

# Penggerombolan Provinsi di Indonesia Berdasarkan Produktivitas Tanaman Pangan Tahun 2005-2015 Menggunakan Metode *K-Error*

Emeylia Safitri\*, I Made Sumertajaya\*, Akbar Rizki\*

\*Departemen Statistika Institut Pertanian Bogor

**Abstrak**—Analisis gerombol merupakan analisis peubah ganda yang bertujuan untuk mengelompokkan objek-objek pengamatan ke dalam beberapa gerombol. Gerombol-gerombol yang terbentuk memiliki kemiripan antar gerombol rendah dan kemiripan di dalam gerombol tinggi. Analisis gerombol klasik memiliki kelemahan yaitu tidak memasukkan informasi kesalahan pengukuran terkait data. Analisis gerombol dengan metode *K-Error* dikembangkan untuk mengatasi masalah kesalahan pengukuran data dalam analisis gerombol klasik. Provinsi-provinsi di Indonesia akan digerombolkan berdasarkan produktivitas tanaman pangan dengan menggunakan metode *K-Error* dan *K-Means*. Metode *K-Error* menghasilkan penggerombolan yang lebih baik dibandingkan dengan metode *K-Means*. Gerombol optimum yang dihasilkan berjumlah 7 gerombol. Gerombol 5 terdiri atas provinsi-provinsi yang memiliki produktivitas tinggi pada hampir seluruh jenis tanaman pangan. Gerombol 2 dan 3 memiliki produktivitas sangat rendah untuk sebagian besar jenis tanaman pangan.

**Kata kunci**—analisis gerombol ;*k-error* ;provinsi ;tanaman pangan

## I. PENDAHULUAN

### A. Latar Belakang

Negara Indonesia dijuluki sebagai negara agraris karena melimpahnya sumber daya alam dan luasnya lahan pertanian yang tersedia. Sektor pertanian berkontribusi besar terhadap pendapatan negara melalui ekspor hasil pertanian, dan dapat mengurangi pengangguran karena menyerap tenaga kerja. Sektor pertanian sangat erat kaitannya dengan tanaman pangan. Tanaman pangan yang ada di Indonesia terdiri atas padi, kacang-kacangan, dan umbi-umbian. Padi menjadi tanaman pangan yang mendominasi di Indonesia karena hampir seluruh masyarakat Indonesia mengonsumsi nasi sebagai makanan pokoknya. Kacang-kacangan dan umbi-

umbian juga dijadikan sebagai makanan pokok bagi sebagian masyarakat Indonesia.

Berdasarkan data Kementerian Pertanian tahun 2015, daerah penghasil padi dengan produktivitas tinggi antara lain Jawa Barat, Sumatera Barat, dan Nusa Tenggara Barat. Tanaman pangan seperti jagung banyak dihasilkan di daerah Jawa Barat, Sumatera Barat, Sumatera Utara, Sumatera Selatan dan Nusa Tenggara Barat. Daerah dengan produktivitas tinggi penghasil kacang hijau antara lain Sulawesi Barat, Gorontalo, dan Jawa Barat sedangkan untuk kedelai banyak dihasilkan di daerah Jawa Timur, Kalimantan Timur, dan Sulawesi Tengah.

Penyebaran tanaman pangan di Indonesia berdasarkan provinsi dapat dibangun menjadi sebuah pemetaan tanaman pangan. Kebijakan terkait dengan pemetaan tanaman pangan di Indonesia perlu dilakukan oleh pemerintah guna meningkatkan potensi sektor tanaman pangan secara nasional. Pemetaan provinsi di Indonesia terkait dengan penyebaran tanaman pangan dapat dilakukan menggunakan analisis gerombol. Analisis gerombol memiliki beberapa jenis metode seperti metode *K-Means* dan metode Ward. Menurut Johnson dan Wichern (2007), metode *K-Means* memiliki dasar yaitu mempartisi data ke dalam gerombol berdasarkan rata-rata gerombol terdekat. Metode tersebut tidak memasukkan informasi kesalahan terkait data.

Pada tahun 2005, Kumar dan Patel mengembangkan metode analisis gerombol yang memasukkan informasi kesalahan terkait data salah satunya metode *K-Error*. Penelitian sejenis juga dilakukan oleh Kurniawan dan Saikhu pada tahun 2010, analisis gerombol berdasarkan kesalahan pengukuran terkait data seperti metode *K-Error* da-

pat menyelesaikan masalah penggerombolan lebih baik dibandingkan dengan analisis gerombol klasik seperti *K-Means* dan Ward. Analisis gerombol mampu memberikan hasil yang lebih akurat dengan memasukkan unsur kesalahan karena data yang digunakan merupakan peubah acak (memiliki unsur kesalahan). Oleh karena itu, dalam penelitian ini akan digunakan metode *K-Error* dan metode lain seperti *K-Means* untuk menggerombolkan provinsi di Indonesia dengan indikator produktivitas tanaman pangan tahun 2005-2015.

