

PENGELOMPOKAN PRODUKSI PADI NASIONAL DENGAN PENDEKATAN DATA MINING KONSEP K-MEANS

CLUSTERING NATIONAL RICE PRODUCTION IN MINING DATA APPROACH CONCEPT OF K-MEANS

Oleh :

Irfan Sudono*)✉, Widya Utami**), Santi Lestari**)

*)Peneliti di Balai Irigasi, Pusat Litbang Sumber Daya Air, Kementerian Pekerjaan Umum

**) Calon Peneliti di Balai Irigasi, Pusat Litbang Sumber Daya Air, Kementerian Pekerjaan Umum

✉Komunikasi penulis, email : irfansudono@yahoo.com

Naskah ini diterima pada 10 September 2013; revisi pada 16 September 2013;
disetujui untuk dipublikasikan pada 26 September 2013

ABSTRACT

One indicator of national food security is rice production, generated from the crop area in irrigation area. Influencing policy is regulation and changes in strategic environment, which are both synergized in the form of irrigation system. JICA Study - FIDP, 1993, indicating that the development of irrigated area will move to region Sumatra, Kalimantan, and Maluku - Papua, but based on record of data production on the last 20 years the regions remains low. This study aims to find solutions the expectations of increasing rice production by analyzing which potential areas have to increase in production. Data Mining approach concept K-means method used in clustering to analyze the province in national rice production, analysis based on the 20 years data record rice production from 1993 to 2012 consists of 33 provinces whole of Indonesia, with the object of observation on the average production to the increase production (slope) and forecast production for year 2013. Based on result of simulation optimization with the K-Means obtained that rice production development sequence consisting of six clusters. Highest national rice production is still dominated by Java and Bali (group 1), meaning that the development in this region are optimally exploit the potential and more effort is needed to maintain existing area by preventing land use change, but due to limited space (potential area development only 62000 Ha), the development of irrigation areas is more rational prioritized in groups 2, namely region Sulawesi, NTB, and West Sumatera. Simulation optimization has accommodate history of production of the dimensions past, current and future production.

Keywords : national rice production, irrigated area, estimated production, K-Means, data mining

ABSTRAK

Salah satu indikator dari ketahanan pangan nasional adalah produksi beras yang dihasilkan dari luas areal panen seluruh daerah irigasi. Kebijakan yang mempengaruhinya adalah peraturan/perundangan dan perubahan lingkungan strategis, yang keduanya bersinergi dalam bentuk sistem irigasi. Studi JICA-FIDP 1993 mengindikasikan perkembangan luas area irigasi akan berpindah ke Pulau Sumatera, Kalimantan, dan Maluku-Papua, namun rekaman data produksi 20 tahun terakhir pada wilayah tersebut masih rendah. Kajian ini bertujuan untuk mencari solusi atas harapan peningkatan produksi beras dengan menganalisis daerah mana saja yang mempunyai potensi peningkatan atau percepatan produksi. Pendekatan data Mining Konsep K-means menjadi metode dalam menganalisis pengelompokan propinsi pada produksi padi Nasional, analisis tersebut berdasarkan rekaman data produksi padi dari tahun 1993 sampai dengan 2012 atau selama 20 tahun dari 33 Propinsi di seluruh Indonesia, dengan objek pengamatan pada rata-rata produksi terhadap peningkatan produksi (*slope/kemiringan*) dan perkiraan (*forecast*) produksi tahun 2013. Berdasarkan simulasi hasil optimasi dengan K-Means didapatkan urutan pengembangan produksi padi yang terdiri dari enam kelompok. Produksi tertinggi padi nasional masih di dominasi oleh Pulau Jawa dan Bali (kelompok 1), artinya pengembangan pada wilayah ini memanfaatkan potensi seoptimal mungkin dan diperlukan upaya lebih untuk mempertahankan luas area dengan mencegah alih fungsi lahan, namun karena keterbatasan lahan (potensi pengembangan hanya 62000 Ha), maka pengembangan area irigasi lebih rasional diprioritaskan pada kelompok 2 yaitu Wilayah Sulawesi, NTB, dan Sumatera Barat. Simulasi hasil optimasi tersebut sudah mengakomodir riwayat produksi pada dimensi masa lampau, saat ini dan perkiraan produksi mendatang berdasarkan peningkatan yang terjadi.

Kata Kunci: produksi padi nasional, lahan irigasi, perkiraan produksi, K-Means, data mining

I. PENDAHULUAN

1.1. Latar Belakang

Lebih dari tiga puluh tahun sudah beras ditempatkan sebagai komoditas utama dalam perekonomian Indonesia (Kasryno dan Pasandaran, 2004) dan karena sampai saat ini pun beras tetap menempati komoditas utama artinya sudah menginjak empat dasawarsa. Berbagai upaya telah dilakukan Pemerintah untuk memberikan dukungan peningkatan produktivitas petani padi dan produksi beras nasional yang meliputi: (1) pengembangan infrastruktur untuk mendukung usaha tani padi; (2) peningkatan akses petani pada sarana produksi dan sumber permodalan; (3) peningkatan mutu intensifikasi usaha padi menggunakan teknologi maju; (4) ekstensifikasi lahan pertanian dilahan kering, rawa pasang surut, lebak, dan daerah bukaan baru; dan (5) peningkatan akses petani terhadap sarana pengolahan pascapanen dan pemasaran, sementara itu ahli-ahli pertanian menyatakan tiga pesan dalam tantangan ekonomi padi dan perberasan yaitu: (1) Pendekatan berspektrum luas, (2) Pendekatan perhatian pada kelestarian lingkungan dan keseimbangan ekosistem, serta (3) Pendekatan berbasis ilmu pengetahuan dan teknologi. (Kasryno, 2004).

Memperhatikan potret irigasi Nasional (Kepmen PU No. 390 Tahun 2007) bahwa luas irigasi 7.47 Juta Ha dengan komposisi berada di Pulau Jawa sebesar 42 %, Sumatera 27%, Sulawesi 14%, dan sisanya di pulau-pulau lain kurang dari 3 % (Buku Pintar Irigasi, Direktorat Irigasi, 2010) dan informasi konversi lahan terutama di Pulau Jawa menimbulkan kekhawatiran penurunan produksi padi karena tekanan pertambahan penduduk. Hasil penelitian menunjukkan bahwa dampak pencetakan sawah baru lebih rendah dari dampak konversi lahan yang mengakibatkan penurunan kapasitas produksi pada periode yang sama (Irawan, 2004)

Salah satu indikator dari ketahanan pangan nasional adalah produksi beras yang dihasilkan dari luas areal panen seluruh daerah irigasi, kebijakan yang mempengaruhinya adalah peraturan/perundangan dan perubahan lingkungan strategis yang keduanya bersinergi dalam bentuk sistem irigasi. Keberlanjutan sistem irigasi ini dibagi dalam dua program kebijakan (PP No. 20 Tahun 2006 Tentang Irigasi) yaitu (1) Pengembangan irigasi yang terdiri dari peningkatan (*upgrading*), lanjutan (*rounding up*) dan pembangunan baru; (2) Pengelolaan irigasi

meliputi operasi pemeliharaan dan rehabilitasi jaringan irigasi. Sedangkan pengembangan system irigasi ditegaskan dalam lima pilar pengembangan irigasi (Sastrohardjono, 2011), yaitu: (1) Peningkatan keandalan penyediaan air irigasi, (2) Perbaikan sarana dan prasarana irigasi, (3) Penyempurnaan system pengelolaan irigasi, (4) Penguatan institusi pengelola irigasi, (5) Pemberdayaan sumber daya manusia pengelola irigasi. (periksa PP No. 20/2006 tentang Irigasi pasal 1, butir 4).

