bahwa *bicluster* yang terbentuk cenderung lebih homogen berdasarkan kriteria *variance.*

Tabel 12. Hasil *Matrix Covarage Plaid Model* dan *Spectral Biclustering*

Metode	Komponen	Nilai
Plaid Model	Jumlah total sel dalam matriks	266
	Jumlah sel yang tercakup oleh bicluster	150
	Matrix Covarage	0.5639
Spectral Biclustering	Jumlah total sel dalam matriks	266
	Jumlah sel yang tercakup oleh bicluster	196
	Matrix Covarage	0.7368

Berdasarkan tabel 12, *Plaid Model* menghasilkan nilai *Matrix Coverage* sebesar 0,5639, sedangkan *Spectral Biclustering* sebesar 0,7368. Nilai tersebut menunjukkan bahwa Spectral Biclustering mencakup proporsi struktur data yang lebih besar dibandingkan *Plaid Mode*l, yaitu 73,68% berbanding 56,39%.

SIMPULAN

Berdasarkan hasil analisis, *Spectral Biclustering* menghasilkan dua bicluster akhir yang menggambarkan keterkaitan indikator pertanian, sosial ekonomi, lingkungan, kesejahteraan, dan infrastruktur. Sementara itu, *Plaid Model* menghasilkan tiga bicluster dengan pola yang lebih spesifik pada kelompok wilayah tertentu. Hasil evaluasi menunjukkan bahwa *Plaid Model* memiliki nilai MSR, ASR, dan *variance* yang lebih rendah, sedangkan *Spectral Biclustering* memiliki *Matrix Coverage* yang lebih tinggi. Dengan demikian, kedua metode mampu mengidentifikasi pola ketahanan pangan dengan karakteristik yang berbeda, dengan *Plaid Model* menunjukkan homogenitas bicluster yang lebih tinggi berdasarkan kriteria evaluasi yang digunakan. Oleh karena itu tujuan penelitian ini adalah mengidentifikasi pola ketahanan pangan yang homogen dan mudah diinterpretasikan, maka metode *Plaid Model* dapat direkomendasikan sebagai metode yang lebih optimal dalam mengidentifikasi pola ketahanan pangan di Indonesia

DAFTAR PUSTAKA

Castanho, E. N., Aidos, H., & Madeira, S. C. (2022). Biclustering fMRI time series : a comparative study. *BMC Bioinformatics*, 1–30. https://doi.org/10.1186/s12859-022-

E-ISSN : 2962-9713
P-ISSN : 2963-928X

04733-8

Castanho, E. N., Aidos, H., & Madeira, S. C. (2024). *Biclustering data analysis : a comprehensive survey*. *25*(4).

Henriques, R., & Madeira, S. C. (2016). BiC2PAM: constraint-guided biclustering for biological data analysis with domain knowledge. *Algorithms for Molecular Biology*, *11*(1), 23. https://doi.org/10.1186/s13015-016-0085-5

Hestina, J. (2011). KETAHANAN PANGAN Heri Suharyanto * Abstrak. *Sosial Humaniora*, *4*(2), 186–194. http://iptek.its.ac.id/index.php/jsh/article/view/633/355

Insecurity, C. F., & Insecurity, T. F. (1996). *An Introduction to the Basic Concepts of Food Security Food Security Information for Action*. 1–3.

Kluger, Y., Basri, R., Chang, J. T., & Gerstein, M. (2003). *Spectral Biclustering of Microarray Data : Coclustering Genes and Conditions*. 703–716. https://doi.org/10.1101/gr.648603.Checkerboard

Lee., Y. (2009). Validation Measures of Bicluster Solutions. *Validation Measures of Bicluster Solutions*.

Majd, H. A., Shahsavari, S., Baghestani, A. R., Tabatabaei, S. M., Bashi, N. K., Tavirani, M. R., & Hamidpour, M. (2016). *Evaluation of Plaid Models in Biclustering of Gene Expression Data*. *2016*. https://doi.org/10.1155/2016/3059767

Patowary, P., Sarmah, R., & Bhattacharyya, D. K. (2020). Developing an effective biclustering technique using an enhanced proximity measure. *Network Modeling Analysis in Health Informatics and Bioinformatics*, *9*(1), 6. https://doi.org/10.1007/s13721-019-0211-7

Saber, H. Ben, & Elloumi, M. (2014). *A C OMPARATIVE S TUDY OF C LUSTERING AND*. *6*(6), 93–111. https://doi.org/10.5121/ijcsit.2014.6607

Siswantining, T., & Swasti, O. (2021). *Biclustering Analysis Using Plaid Model on Gene Expression Data of Colon Cancer*. *50*(July), 101–114. https://doi.org/10.17713/ajs.v50i5.1195