### B. Tujuan

Tujuan dari penelitian ini adalah :

- 1) Membandingkan hasil metode *K-Error* dan metode *K-Means*
- 2) Menggerombolkan provinsi di Indonesia dengan *K-Error* untuk mengetahui karakteristik provinsi-provinsi di Indonesia berdasarkan produktivitas tanaman pangan pada tahun 2005-2015.

## II. TINJAUAN PUSTAKA

### A. Error-based clustering

Metode *Error-based clustering* adalah salah satu metode penggerombolan objek yang diteliti oleh Kumar dan Patel pada tahun 2005. Metode penggerombolan klasik seperti *K-Means* dan Ward memiliki kekurangan yaitu tidak mempertimbangkan informasi kesalahan terkait data yang dapat menghasilkan gerombol yang tidak akurat. Penelitian yang dilakukan oleh Kumar dan Patel (2005), metode penggerombolan dengan memasukkan kesalahan pengukuran terkait data dapat menghasilkan hasil gerombol yang lebih baik dan berbeda jika dibandingkan dengan hasil analisis gerombol klasik. Metode penggerombolan klasik mengasumsikan bahwa rataan dan matriks ragam peragam tidak diketahui, sedangkan metode *Error-based clustering* mengasumsikan bahwa matriks ragam peragam diketahui sedangkan rataan tidak diketahui.

### B. Algoritma K-Error

Metode *K-Error* merupakan salah satu metode analisis gerombol yang memasukkan informasi kesalahan terkait data. Metode *K-Error* digunakan

saat diketahui jumlah gerombol sejak awal. Metode *K-Error* bersifat iteratif dengan tahap awal yaitu inialisasi partisi data dengan algoritme sebagai berikut (Kumar and Patel (2005)):

- 1) Tentukan banyaknya gerombol yaitu  $k$ , untuk masing-masing gerombol hitung pusat gerombol dengan menggunakan rumus sebagai berikut:

$$\widehat{\theta}_k = \left( \sum_{i \in R_k} S_i^{-1} \right)^{-1} \left( \sum_{i \in R_k} S_i^{-1} x_i \right) \quad (1)$$

Keterangan:

$\widehat{\theta}_k$  : vektor penduga pusat gerombol ke- $k$ ;  
 $i$  : 1,2,3,..., $n$ ;  
 $n$  : banyaknya objek pengamatan;  
 $k$  : banyaknya gerombol;  
 $C_k$  : gerombol ke- $k$ ;  
 $x_i$  : vektor objek pengamatan ke- $i$   
 $S_i$  : matriks ragam peragam objek pengamatan ke- $i$

$$S_i = \begin{bmatrix} S_{xi1}^2 & S_{i1,i2} & \cdots & S_{i1,ip} \\ S_{i2,i1} & S_{xi2}^2 & \cdots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ S_{i1,ip} & S_{i2,ip} & \cdots & S_{xip}^2 \end{bmatrix} \quad (2)$$

Keterangan:

$j$  : 1,2,3,..., $p$ ;  
 $S_{xij}^2$  : ragam objek pengamatan ke- $i$ , peubah ke- $j$ ;  
 $S_{ij_1,ij_2}$  : peragam antara objek pengamatan ke- $i$ , peubah ke- $j_1$  dan objek pengamatan ke- $i$ , peubah ke- $j_2$ , dengan  $j_1 \neq j_2$

- 2) Hitung kembali jarak setiap objek pengamatan ke masing-masing pusat gerombol, masukkan objek pengamatan ke dalam gerombol dengan jarak terdekat. Rumus jarak yang digunakan sebagai berikut :

$$d_{ik} = (x_i - \widehat{\theta}_k)^t S_i^{-1} (x_i - \widehat{\theta}_k) \quad (3)$$

Keterangan :

$d_{ik}$  : jarak  $x_i$  dari pusat gerombol ke- $k$ ;  
 $\widehat{\theta}_k$  : vektor penduga pusat gerombol ke- $k$ ;  
 $x_i$  : vektor objek pengamatan ke- $i$ ;  
 $S_i$  : matriks ragam peragam objek

Algoritma K-Error tersebut akan terus diulang sampai tidak ada lagi objek pengamatan yang berpindah gerombol. K-Error memiliki kekurangan yaitu jumlah gerombol yang dihasilkan bisa saja berkurang dari jumlah gerombol yang ditentukan dari awal karena gerombol tersebut tidak memiliki anggota objek.