Upaya struktural ketahanan pangan nasional berupa peningkatan areal panen daerah irigasi diantaranya adalah dukungan Direktorat Irigasi Dirjen Sumber Daya Air melalui program rehabilitasi, Operasi dan Pemeliharaan irigasi pusat, dan Pembangunan lanjutan telah menghasilkan perluasan areal panen dari 11.893.000 Ha pada tahun 2005 menjadi 13.472.000 Ha pada tahun 2012. (BPS 2013 dan Buku pintar irigasi untuk data hingga 2010)

1.2. Tujuan

Kajian ini bertujuan untuk mencari solusi atas harapan peningkatan produksi beras dengan menganalisis daerah mana saja yang mempunyai potensi peningkatan atau percepatan produksi. Analisis dilakukan dengan pengelompokan propinsi pada produksi padi Nasional dengan pendekatan Data Mining Konsep K-Means, berdasarkan rekaman data produksi padi dari tahun 1993 sampai dengan 2012 atau selama 20 tahun dari 33 Propinsi di seluruh Indonesia, dengan objek pengamatan pada rata-rata produksi terhadap peningkatan produksi (*slope/kemiringan*) dan perkiraan produksi tahun 2013.

II. KAJIAN PUSTAKA

2.1. Data Mining

Data mining adalah proses untuk menemukan pola atau pengetahuan yang bermanfaat secara otomatis dari sekumpulan data yang berjumlah banyak, data mining sering dianggap sebagai bagian dari *Knowledge Discovery in Database* (KDD) yaitu sebuah proses mencari pengetahuan yang bermanfaat dari data, proses KDD secara garis besar dilakukan dalam bentuk (1) Pemilihan Data (*Data Selection*), (2) Pra-proses (*Pre-processing/Cleaning*), (3) Transformasi (*Transformation*), (4) Sajian yang mudah dimengerti dan Evaluasi (*Interpretation/Evaluation*) (Sujana, 2010).

Pada dasarnya Data Mining merupakan rangkaian perlakuan mulai dari pemilihan data dari database yang terpilih, pemeriksaan data yang inkonsisten, duplikasi, memperkaya data (dengan informasi lain bila perlu untuk memperkuat KDD), Coding yang kreatif tergantung jenis dan pola, dan proses sesuai metode dan muara nya adalah pola informasi yang dihasilkan perlu disajikan dalam bentuk yang mudah dimengerti oleh para pihak yang terlebih dahulu diperiksa apakah pola atau informasi ini bertentangan dengan fakta atau hipotesa yang ada sebelumnya.

2.2. Konsep K-Means

K-Means merupakan salah satu metode pengelompokan data non-hierarkhi, berusaha mempartisi data yang ada dalam bentuk dua atau lebih kelompok. Metode ini akan mempartisi data dalam kelompok sehingga data yang berkarakteristik sama dimasukkan dalam kelompok yang sama dan sebaliknya untuk kelompok berbeda. Adapun tujuan dari pengelompokan adalah meminimalkan fungsi objektif yang diset dalam pemrosesan, secara sederhana berusaha meminimalkan variasi dalam suatu kelompok dan memaksimalkan variasi antar kelompok (Tan, 2006)

K-Means adalah metode analisis kelompok mengarah pada pemartisian N objek pengamatan menjadi K kelompok (Cluster), dimana setiap objek pengamatan diarahkan pada kelompok dengan mean (rata-rata) terdekat (Prasetyo, 2012)

2.3. Kemiripan dan Ketidak miripan

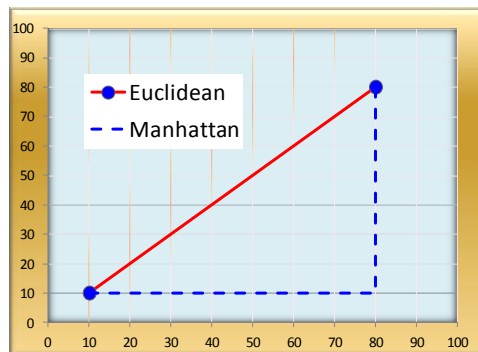
Istilah ketidakmiripan antara dua data dalam data mining disebut sebagai jarak antara dua data tersebut, sehingga jika s ukuran kemiripan dan d ukuran ketidakmiripan maka $s + d = 1$. Pada umumnya dalam klasifikasi (pengelompokan data) disarankan nilai interval ketidak miripan data ditransformasikan dalam interval yang ternormalisasi $[0,1]$, atau agar nilainya tampak lebih mudah dapat dituliskan dalam $[0,100]$. (Prasetyo, 2012)

Adapun perlakuan yang dilakukan adalah perubahan pengelompokan setelah pasangan data dibuat dalam bentuk standar besaran yang sama melalui transformasi

$$Y_i = \frac{(Y_{O_i} - Y_{min})}{(Y_{max} - Y_{min})} \times 100, \dots \dots \dots \text{Pers. 1}$$

$$X_i = \frac{(X_{O_i} - X_{min})}{(X_{max} - X_{min})} \times 100, \dots \dots \dots \text{Pers. 2}$$

Terdapat beberapa metode untuk mengukur ketidakmiripan data yaitu dengan cara Euclidean atau Manhattan lihat Gambar 1.



Gambar 1 Jarak dua data dalam dua dimensi

Pengukuran Jarak dengan Euclidean (Bezdek, 1981) :

$$D_{ij} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}, \dots \dots \dots \text{Pers. 3}$$

sedangkan pengukuran jarak Manhattan (Miyamoto, Agus 1995),

$$D_{ij} = \sqrt{(X_i - X_j)^2} + \sqrt{(Y_i - Y_j)^2} \dots \dots \dots \text{Pers. 4}$$

Metode lain yang dapat digunakan adalah metode Minkowsky dengan formula

$$D_{ij} = \sqrt[\lambda]{(X_i - X_j)^\lambda + (Y_i - Y_j)^\lambda} \dots \dots \dots \text{Pers. 5}$$

2.4. Titik Pusat atau Centroid masing-masing kelompok

Titik pusat Centroid adalah titik pilihan sedemikian rupa sehingga menghasilkan jumlah jarak terdekat dengan kelompoknya, umumnya berupa rata-rata koordinat Kelompok K. Nilai rata-rata fitur X dari seluruh anggota kelompok K dan nilai rata-rata fitur Y dari seluruh anggota kelompok K selanjutnya akan dinyatakan dalam (CX_j, CY_j) diperoleh dengan formula berikut.

$CX_j = \text{Average}(X_{ij})$, dan $CY_j = \text{Average}(Y_{ij})$, $j = 1$ sampai K, dan i dari 1 sampai K_i

Atau

$$CX_j = \sum_{i=1}^{K_j} \frac{X_{ij}}{K_j}, \quad j = 1 \text{ sampai } K, \dots \dots \dots \text{Pers. 6}$$

$$CY_j = \sum_{i=1}^{K_j} \frac{Y_{ij}}{K_j}, \quad j = 1 \text{ sampai } K, \dots \dots \dots \text{Pers. 7}$$

Keterangan :

- K : Jumlah kelompok
- K_j : Jumlah anggota kelompok ke-J

2.5. Fungsi Objektif

Fungsi Objektif atau fungsi tujuan adalah fungsi sasaran untuk mendapatkan nilai optimal, umumnya dalam bentuk nilai maksimum suatu fungsi atau minimum suatu fungsi (Taha 2010).

Adapun yang akan digunakan dalam Fungsi Objektif Data Mining Konsep K-Means adalah Minimum Jumlah jarak dari seluruh anggota kelompok terhadap titik-titik pusatnya (centroid)

$$\text{Min } F = \sum_{j=1}^K \sum_{i=1}^{K_j} \sqrt{(X_{ij} - CX_j)^2 + (Y_{ij} - CY_j)^2}, \dots \text{ Pers. 8}$$

2.6. Ambang Batas

Ambang batas (threshold) adalah batas untuk menentukan simulasi berhenti yang dinyatakan dengan besaran suatu nilai. Setiap akhir simulasi dihitung harga mutlak perbedaan antara Fungsi Objektif Simulasi dengan Fungsi Objektif sebelumnya (Prasetyo, 2012), jika nilai ini kurang dari nilai ambang batas (Threshold) maka simulasi dihentikan. Ambang batas untuk pengelompokan data produksi padi ini ditentukan nilai Thresholdnya adalah 0.01 (10^{-2})

2.7. Kemiringan (Slope)

Simulasi Data mining pendekatan K-Means akan diujikan menggunakan set data dua dimensi untuk itu memerlukan satu variabel independent untuk mendampingi variabel dependent yang dalam hal ini telah ditetapkan berupa rata-rata produksi padi (Ton/Ha). Berdasarkan pengamatan secara visual dan statistik dasar bahwa perubahan/peningkatan produksi yang berbeda-beda dikarenakan tingkat kemiringan (slope) yang berbeda untuk itu dipilih Slope sebagai variabel bebas.

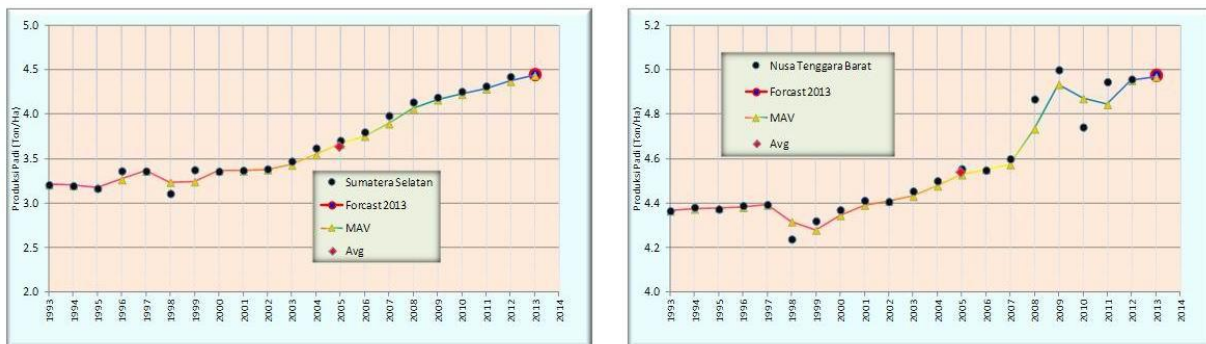
Slope diperoleh dari analisis regresi linear produksi padi masing-masing propinsi pada seluruh tahun dan dinyatakan dalam regresi linear, nilai gradien (faktor tangent sudut) yang

dibentuk ini yang dipilih sebagai slope. Atau koefisien a dari persamaan $Y = aX + b$, diperoleh dengan memanfaatkan fungsi excell SLOPE(range cell known_y's, range cell known_x's). Adapun nilai statistik dasar produksi padi nasional untuk analisis data periode tahun 1993-2012 dapat dilihat pada Tabel 2 kolom 6, sedangkan untuk periode 2001-2012 pada kolom 10.

2.8. Perkiraan Produksi 2013

Variabel dependen lain yang dipilih adalah prediksi padi pada tahun 2013, berdasarkan forecast menggunakan data produksi sebelumnya dengan anggapan bahwa tidak terjadi kejadian luar biasa yang dapat mengubah produksi seperti rekaman produksi sebelumnya (Makridakis, 1999), untuk rujukan nilai prediksinya menggunakan fungsi formula dari Excell $Y = \text{FORECAST}(\text{nilai } x, \text{range cell known_y's, range cell known_x's})$, dengan panjang data terlebih dahulu memilih data dari tahun terakhir mundur sampai tahun berapa dimana diperoleh nilai korelasi tinggi sehingga jika sampai tahun 2003 berarti 10 data, tahun 2001 berarti 12 data, 1993 berarti 20 data sehingga masing-masing propinsi dapat berbeda panjang data yang dipilih.

Nilai forecast yang diperoleh tidak langsung diterima, tetapi terlebih dahulu divalidasi dengan moving average apakah terjadi perbedaan signifikan antara prediksi (2013) dengan data sebelumnya (2012) lebih dari tingkat kemiringan lima tahun terakhir, jika terjadi perlakuan bahwa prediksi akhir (yang akan dipilih) adalah setengah dari penjumlahan antara prediksi awal dan data tahun 2012, lihat lagi moving averagenya dan lakukan hal yang sama. Sebagai ilustrasi berikut ini disajikan nilai forecast 2013 untuk propinsi Sumatera Selatan dan Nusa Tenggara Barat, hasil seluruh forecast 2013 produksi padi disajikan pada Tabel 2, Bersandingan dengan produksi tahun 2012.



Gambar 2 Validasi Forecasting 2013 berdasarkan data runtut waktu dan Moving average

III. METODE PENELITIAN

3.1. Perlakuan Pra Simulasi Data Mining Konsep K-Means

Sebelum melakukan simulasi Data Mining terlebih dahulu dilakukan analisis sederhana seperti statistik dasar, penentuan derajat kemiringan (Slope), dan peramalan (forecast) untuk produksi tahun 2013 berdasarkan rekaman data 1993 sampai dengan 2012.

Penentuan variabel tidak bebas juga merupakan bagian penting dari analisis, pemilihan variabel yang tepat akan menentukan ketepatan pengelompokan data, telah dilakukan berbagai pemilihan variabel tidak bebas seperti produksi tahun 2012 (tahun terakhir data), rata-rata lima tahun terakhir, dan lain-lain yang akhirnya untuk mengelompokan data dengan mempertimbangkan produksi masa lalu (lebih sepuluh tahun kebelakang), saat ini (lima tahun terakhir) dan perkiraan mendatang dipilih adalah nilai kemiringan (Slope) dan perkiraan produksi tahun 2013 dengan dasar variabel bebas berupa rata-rata periode 1993-2012 dan periode 2001-2012.

3.2. Prinsip Kerja K-Means

K-Means adalah metode analisis kelompok mengarah pada pemartisian N objek pengamatan menjadi K kelompok (Cluster), dimana setiap objek pengamatan diarahkan pada kelompok dengan mean (rata-rata) terdekat (Prasetyo, 2012), berikut ini Algoritma K-Means pada pengelompokan data produksi padi.

1. Kumpulkan data dan tentukan fiturnya dalam dua dimensi

Kumpulkan seluruh data katakan terdiri dari N data, Pada setiap data tentukan nilai fitur X dan nilai fitur Y, sehingga diperoleh pasangan data $(X_1, Y_1), (X_2, Y_2), \dots, (X_n, Y_n)$

2. Tentukan Jumlah Kelompok dan anggota masing-masing kelompok

Bagi data dalam kelompok, misalkan K kelompok dan ambil anggotanya secara acak sehingga diperoleh:

Kelompok 1	Jumlah anggota	K1
Kelompok 2	Jumlah anggota	K2
.....
.....
.....
Kelompok K	Jumlah anggota	KK

Perhatikan bahwa: $K_1 + K_2 + K_3 + \dots + K_K = N$
 Sehingga fitur X dan Y masing-masing anggota dinyatakan sebagai berikut:

(X_{ij}, Y_{ij}) , $i = 1$ sampai K, dan j untuk tiap i adalah dari 1 sampai K_i , untuk lebih jelas:

- Fitur Anggota Kelompok 1 adalah;
 $(X_{11}, Y_{11}), (X_{12}, Y_{12}), (X_{13}, Y_{13}) \dots (X_{1k_1}, Y_{1k_1})$
- Fitur Anggota Kelompok 2 adalah;
 $(X_{21}, Y_{21}), (X_{22}, Y_{22}), (X_{23}, Y_{23}) \dots (X_{2k_2}, Y_{2k_2})$

- Fitur Anggota Kelompok K adalah;
 $(X_{j1}, Y_{j1}), (X_{j2}, Y_{j2}), (X_{j3}, Y_{j3}) \dots (X_{jk}, Y_{jk})$

3. Hitung pusat masing-masing kelompok Centroid atau rata-rata Kelompok K adalah nilai rata-rata fitur X dari seluruh anggota kelompok K dan nilai rata-rata fitur Y dari seluruh anggota kelompok K

$CX_j = \text{Average}(X_{ij})$, $j = 1$ sampai K, dan i adalah dari 1 sampai K_i

$CY_j = \text{Average}(Y_{ij})$, $j = 1$ sampai K, dan i adalah dari 1 sampai K_i

Pusat atau centroid masing-masing kelompok adalah (CX_k, CY_k)

4. Alokasikan masing-masing data ke Centroid/Rata-rata terdekat.

Hitung jarak tiap data terhadap Centroid masing-masing kelompok.