C. Ukuran Keragaman

Keragaman dari gerombol yang terbentuk dapat dijadikan acuan dalam melakukan validasi gerombol. Terdapat dua ukuran yang digunakan dalam melakukan validitas gerombol, sebagai berikut (Kalkstein et al. (1987)):

- 1) Nilai keragaman di dalam gerombol yang dapat diukur dengan menggunakan rata-rata standar deviasi dalam gerombol dengan rumus sebagai berikut:

$$\hat{\sigma}_w = \frac{1}{k} \sum_{k=1}^g \hat{\sigma}_{jk} \tag{4}$$

Keterangan :

$\hat{\sigma}_w$  : rata-rata simpangan baku dalam gerombol

$\hat{\sigma}_{jk}$  : simpangan baku dalam gerombol ke-*k* peubah ke-*j*

- 2) Nilai keragaman antar gerombol yang dapat diukur dengan menggunakan rata-rata standar deviasi antar gerombol dengan rumus sebagai berikut :

$$\hat{\sigma}_b = \left[ \frac{1}{k} \sum_{k=1}^g (\bar{x}_{jk} - \bar{\bar{x}}_j)^2 \right]^{1/2} \tag{5}$$

Keterangan :

$\hat{\sigma}_b$  : rata-rata simpangan baku antar gerombol

$\bar{x}_{jk}$  : nilai rata-rata peubah ke-*j* pada gerombol ke-*k*

$\bar{\bar{x}}_j$  : nilai rata-rata peubah ke-*j* untuk seluruh gerombol

Nilai  $\hat{\sigma}_w$  akan menunjukkan hasil yang semakin baik ketika nilainya semakin kecil. Sedangkan nilai  $\hat{\sigma}_b$  akan menunjukkan hasil yang baik ketika nilainya semakin besar. Oleh karena itu, nilai rasio antara  $\hat{\sigma}_w/\hat{\sigma}_b$  akan semakin baik jika nilainya semakin kecil (Alfina et al. (2012)).

III. METODE

A. Data

Data yang digunakan dalam penelitian ini adalah data terkait produksi tanaman pangan pada setiap provinsi di Indonesia tahun 2005-2015. Data yang digunakan merupakan data sekunder yang bersumber dari laman web BPS (Badan Pusat Statistik). Adapun jumlah provinsi di Indonesia yaitu 34 provinsi akan tetapi jumlah provinsi yang digunakan dalam penelitian ini berjumlah 32. Provinsi yang tidak digunakan sebagai objek pengamatan yaitu provinsi DKI Jakarta dan Kalimantan Utara. Provinsi DKI Jakarta tidak diikutsertakan karena DKI Jakarta merupakan ibu kota negara dan telah menjadi kota besar sehingga tidak memiliki lahan untuk bertani. Provinsi Kalimantan Utara tidak diikutsertakan karena merupakan provinsi baru yang resmi menjadi sebuah provinsi pada tahun 2012 sehingga tidak dapat menduga matriks ragam peragam.

Peubah yang digunakan dalam penelitian ini adalah produktivitas tanaman pangan. Tanaman pangan yang dijadikan peubah yaitu produktivitas jagung (X1), kacang tanah (X2), kacang hijau (X3), kedelai (X4), padi (X5), ubi jalar (X6), ubi jalar (X7). Seluruh peubah yang digunakan memiliki satuan kuintal/hektar.

B. Metode

Tahapan analisis data pada penelitian ini adalah sebagai berikut :

- 1) Persiapan data yaitu dengan menghitung rata-rata dan matriks ragam-peragam untuk masing-masing provinsi yang merupakan rataan dan ragam dari produktivitas tanaman pangan tahun 2005-2015. Melakukan uji Boxs untuk menguji kesamaan matriks ragam peragam tiap provinsi. Dengan hipotesis sebagai berikut :

$$H_0 = \sum_1 = \sum_2 = \dots = \sum_{32} = \sum$$

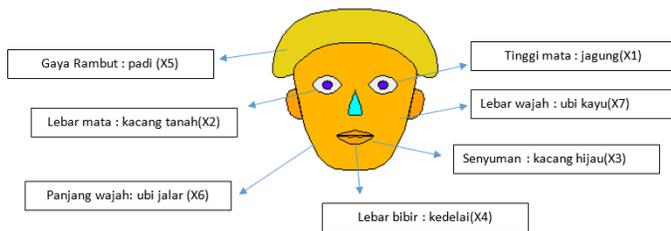
$$H_1 = \text{minimal ada satu } i, \sum_i \neq \sum,$$

dengan  $\sum_i$  adalah matriks ragam peragam provinsi ke-*i*

Dalam penelitian ini digunakan dua jenis pendugaan terhadap matriks ragam peragam yaitu dengan menggunakan matriks ragam peragam secara keseluruhan dan matriks

ragam peragam data deret waktu tiap provinsi.