5. Periksa apakah ada data yang jarak ke Centroid/Rata-rata lebih dekat dari Centroid kelompok sebelumnya (misal data dari kelompok 1 ternyata setelah dicari jarak terhadap masing-masing centroid kelompok, data tersebut paling dekat ke Centroid kelompok 3) , jika Ya berikan tanda dan pindahkan data tersebut dari kelompok sebelumnya menjadi anggota kelompok baru.

6. Hitung Nilai Fungsi Objektifnya, periksa perubahan nilai objektif dengan nilai objektif sebelumnya apakah masih diambang batas yang ditentukan

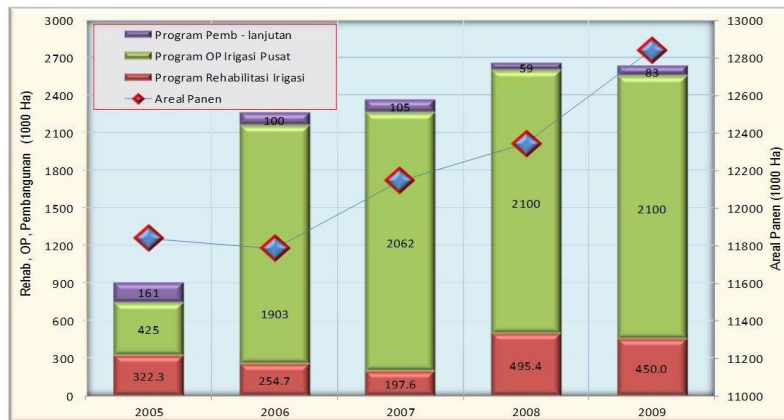
7. Ulangi (loop) ke langkah ke-3 lagi dan nyatakan sebagai iterasi, sampai dengan langkah 5 tidak terdapat data yang pindah kelompok dan langkah ke-6 memenuhi ambang batas. Proses selesai.

IV. HASIL DAN PEMBAHASAN

4.1. Kondisi luas area irigasi saat ini dan perkiraan mendatang

Bersumber dari Studi *Japan Interntional Cooperation Agency – Formulation of Irrigation Development Program* (JICA-FIDP) dan informasi luas irigasi masing-masing Balai Besar/Balai Wilayah Sungai (BBWS/BWS), Direktorat Irigasi Dirjen SDA menetapkan luas area Irigasi di Indoensia (2010) kurang lebih 7.47 Juta Ha, pada saat itu luas area panen 13.253 Juta Ha dan

produksi 66.47 Juta Ton atau jika di rata-rata kan produksi padi Nasional adalah 5.015 Ton/Ha, (Buku pintar irigasi, 2010) dan intensitas tanam 1.78, berbagai usaha dilakukan pemerintah dalam hal ini Kementerian Pekerjaan Umum, melalui Direktorat Irigasi dan Rawa, Dirjen Sumber Daya Air telah mendukung melalui program Operasi dan Pemeliharaan Irigasi Pusat, Rehabilitasi Irigasi, dan Program Pembangunan Lanjutan, serta memberikan dampak peningkatan luas area panen, seperti dapat dilihat pada Gambar 3.

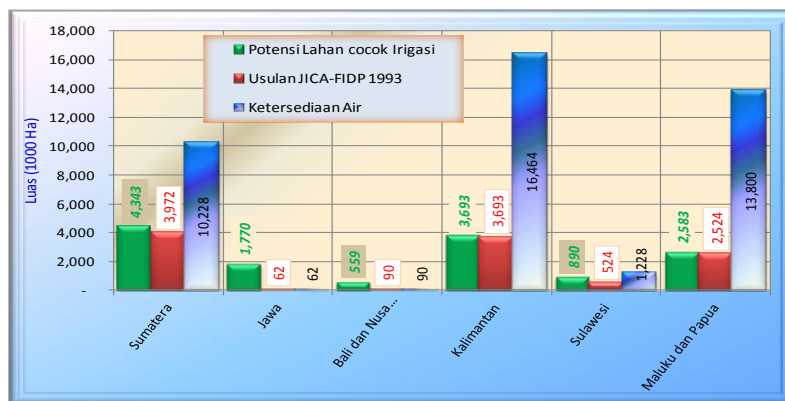


Sumber : Buku pintar Irigasi 2010, Direktorat Irigasi-Dirjen SDA

Gambar 3 Dukungan Direktorat Irigasi Dirjen SDA terhadap areal panen produksi beras 2005-2009

Studi JICA-FIDP 1993 yang merupakan studi koordinasi antar lembaga yaitu Kementerian Pekerjaan Umum dan Bappenas dibantu *expert* dari Jepang (JICA) mengusulkan pengembangan areal irigasi di Indonesia tahun 2020 sebesar 10.865 Juta Ha untukantisipasi kebutuhan pangan karena peningkatan jumlah penduduk dengan skala moderat. Potensi pengembangan dipertimbangkan dari potensi ketersediaan air dan potensi/kecocokan lahan untuk area irigasi sebagai bagian dari 8 syarat pengembangan

irigasi yang digariskan Ditjen Sumber Daya Air, disamping itu dipertimbangkan pula 6 syarat lainnya: (1) ketersediaan petani, (2) banjir tidak sulit diatasi, (3) assesmen mudah, (4) terdapat daerah pemasaran, (5) pembebasan lahan tidak sulit, (6) sinkron dengan program prioritas pemerintah. Potensi tersebut disajikan pada Gambar 4, adapun usulan pengembangan areal irigasi FIDP-JICA 2010 seperti disajikan pada tabel 1.



Gambar 4 Potensi berdasarkan ketersediaan air dan lahan, serta usulan JICA-FIDP 1993

Tabel 1 Usulan Pengembangan Areal Irigasi FIDP-JICA 93

Wilayah	2010		2020
	Luas Area Irigasi	Persentase	Potensi Pengembangan
Sumatera	1,988,844	26.63	3,972,000
Jawa	3,157,394	42.27	62,000
Bali	143,008	1.91	90,000
Nusa Tenggara	486,306	6.51	
Kalimantan	480,320	6.43	3,693,000
Sulawesi	1,021,921	13.68	524,000
Maluku dan Papua	192,004	2.57	2,524,000
	7,469,797	100.00	10,865,000

Sumber : Buku pintar Irigasi 2010, Direktorat Irigasi-Dirjen SDA

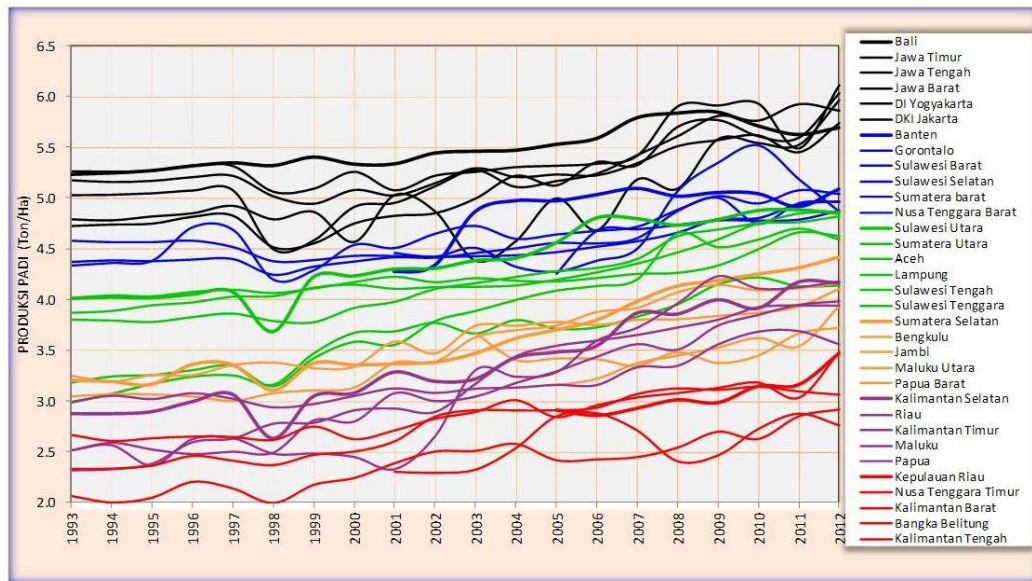
Apabila pada tahun 2010 menunjukkan bahwa sebaran area irigasi masih dominan di Pulau Jawa sebesar 42.3% (3.157 Ha), maka merujuk hasil studi JICA-FIDP (Formulation of Irrigation Program 1993) bahwa potensi pengembangan Irigasi di Indonesia sampai dengan pada tahun

2020 dan seterusnya akan berada di Sumatera, Kalimantan, dan Maluku Papua.

Pertanyaannya siapkah kita dan bagaimana kelanjutan produksinya, karena menurut rekaman data produksi 20 tahun pada wilayah tersebut produksi masih lebih rendah dari produksi di Jawa, terutama Kalimantan dan Maluku-Papua yang dalam pengurutan produksi berada jauh dibawah diproduksi rata-rata nasional. Nusa Tenggara-Bali dari sudut produksi lebih baik tetapi potensi sudah maksimal dikembangkan, terutama karena keterbatasan lahan, seperti dapat dilihat pada gambar 4.

4.2. Rekaman Produksi Padi Tahun 1993-2012

Sebelum melakukan analisis perlu dituliskan terlebih dahulu data masukan yang akan digunakan dalam analisis K-Means, seri data luas dan produksi padi dari seluruh propinsi di Indonesia dari tahun 1993 sampai dengan 2012 yang bersumber dari BPS 2013 akan digunakan sebagai masukan. Data produksi padi ini merupakan penyederhanaan bentuk dari data luas panen (Ha), berbeda dengan luas area irigasi dan produksi (Ton) tiap propinsi selama dua puluh tahun terakhir.



Sumber: BPS 2013, (www.bps.go.id diakses : 27 Februari 2013)

Gambar 5 Produksi Padi Nasional perpropinsi tahun 1993-2012

Berdasarkan Gambar 5, terlihat jelas bahwa terjadi kenaikan produksi padi disemua propinsi dengan tingkat kenaikan (kemiringan yang

berbeda), setelah melalui analisis sederhana perbedaan rata-rata akhirnya untuk analisis lanjut akan di buat dalam dua bagian menggunakan

seluruh data 1993-2012 dan data dengan rentang 2001-2012 karena pada tahun ini kecenderungan terjadi perubahan produksi yang signifikan. Perbedaan dua kelompok besar ini dapat dilihat pada tabel 2.

4.3. Kasus Rata-rata dan Slope, skenario data 1993-2012

Simulasi optimasi pertama (Kasus-01) dilakukan pengelompokan 33 data produksi padi dari seluruh propinsi di Indonesia, menggunakan set data dua dimensi hal ini untuk memudahkan visualisasi dalam koordinat Cartesius, fitur yang digunakan dalam pengelompokan adalah X dan Y, adapun masing-masing variabel ini adalah :

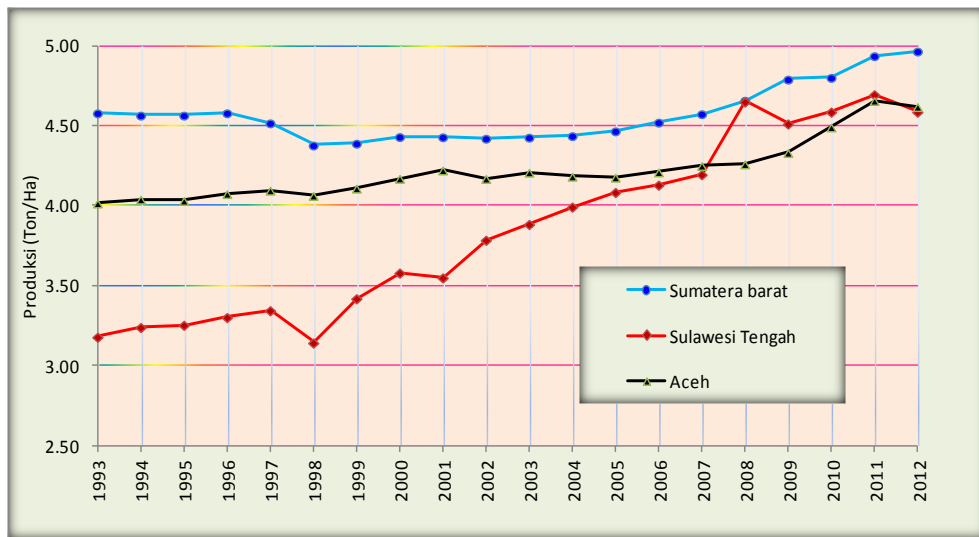
X: Rata-rata besaran produksi padi dari tahun 1993 sampai dengan 2012 dalam unit (Ton/Ha); (kolom 3, pada Tabel statistik dasar produksi padi nasional).

Y: Nilai slope (kemiringan) produksi padi dari tahun 1993 sampai dengan 2012 tidak bersatuan, (kolom 6, pada Tabel statistik dasar produksi padi nasional).

Mengapa dua variabel tersebut yang dipilih adalah karena pertimbangan perubahan posisi urutan kurva produksi seperti secara rata-rata dapat saja suatu propinsi mempunyai produksi relatif tinggi tetapi rekaman data terakhir 2012 tidak menunjukkan perbedaan signifikan dengan data tahun sebelumnya (Aceh) dan ini dapat dilihat bahwa nilai slopenya rendah, sebaliknya ada propinsi dengan produk rata-rata medium tetapi lima data terakhir menunjukkan peningkatan yang signifikan seperti Sulawesi Tengah, demikian halnya ada propinsi dengan produksi rata-rata dan peningkatan yang relatif biasa (Sumatera Barat) seperti dapat dilihat pada Gambar 6.

Tabel 2 Statistik Dasar Produksi Padi Nasional (Ton/Ha)