- 2) Melakukan eksplorasi data yaitu dengan menghitung rata-rata dan ragam tanaman pangan, membuat boxplot untuk melihat karakteristik tanaman pangan, dan membuat *suffer face plot* dengan keterangan yang disajikan pada Gambar 1, untuk mengetahui karakteristik masing-masing provinsi.



Gambar 1. Keterangan dari *Suffer face plot*

- 3) Melakukan penggerombolan provinsi menggunakan metode *K-Error*. Nilai  $k$  yang digunakan yaitu  $k=2,3,\dots,8$  Nilai  $k$  merupakan banyaknya gerombol yang terbentuk. Algoritme *K-Error* yaitu sebagai berikut:
  - a) Mempartisi objek pengamatan menjadi  $k$ -gerombol secara acak.
  - b) Menghitung pusat gerombol untuk masing-masing  $k$ -gerombol menggunakan persamaan (1).
  - c) Menghitung jarak objek pengamatan terhadap pusat gerombolnya dengan menggunakan persamaan (3).
  - d) Memasukkan objek pengamatan ke dalam gerombol baru yang memiliki jarak terdekat dengan pusat gerombolnya.
  - e) Mengulangi tahap b sampai d sehingga tidak ada lagi objek pengamatan atau provinsi yang berpindah gerombol.
  - f) Menghitung perbandingan antara  $S_w$  dengan  $S_b$  dengan persamaan (4) dan (5).
  - g) Ulangi tahap a sampai f untuk nilai  $k=2,3,\dots,8$ .
  - h) Menentukan nilai terkecil perbandingan

antara  $S_w$  dengan  $S_b$  untuk nilai  $k$  yang digunakan.

- 4) Melakukan analisis gerombol dengan metode *K-Means* dengan tahap sebagai berikut :
  - a) Melakukan analisis gerombol dengan metode *K-Means* dengan nilai  $k=2,3,\dots,8$ .
  - b) Menghitung perbandingan antara  $S_w$  dengan  $S_b$  dengan persamaan (4) dan (5).
- 5) Melakukan ulasan perbandingan hasil dari metode yang digunakan yaitu *K-Error* dan *K-Means*.
- 6) Mendeskripsikan hasil gerombol terbaik yang didapatkan dengan metode *K-Error*.

#### IV. HASIL DAN PEMBAHASAN

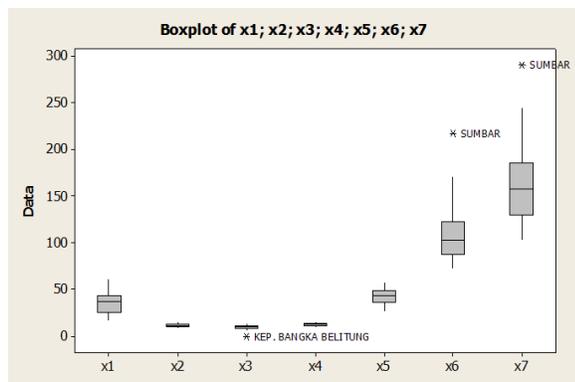
##### A. Statistik Deskriptif

Berdasarkan Tabel 1, terlihat bahwa produktivitas tanaman pangan tertinggi yaitu ubi kayu (X7) dengan rata-rata sebesar 162.8592 kuintal/ha. Kacang hijau (X3) dengan rata-rata produktivitas sebesar 9.943964 kuintal/ha menjadi tanaman pangan yang memiliki produktivitas terendah. Secara keseluruhan rata-rata produktivitas dari seluruh peubah tidak terlalu berbeda jauh, hanya ubi jalar (X6) dan ubi kayu (X7) yang memiliki rata-rata produktivitas lebih dari 100 kuintal/ha. Keragaman produktivitas tanaman pangan tertinggi yaitu tanaman jagung (X1) karena memiliki koefisien keragaman sebesar 31%, sedangkan tanaman pangan yang memiliki keragaman produktivitas terendah yaitu kedelai (X4).

Tabel I  
DESKRIPTIF MASING-MASING TANAMAN PANGAN

Peubah	Rataan	Ragam	KK
Jagung(X1)	36.26	126.20	31%
Kacang Tanah(X2)	11.86	2.81	14%
Kacang Hijau(X3)	9.94	6.56	26%
Kedelai (X4)	12.54	2.68	13%
Padi(X5)	43.41	77.76	20%
Ubi Jalar(X6)	112.00	1058.19	29%
Ubi Kayu(X7)	162.86	1752.65	26%

Karakteristik masing-masing peubah dapat dilihat menggunakan boxplot yang disajikan pada Gambar 2. Secara keseluruhan seluruh jenis tanaman pangan memiliki karakteristik yang berbeda-beda jika dilihat dari range masing-masing peubah. Untuk kacang tanah (X2), kacang hijau (X3), kedelai (X4) memiliki range yang sempit dalam hal ini berarti produktivitas untuk masing-masing provinsi tidak terlalu beragam. Untuk kedelai terdapat pencilan bawah yaitu provinsi Kepulauan Bangka Belitung. Ubi jalar (X6) dan ubi kayu (X7) memiliki keragaman yang besar dan memiliki pencilan atas yaitu provinsi Sumatera Barat.



Gambar 2. Boxplot untuk masing-masing tanaman pangan

Karakteristik untuk masing-masing provinsi dapat digambarkan melalui *suffer face plot* yang disajikan pada Gambar 3. *Suffer face plot* menggambarkan masing-masing tanaman pangan (X1-X7). Tanaman jagung (X1) digambarkan dengan tinggi mata, tanaman kacang tanah (X2) dengan lebar mata, kacang hijau (X3) dengan senyuman, kedelai (X4) dengan bentuk senyuman, padi (X5) dengan gaya rambut, ubi jalar (X6) dengan panjang wajah, dan ubi kayu (X7) dengan lebar wajah. Berdasarkan Gambar 3, terlihat bahwa masing-masing provinsi memiliki karakteristik yang berbeda-beda, namun sebagian memiliki kemiripan dan membentuk gerombol-gerombol. Secara eksploratif provinsi-provinsi tersebut akan membentuk gerombol berdasarkan kemiripan seperti provinsi Jawa Barat, Jawa Tengah, Jawa Timur, dan Sulawesi Selatan yang memiliki kemiripan karena memiliki produktivitas yang tinggi untuk beberapa jenis tanaman pangan. Provinsi Riau, Kep. Bangka Belitung, dan Sulawesi Tenggara pun memiliki kemiripan

berdasarkan *suffer face plot* yang terbentuk. Oleh karena itu, sangat relevan jika analisis selanjutnya yang dilakukan yaitu analisis gerombol.



Gambar 3. *Suffer face plot* untuk masing-masing provinsi

Kesamaan matriks ragam peragam masing-masing provinsi diuji menggunakan uji Box's M. Pengujian menghasilkan nilai p-value sebesar 0.000. Nilai alpha yang digunakan yaitu sebesar 5%, sehingga kesimpulan yang diambil adalah menolak hipotesis nol. Hal ini berarti matriks ragam peragam antara provinsi berbeda sehingga pendugaan awal masing-masing matriks ragam peragam menggunakan matriks ragam peragam tiap provinsi.

**B. Penggerombolan provinsi di Indonesia dengan K-Error**

Metode *K-Error* merupakan salah satu metode penggerombolan dengan memasukkan unsur ragam peragam. Dalam penelitian ini akan dicobakan dua pendugaan terhadap matriks ragam peragam. Pendugaan pertama yaitu menggunakan ragam keseluruhan untuk masing-masing tanaman pangan sedangkan pendugaan kedua yaitu menggunakan matriks ragam peragam dengan menggunakan data deret waktu untuk tiap provinsi. Evaluasi hasil gerombol, dapat dilihat dari perbandingan antara jarak di dalam gerombol ( $S_w$ ) dengan jarak antar gerombol nya ( $S_b$ ) untuk penentuan jumlah gerombol optimum. Nilai perbandingan  $S_w/S_b$  yang kecil mengindikasikan jumlah gerombol optimum. Nilai

Sw/Sb untuk beberapa nilai jumlah gerombol disajikan pada Tabel 2. Jumlah gerombol yang digunakan yaitu dari dua sampai dengan delapan gerombol. Jumlah gerombol yang digunakan hanya sampai 8 gerombol, hal ini disebabkan terdapat gerombol yang tidak memiliki anggota jika dinaikkan jumlah gerombolnya. Tabel 2 menunjukkan nilai Sw/Sb selalu kurang dari 1, berarti jarak di dalam gerombol selalu lebih kecil dibandingkan dengan jarak antar gerombol yang terbentuk. Hal ini sesuai dengan prinsip dasar penggerombolan yaitu jarak di dalam gerombol harus lebih kecil dibandingkan jarak antar gerombolnya.