No	Propinsi	1993 - 2012				2001 - 2012				2012	2013
		AVG	STDEV	SKEW	SLOPE	AVG	STDEV	SKEW	SLOPE		
	Indonesia	4.58	0.28	0.71	0.0420	4.74	0.25	0.21	0.0676	5.12	5.12
1	2	3	4	5	6	7	8	9	10	11	12
1	Aceh	4.22	0.18	1.34	0.0273	4.32	0.17	1.26	0.0411	4.62	4.64
2	Sumatera Utara	4.23	0.29	0.88	0.0468	4.39	0.28	0.50	0.0740	4.82	4.87
3	Sumatera barat	4.57	0.18	1.12	0.0183	4.62	0.21	0.66	0.0544	4.97	4.99
4	Riau	3.20	0.24	1.08	0.0355	3.33	0.24	0.55	0.0601	3.56	3.68
5	Jambi	3.54	0.45	0.15	0.0731	3.85	0.28	-0.59	0.0753	4.16	4.26
6	Sumatera Selatan	3.64	0.43	0.60	0.0685	3.88	0.38	-0.08	0.1041	4.42	4.45
7	Bengkulu	3.59	0.28	0.06	0.0452	3.78	0.16	-0.07	0.0398	4.10	4.06
8	Lampung	4.19	0.39	0.55	0.0632	4.43	0.31	0.13	0.0840	4.86	4.90
9	Bangka Belitung	2.59	0.23	-0.09	0.0388	2.59	0.23	-0.09	0.0388	2.76	2.80
10	Kepulauan Riau	3.06	0.19	1.40	0.0706	3.06	0.19	1.40	0.0706	3.46	3.37
11	DKI Jakarta	4.99	0.39	0.97	0.0448	5.12	0.46	0.20	0.1050	5.96	5.91
12	Jawa Barat	5.22	0.40	0.26	0.0541	5.44	0.34	0.19	0.0897	5.86	5.93
13	Jawa Tengah	5.26	0.21	0.75	0.0275	5.36	0.21	0.26	0.0544	5.73	5.70
14	DI Yogyakarta	5.10	0.44	0.62	0.0678	5.36	0.38	0.11	0.1001	6.03	5.98
15	Jawa Timur	5.41	0.31	1.25	0.0376	5.52	0.34	0.59	0.0816	6.11	6.04
16	Banten	4.89	0.28	-1.81	0.0554	4.89	0.28	-1.81	0.0554	5.08	5.17
17	Bali	5.49	0.20	0.53	0.0309	5.61	0.17	0.00	0.0363	5.69	5.75
18	Nusa Tenggara Barat	4.54	0.23	0.95	0.0343	4.66	0.22	0.40	0.0571	4.96	4.98
19	Nusa Tenggara Timur	2.86	0.24	0.87	0.0374	3.01	0.20	0.91	0.0486	3.48	3.38
20	Kalimantan Barat	2.75	0.31	-0.13	0.0503	2.97	0.16	-1.19	0.0372	3.06	3.16
21	Kalimantan Tengah	2.39	0.27	0.22	0.0430	2.58	0.17	1.02	0.0375	2.92	2.89
22	Kalimantan Selatan	3.38	0.47	0.36	0.0752	3.67	0.36	0.01	0.0961	4.16	4.20
23	Kalimantan Timur	3.16	0.58	-0.02	0.0963	3.54	0.37	-0.78	0.0992	3.94	3.99
24	Sulawesi Utara	4.40	0.36	-0.10	0.0574	4.64	0.23	-0.50	0.0597	4.85	4.93
25	Sulawesi Tengah	3.86	0.55	0.24	0.0902	4.22	0.38	-0.26	0.1009	4.59	4.72
26	Sulawesi Selatan	4.65	0.26	0.04	0.0362	4.79	0.19	0.27	0.0492	5.04	5.06
27	Sulawesi Tenggara	3.64	0.39	-0.20	0.0631	3.90	0.21	0.46	0.0516	4.14	4.21
28	Gorontalo	4.74	0.44	0.70	0.0912	4.74	0.44	0.70	0.0912	4.87	5.10
29	Sulawesi Barat	4.70	0.19	-2.22	0.0630	4.70	0.19	-2.22	0.0630	4.88	4.93
30	Maluku	3.14	0.71	0.39	0.1099	3.56	0.62	-0.80	0.1618	4.18	4.30
31	Maluku Utara	3.49	0.14	0.82	0.0146	3.49	0.14	0.82	0.0146	3.71	3.61
32	Papua Barat	3.48	0.24	0.63	0.0931	3.48	0.24	0.63	0.0931	3.93	3.86
33	Papua	3.11	0.52	0.27	0.0865	3.47	0.36	0.13	0.0968	3.98	4.01



Gambar 6 Produksi Padi Propinsi Sumatera Barat, Sulawesi Tengah dan Aceh

Proses Simulasi dimulai dengan penentuan nilai masing-masing variabel, untuk itu pilih kolom 3 dan 6 dari Tabel 2 dan karena dua variabel ini memiliki standard nilai besaran yang sangat berbeda sedang syarat perlakuan analisis Data Mining mengharuskan data dalam besaran standard yang sama, untuk itu terlebih dahulu dilakukan perubahan atau transformasi data pasangan dalam bentuk standar besaran yang sama melalui transformasi :

$$\left. \begin{aligned}
 Y_i &= \frac{(YO_i - Y_{min})}{(Y_{max} - Y_{min})} \times 100, \text{ dan} \\
 X_i &= \frac{(XO_i - X_{min})}{(X_{max} - X_{min})} \times 100,
 \end{aligned} \right\} \text{.....Pers. 9}$$

Sehingga data untuk Propinsi Aceh dengan nilai rata-rata 4.22 (X) dan Slope 0.0273 (Y) menjadi nilai baru $X = (4.22 - 2.39) / (5.49 - 2.39) * 100 = 59.04$, dan $Y = (0.0273 - 0.00146) / (0.1099 - 0.00146) * 100 = 13.35$, dan seterusnya untuk seluruh propinsi pada Tabel 2, lakukan dengan proses yang sama hasilnya dapat dilihat pada Tabel 3.

Langkah kedua adalah penentuan jumlah kelompok dan anggotanya sebagai inialisasi, Tabel 3, di bawah ini menyajikan pembagian kelompok, dimana 33 Propinsi dibagi dalam 6 kelompok untuk case-01 pembagiannya berdasarkan produksi, Kelompok 1 adalah kelompok dengan produksi padi rata-rata mulai 4.99 sampai dengan 5.49 Ton/Ha dan terdiri dari enam anggota yaitu Bali, Jawa Timur, Jawa Tengah, Jawa Barat, DIY, dan DKI Jakarta.

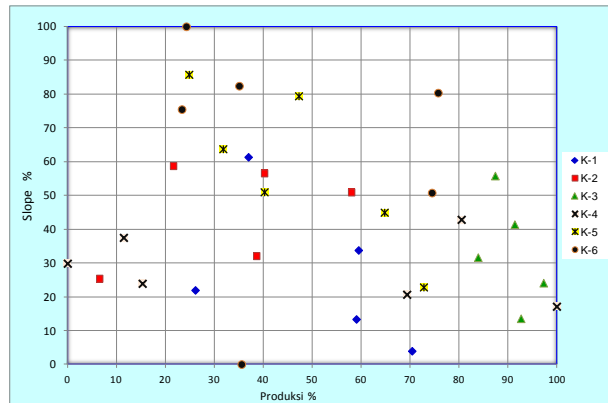
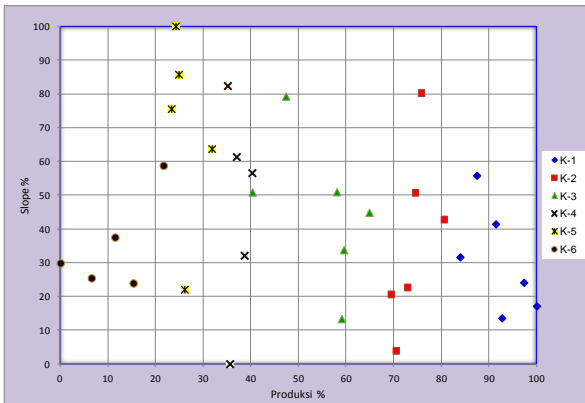
Semua kelompok ini mempunyai anggota, tiga kelompok pertama mempunyai 6 (enam) anggota dan tiga kelompok lain (4,5, dan 6) mempunyai 5 (lima) anggota, tidak usah dipikirkan mengapa pembagiannya seperti tersebut di atas, ini adalah initial saja yang akan berganti batasan dan anggota melalui proses optimasi. Sebagai pembandingan pada Case-02 tiga kelompok pertama mempunyai 5 anggota dan tiga sisanya 6 anggota, dan pembagiannya tidak berdasarkan besaran produksi tetapi urutan propinsi dari Barat ke Timur.