Tabel II  
NILAI SW/SB K-Error DENGAN MATRIKS RAGAM PERAGAM KESELURUHAN

Jumlah gerombol (k)	Sw/Sb
2	0.94
3	0.82
4	0.75
5	0.72
6	0.60
7	0.56
8	0.64

### C. Penggerombolan provinsi di Indonesia dengan K-Means

Tabel III  
NILAI SW/SB DENGAN METODE K-Means

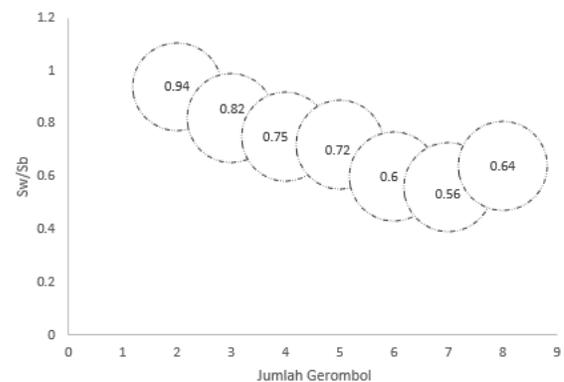
Jumlah gerombol (k)	Sw/Sb
2	1,06
3	1,56
4	0,79
5	0,68
6	0,70
7	0,64
8	0,63

Jumlah gerombol yang digunakan pada metode K-Means sama dengan jumlah gerombol yang digunakan pada metode K-Error. Jumlah gerombol optimum yang dihasilkan dengan metode K-Means yaitu sebanyak 8 gerombol karena memiliki nilai Sw/Sb terkecil yaitu sebesar 0.63. Hal ini dapat dilihat dari perbandingan Sw dengan Sb yang disajikan pada Tabel 3. Nilai perbandingan Sw/Sb secara

keseluruhan dengan menggunakan metode K-Means fluktuatif namun cenderung menurun.

### D. Evaluasi hasil metode K-Error dan metode K-Means

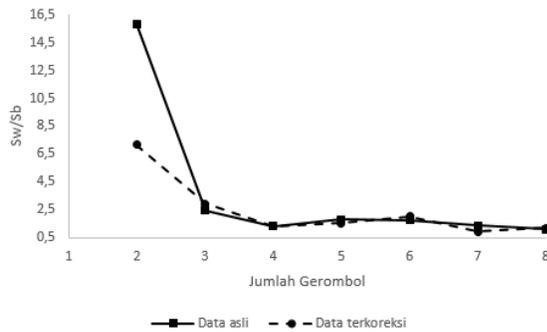
Nilai Sw/Sb K-Error dapat digunakan untuk membandingkan hasil metode K-Error dengan K-Means. Gambar 4 menunjukkan nilai Sw/Sb dengan matriks ragam peragam keseluruhan secara keseluruhan menurun seiring dengan naiknya jumlah gerombol. Nilai Sw/Sb menurun sampai dengan gerombol berjumlah 7 sedangkan ketika jumlah gerombol sebanyak 8 nilai Sw/Sb meningkat. Hasil evaluasi tersebut dapat disimpulkan jumlah gerombol optimum yang dihasilkan dengan metode K-Error menggunakan matriks ragam peragam keseluruhan sebanyak 7 gerombol.



Gambar 4. Pola nilai Sw/Sb K-Error dengan matriks ragam peragam keseluruhan

Perbandingan nilai Sw/Sb K-Error dengan matriks ragam peragam deret waktu tiap provinsi dapat dihitung dengan menggunakan data asli dan data terkoreksi. Data terkoreksi merupakan data asli yang dikalikan dengan invers matriks ragam peragamnya. Berdasarkan Gambar 5, secara keseluruhan nilai perbandingan Sw dengan Sb dengan data asli maupun data terkoreksi cenderung fluktuatif. Penurunan nilai perbandingan yang signifikan yaitu pada saat jumlah gerombol sebanyak 2 gerombol menjadi 3 gerombol. Pada jumlah gerombol sebanyak 3 gerombol sampai dengan 8 gerombol, nilai Sw/Sb cukup stabil tidak terjadi kenaikan maupun penurunan yang signifikan. Dari kedua data tersebut menghasilkan gerombol optimum sebanyak 7

gerombol karena memiliki nilai Sw/Sb terkecil.

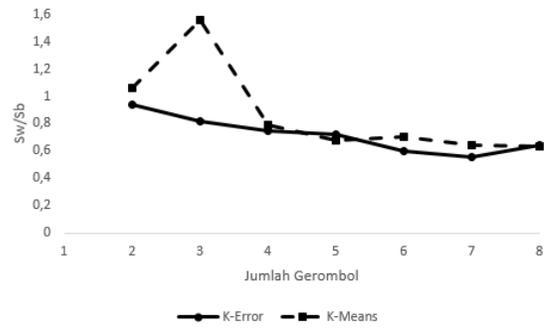


Gambar 5. Pola nilai Sw/Sb K-Error dengan matriks ragam peragam deret waktu tiap provinsi

Metode K-Error dengan matriks ragam peragam keseluruhan lebih tepat jika dibandingkan dengan K-Error dengan matriks ragam peragam deret waktu tiap provinsi. Oleh karena itu metode K-Error yang dipilih adalah hasil K-Error dengan menggunakan matriks ragam peragam keseluruhan. Selanjutnya metode K-Error yang digunakan merupakan K-Error dengan matriks ragam peragam keseluruhan. Gambar 6, menunjukkan perbedaan nilai Sw/Sb hasil metode K-Error dibandingkan dengan K-Means. Perbandingan antara Sw dengan Sb cenderung berbeda secara keseluruhan. Ketika jumlah gerombol sebanyak 3, perbedaan nilai Sw/Sb sangat signifikan dibandingkan dengan jumlah gerombol lainnya. Secara keseluruhan nilai Sw/Sb metode K-Error dengan matriks ragam peragam keseluruhan lebih kecil dibandingkan dengan metode K-Means. Oleh karena itu, dapat disimpulkan bahwa Metode K-Error lebih tepat dalam menggerombolkan provinsi di Indonesia dengan indikator produktivitas tanaman pangan tahun 2005-2015 dengan jumlah gerombol optimum sebanyak 7 gerombol.

E. Deskripsi Gerombol Optimum Metode K-Error

Jumlah gerombol optimum yang dihasilkan dengan Metode K-Error yaitu 7 gerombol. Pada Tabel 4, disajikan karakteristik dari masing-masing gerombol yang terbentuk berdasarkan nilai rata-rata terkoreksi. Tingkat produktivitas tanaman pangan dapat ditentukan berdasarkan nilai rata-rata setiap tanaman pangan. Gerombol yang dikategorikan sangat tinggi yaitu gerombol 5 karena memiliki nilai rata-rata produktivitas tertinggi untuk seluruh tanaman



Gambar 6. Perbandingan K-Error ragam peragam keseluruhan dengan K-Means

Tabel IV  
NILAI RATA-RATA TERKOREKSI PEUBAH PADA SETIAP GEROMBOL METODE K-Error

Gerombol	X1	X2	X3	X4	X5	X6	X7
1	35.85	13.08 <sup>b</sup>	8.57	12.88 <sup>c</sup>	51.87 <sup>b</sup>	119.09 <sup>b</sup>	155.06
2	23.96	9.86	9.29	10.57	36.31	90.76	135.66
3	32.60	10.79	5.15	11.47	28.44	78.66	143.07
4	46.63 <sup>b</sup>	11.95	9.88	11.56	47.05 <sup>c</sup>	106.90	241.44 <sup>a</sup>
5	52.87 <sup>a</sup>	13.53 <sup>a</sup>	11.57 <sup>b</sup>	14.52 <sup>a</sup>	54.47 <sup>a</sup>	175.57 <sup>a</sup>	220.88 <sup>b</sup>
6	41.79 <sup>c</sup>	12.48 <sup>c</sup>	12.09 <sup>a</sup>	13.79 <sup>b</sup>	45.46	113.09 <sup>c</sup>	151.34
7	25.10	11.57	10.69 <sup>c</sup>	12.29	38.22	98.81	160.08 <sup>c</sup>

<sup>a</sup> Rata-rata tertinggi ke-1; <sup>b</sup> Rata-rata tertinggi ke-2; <sup>c</sup> Rata-rata tertinggi ke-3

pangan. Untuk tingkat produktivitas gerombol lainnya dapat juga ditentukan dari nilai rata-rata yang disajikan pada Tabel 4.