Tabel 3 Penyetaraan standard dengan Transformasi pada variabel X dan Y

Case 01								Case 02							
Kode	Avg	Slope	K	No	Lokasi	X	Y	Kode	Avg	Slope	K	No	Lokasi	X	Y
1	5.49	0.0309	1	17	Bali	100.00	17.12	1	4.22	0.0273	1	1	Aceh	59.04	13.35
2	5.41	0.0376	1	15	Jawa Timur	97.31	24.10	2	4.23	0.0468	1	2	Sumatera Utara	59.49	33.78
3	5.26	0.0275	1	13	Jawa Tengah	92.69	13.56	3	4.57	0.0183	1	3	Sumatera barat	70.42	3.93
4	5.22	0.0541	1	12	Jawa Barat	91.40	41.43	4	3.20	0.0355	1	4	Riau	26.11	21.95
5	5.10	0.0678	1	14	DI Yogyakarta	87.42	55.81	5	3.54	0.0731	1	5	Jambi	36.97	61.36
6	4.99	0.0448	1	11	DKI Jakarta	83.92	31.66	6	3.64	0.0685	2	6	Sumatera Selatan	40.23	56.60
7	4.89	0.0554	2	16	Banten	80.53	42.85	7	3.59	0.0452	2	7	Bengkulu	38.62	32.11
8	4.74	0.0912	2	28	Gorontalo	75.75	80.36	8	4.19	0.0632	2	8	Lampung	58.03	50.96
9	4.70	0.0630	2	29	Sulawesi Barat	74.49	50.77	9	2.59	0.0388	2	9	Bangka Belitung	6.52	25.42
10	4.65	0.0362	2	26	Sulawesi Selatan	72.87	22.72	10	3.06	0.0706	2	10	Kepulauan Riau	21.62	58.77
11	4.57	0.0183	2	3	Sumatera barat	70.42	3.93	11	4.99	0.0448	3	11	DKI Jakarta	83.92	31.66
12	4.54	0.0343	2	18	Nusa Tenggara Barat	69.38	20.67	12	5.22	0.0541	3	12	Jawa Barat	91.40	41.43
13	4.40	0.0574	3	24	Sulawesi Utara	64.85	44.87	13	5.26	0.0275	3	13	Jawa Tengah	92.69	13.56
14	4.23	0.0468	3	2	Sumatera Utara	59.49	33.78	14	5.10	0.0678	3	14	DI Yogyakarta	87.42	55.81
15	4.22	0.0273	3	1	Aceh	59.04	13.35	15	5.41	0.0376	3	15	Jawa Timur	97.31	24.10
16	4.19	0.0632	3	8	Lampung	58.03	50.96	16	4.89	0.0554	4	16	Banten	80.53	42.85
17	3.86	0.0902	3	25	Sulawesi Tengah	47.34	79.32	17	5.49	0.0309	4	17	Bali	100.00	17.12
18	3.64	0.0631	3	27	Sulawesi Tenggara	40.30	50.89	18	4.54	0.0343	4	18	Nusa Tenggara Barat	69.38	20.67
19	3.64	0.0685	4	6	Sumatera Selatan	40.23	56.60	19	2.86	0.0374	4	19	Nusa Tenggara Timur	15.31	23.91
20	3.59	0.0452	4	7	Bengkulu	38.62	32.11	20	2.75	0.0503	4	20	Kalimantan Barat	11.47	37.51
21	3.54	0.0731	4	5	Jambi	36.97	61.36	21	2.39	0.0430	4	21	Kalimantan Tengah	0.00	29.86
22	3.49	0.0146	4	31	Maluku Utara	35.57	0.00	22	3.38	0.0752	5	22	Kalimantan Selatan	31.87	63.65
23	3.48	0.0931	4	32	Papua Barat	35.09	82.41	23	3.16	0.0963	5	23	Kalimantan Timur	24.90	85.70
24	3.38	0.0752	5	22	Kalimantan Selatan	31.87	63.65	24	4.40	0.0574	5	24	Sulawesi Utara	64.85	44.87
25	3.20	0.0355	5	4	Riau	26.11	21.95	25	3.86	0.0902	5	25	Sulawesi Tengah	47.34	79.32
26	3.16	0.0963	5	23	Kalimantan Timur	24.90	85.70	26	4.65	0.0362	5	26	Sulawesi Selatan	72.87	22.72
27	3.14	0.1099	5	30	Maluku	24.27	100.00	27	3.64	0.0631	5	27	Sulawesi Tenggara	40.30	50.89
28	3.11	0.0865	5	33	Papua	23.36	75.48	28	4.74	0.0912	6	28	Gorontalo	75.75	80.36
29	3.06	0.0706	6	10	Kepulauan Riau	21.62	58.77	29	4.70	0.0630	6	29	Sulawesi Barat	74.49	50.77
30	2.86	0.0374	6	19	Nusa Tenggara Timur	15.31	23.91	30	3.14	0.1099	6	30	Maluku	24.27	100.00
31	2.75	0.0503	6	20	Kalimantan Barat	11.47	37.51	31	3.49	0.0146	6	31	Maluku Utara	35.57	0.00
32	2.59	0.0388	6	9	Bangka Belitung	6.52	25.42	32	3.48	0.0931	6	32	Papua Barat	35.09	82.41
33	2.39	0.0430	6	21	Kalimantan Tengah	0.00	29.86	33	3.11	0.0865	6	33	Papua	23.36	75.48
Max	5.49	0.1099						Max	5.49	0.1099					
Min	2.39	0.0146						Min	2.39	0.0146					

Perbedaan inisiasi ini dapat dilihat lebih jelas pada gambar 8, di bawah ini yang merupakan perwujudan dari Case-01 dan Case-02, pada Gambar 7 bagian kiri (Case-01) terlihat bahwa

pembagian kelompok jelas berdasarkan besaran produksi sedangkan Case-02 besarnya acak (urutan bebas).



Gambar 7 Perbedaan inisiasi pengelompokan (kiri diurut berdasarkan produksi, kanan urutan bebas)

Benar bahwa perbedaan inisialisasi dapat menghasilkan urutan yang berbeda, hal ini karena kombinasi hasil optimasi tidak unik (tunggal) dapat saja dengan kombinasi yang berbeda

menghasilkan nilai fungsi objektif yang sama, untuk itu selanjutnya pada kasus ini akan menggunakan inisialisasi pengelompokan berdasarkan produksi (Case-01).

Langkah ke tiga analisis adalah penentuan titik pusat masing-masing kelompok dengan mencari rata-rata koordinat X dan Y nya seperti diformulasikan pada persamaan 6 dan 7 (lihat Sub-bab 2.4).

Hasil perhitungan penentuan centroid menghasilkan centroid kelompok 1 (92.1, 30.6), kelompok 2 (73.9,36.9), dan seterusnya hingga kelompok 6 bertitik pusat pada (11.0, 35.1), dilanjutkan dengan penentuan jarak tiap titik anggota kelompok dengan titik pusat kelompoknya seperti disajikan pada Tabel 4 ini jumlahkan untuk masing-masing kelompok dan jumlahkan juga hasilnya untuk menghasilkan nilai Fungsi Objektif seperti dirumuskan pada formula

persamaan 8 (lihat Sub-bab II.5) hasilnya adalah Nilai Fungsi Objektif (FO) 622.68 dan karena FO awalnya tidak ada jadi dianggap nol maka perbedaan atau Threshold = FO (awal) - FO (akhir) = 622.68 - 0 = 622.68 $\geq 10^{-2}$, artinya simulasi belum optimal.

Langkah ke-empat hitung jarak masing-masing titik ketitik pusat semua kelompok, ini dilakukan untuk mengetahui bahwa ada titik anggota suatu kelompok ternyata lebih dekat ke kelompok lain sehingga harus dilakukan pemindahan anggota kelompok pada titik tersebut. Perhitungan jarak ke semua titik pusat ini disajikan pada Tabel 5 Jarak titik pada Centroid Seluruh Kelompok.