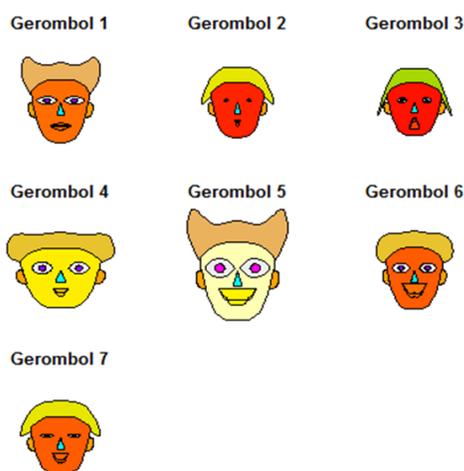
Pada Tabel 5, disajikan tingkat produktivitas untuk masing-masing gerombol beserta anggota gerombolnya. Gerombol 5 memiliki karakteristik yaitu memiliki anggota provinsi-provinsi yang memiliki produktivitas tinggi untuk semua jenis tanaman pangan. Gerombol 6 termasuk dalam kategori produktivitas sedang karena hanya memiliki rata-rata produktivitas tinggi pada tanaman kacang hijau (X3). Untuk gerombol 2 dan 3 merupakan gerombol yang dicirikan dengan produktivitas rendah untuk sebagian besar tanaman pangan.

Karakteristik untuk masing-masing gerombol disajikan dalam bentuk *suffer face plot* yang disajikan pada Gambar 7. *Suffer face plot* tersebut menggam-

Tabel V  
ANGGOTA DAN PRODUKTIVITAS HASIL GEROMBOL METODE  
K-Error

Gerombol	Produktivitas	Anggota Gerombol
1	sedang	DIY, Banten, Bali, NTB, Sulawesi Tengah
2	sangat rendah	Riau, Bengkulu, Kep.Riau, NTT, Sulawesi Tenggara, Papua Barat, Papua
3	sangat rendah	Kep. Bangka Belitung, Kalimantan Barat, Kalimantan Tengah
4	tinggi	Sumatera Utara, Lampung
5	sangat tinggi	Sumatera Barat, Jawa Barat, Jawa Tengah, Jawa Timur
6	sedang	Aceh, Jambi, Sumatera Selatan, Kalimantan Selatan, Sulawesi Utara, Sulawesi Selatan, Gorontalo, Sulawesi Barat
7	rendah	Kalimantan Timur, Maluku, Maluku Utara

barkan rata-rata masing-masing tanaman pangan (X1-X7). Gerombol-gerombol yang dihasilkan dengan metode K-Error memiliki *suffer face plot* yang berbeda-beda untuk masing-masing gerombol sehingga dapat disimpulkan bahwa gerombol-gerombol yang terbentuk memiliki karakteristik yang khas dalam hal tingkat produktivitas tanaman pangan.



Gambar 7. *Suffer face plot* rata-rata masing-masing tanaman pangan hasil gerombol K-Error

## V. SIMPULAN

Pengerombolan provinsi di Indonesia berdasarkan produktivitas tanaman pangan dengan metode K-Error dan K-Means menghasilkan hasil gerombol yang berbeda. Metode K-Error lebih tepat dibandingkan dengan metode K-Means dalam pengerombolan provinsi di Indonesia berdasarkan produktivitas tanaman pangan. Jumlah gerombol optimum yang dihasilkan dengan metode K-Error sebanyak 7 gerombol. Gerombol 5 merupakan gerombol yang terdiri dari provinsi-provinsi yang memiliki produktivitas tinggi untuk seluruh tanaman pangan. Gerombol 2 dan 3 terdiri atas provinsi yang memiliki tingkat produktivitas yang sangat rendah. Gerombol 5 memiliki produktivitas rendah pada kacang hijau (X3), padi (X5), dan ubi jalar (X6). Untuk gerombol 7 memiliki ciri produktivitas jagung (X1), kacang tanah (X2), kedelai (X4), dan ubi kayu (X7) yang sangat rendah.

## DAFTAR PUSTAKA

- Alfina, T., B. Santosa, and A. R. Barakbah (2012). Analisa perbandingan metode hierarchical clustering, k-means dan gabungan keduanya dalam cluster data (studi kasus: Problem kerja praktek teknik industri its). *Jurnal Teknik ITS* 1(1), A521–A525.
- Kalkstein, L. S., G. Tan, and J. A. Skindlov (1987). An evaluation of three clustering procedures for use in synoptic climatological classification. *Journal of climate and applied meteorology* 26(6), 717–730.
- Kumar, M. and N. Patel (2005). *Clustering Data With Measurement Errors*. New Jersey(US): Rutor Research Report of Rutgers University.