Tabel 4 Titik Pusat (Centroid) dan Jarak Titik anggota kelompok terhadap centroid

K	No	Lokasi	X	Y	PENENTUAN TITIK PUSAT KELOMPOK						JARAK TITIK KE CENTROID							
					1	2	3	4	5	6	1	2	3	4	5	6		
1	1	Bali	100.00	17.12	100.0	17.1							15.6					
1	2	Jawa Timur	97.31	24.10	97.3	24.1							8.3					
1	3	Jawa Tengah	92.69	13.56	92.7	13.6							17.1					
1	4	Jawa Barat	91.40	41.43	91.4	41.4							10.8					
1	5	DI Yogyakarta	87.42	55.81	87.4	55.8							25.6					
1	6	DKI Jakarta	83.92	31.66	83.9	31.7							8.3					
2	7	Banten	80.53	42.85			80.5	42.9						8.9				
2	8	Gorontalo	75.75	80.36			75.8	80.4						43.5				
2	9	Sulawesi Barat	74.49	50.77			74.5	50.8						13.9				
2	10	Sulawesi Selatan	72.87	22.72			72.9	22.7						14.2				
2	11	Sumatera barat	70.42	3.93			70.4	3.9						33.1				
2	12	Nusa Tenggara Barat	69.38	20.67			69.4	20.7						16.8				
3	13	Sulawesi Utara	64.85	44.87					64.9	44.9					10.0			
3	14	Sumatera Utara	59.49	33.78					59.5	33.8					12.6			
3	15	Aceh	59.04	13.35					59.0	13.4					32.5			
3	16	Lampung	58.03	50.96					58.0	51.0					6.3			
3	17	Sulawesi Tengah	47.34	79.32					47.3	79.3					34.6			
3	18	Sulawesi Tenggara	40.30	50.89					40.3	50.9					15.5			
4	19	Sumatera Selatan	40.23	56.60							40.2	56.6				10.5		
4	20	Bengkulu	38.62	32.11							38.6	32.1				14.4		
4	21	Jambi	36.97	61.36							37.0	61.4				14.9		
4	22	Maluku Utara	35.57	0.00							35.6	0.0				46.5		
4	23	Papua Barat	35.09	82.41							35.1	82.4				36.0		
5	24	Kalimantan Selatan	31.87	63.65									31.9	63.6			8.1	
5	25	Riau	26.11	21.95									26.1	21.9			47.4	
5	26	Kalimantan Timur	24.90	85.70									24.9	85.7			16.4	
5	27	Maluku	24.27	100.00									24.3	100.0			30.7	
5	28	Papua	23.36	75.48									23.4	75.5			6.7	
6	29	Kepulauan Riau	21.62	58.77											21.6	58.8		26.0
6	30	Nusa Tenggara Timur	15.31	23.91											15.3	23.9		12.0
6	31	Kalimantan Barat	11.47	37.51											11.5	37.5		2.5
6	32	Bangka Belitung	6.52	25.42											6.5	25.4		10.7
6	33	Kalimantan Tengah	0.00	29.86											0.0	29.9		12.2

5.49 0.11 92.1 30.6 73.9 36.9 54.8 45.5 37.3 46.5 26.1 69.4 11.0 35.1 85.7 130.5 111.5 122.3 109.3 63.2
2.39 0.01

FO = 622.68
Perubahan FO = 622.68

Perhatikan Tabel 5 dibawah ini, Bali asal kelompok 1 dan ternyata titik yang mewakili Bali juga mempunyai jarak terdekat dengan Centroid Kelompok 1 artinya Bali tetap anggota Kelompok 1, sekarang lihat DIY asal kelompok 1, hasil perhitungan jarak Euclidean dengan seluruh kelompok menunjukkan bahwa titik yang mewakilinya paling dekat dengan Centroid Kelompok 2 yaitu 23.25 dari pada ke Centroid kelompoknya 25.63, oleh sebab itu DIY harus berpindah ke Kelompok 2, dan berikan tanda kode perubahan (Code Change/CC).

Hal yang sama terjadi pada DIY untuk iterasi pertama ini terjadi pada sembilan Propinsi yang lain yaitu Gorontalo, Aceh, Sulawesi Tengah, Jambi, Maluku Utara, Papua Barat, Riau dan Kepulauan Riau, perpindahan ini bukan hanya perpindahan menurun seperti DIY dari kelompok 1 ke Kelompok 2 tetapi juga perpindahan yang meningkat seperti Aceh dari Kelompok 3 ke Kelompok 2, untuk lebih jelas dapat dilihat pada Tabel 5, yang dengan penandaan *highlight*.

Tabel 5 Jarak Titik pada Centroid Seluruh Kelompok

K	No	Lokasi	X	Y	JARAK TITIK KE SELURUH CENTROID						K BARU	CC
					1	2	3	4	5	6		
1	1	Bali	100.00	17.12	15.62	32.73	53.35	69.24	90.49	90.81	1	0
1	2	Jawa Timur	97.31	24.10	8.32	26.66	47.57	64.05	84.37	87.02	1	0
1	3	Jawa Tengah	92.69	13.56	17.06	29.94	49.54	64.44	86.87	84.49	1	0
1	4	Jawa Barat	91.40	41.43	10.84	18.08	36.79	54.34	71.02	80.67	1	0
1	5	DI Yogyakarta	87.42	55.81	25.63	23.25	34.16	50.98	62.79	79.19	2	1
1	6	DKI Jakarta	83.92	31.66	8.27	11.29	32.22	48.93	69.02	73.02	1	1
2	7	Banten	80.53	42.85	16.86	8.91	25.83	43.38	60.54	69.98	2	1
2	8	Gorontalo	75.75	80.36	52.37	43.52	40.63	51.24	50.86	79.02	3	2
2	9	Sulawesi Barat	74.49	50.77	26.78	13.90	20.34	37.44	51.83	65.42	2	2
2	10	Sulawesi Selatan	72.87	22.72	20.80	14.20	29.07	42.79	66.04	63.11	2	2
2	11	Sumatera barat	70.42	3.93	34.39	33.14	44.42	53.93	79.02	67.11	2	2
2	12	Nusa Tenggara Barat	69.38	20.67	24.82	16.84	28.80	41.19	65.14	60.15	2	2
3	13	Sulawesi Utara	64.85	44.87	30.77	12.08	10.03	27.60	45.84	54.75	3	2
3	14	Sumatera Utara	59.49	33.78	32.79	14.75	12.64	25.58	48.79	48.52	3	2
3	15	Aceh	59.04	13.35	37.31	27.84	32.45	39.64	64.97	52.75	2	3
3	16	Lampung	58.03	50.96	39.70	21.22	6.30	21.21	36.85	49.65	3	3
3	17	Sulawesi Tengah	47.34	79.32	66.17	50.07	34.62	34.32	23.45	57.25	5	4
3	18	Sulawesi Tenggara	40.30	50.89	55.65	36.41	15.50	5.32	23.29	33.30	4	5
4	19	Sumatera Selatan	40.23	56.60	58.03	39.02	18.33	10.52	19.03	36.31	4	5
4	20	Bengkulu	38.62	32.11	53.52	35.61	21.05	14.45	39.29	27.80	4	5
4	21	Jambi	36.97	61.36	63.14	44.31	23.87	14.86	13.49	36.95	5	6
4	22	Maluku Utara	35.57	0.00	64.31	53.20	49.44	46.53	70.00	42.85	6	7
4	23	Papua Barat	35.09	82.41	77.04	59.83	41.84	35.98	15.85	53.11	5	8
5	24	Kalimantan Selatan	31.87	63.65	68.72	49.84	29.26	17.99	8.11	35.38	5	8
5	25	Riau	26.11	21.95	66.57	50.07	37.17	26.98	47.41	20.04	6	9
5	26	Kalimantan Timur	24.90	85.70	86.91	69.17	50.10	41.11	16.39	52.49	5	9
5	27	Maluku	24.27	100.00	97.05	80.30	62.46	55.07	30.70	66.25	5	9
5	28	Papua	23.36	75.48	82.11	63.60	43.45	32.16	6.71	42.25	5	9
6	29	Kepulauan Riau	21.62	58.77	75.91	56.68	35.76	19.90	11.50	25.96	5	10
6	30	Nusa Tenggara Timur	15.31	23.91	77.11	60.02	45.06	31.52	46.71	11.99	6	10
6	31	Kalimantan Barat	11.47	37.51	80.95	62.45	44.11	27.35	35.05	2.46	6	10
6	32	Bangka Belitung	6.52	25.42	85.76	68.35	52.34	37.30	48.10	10.65	6	10
6	33	Kalimantan Tengah	0.00	29.86	92.13	74.24	57.04	40.84	47.35	12.17	6	10

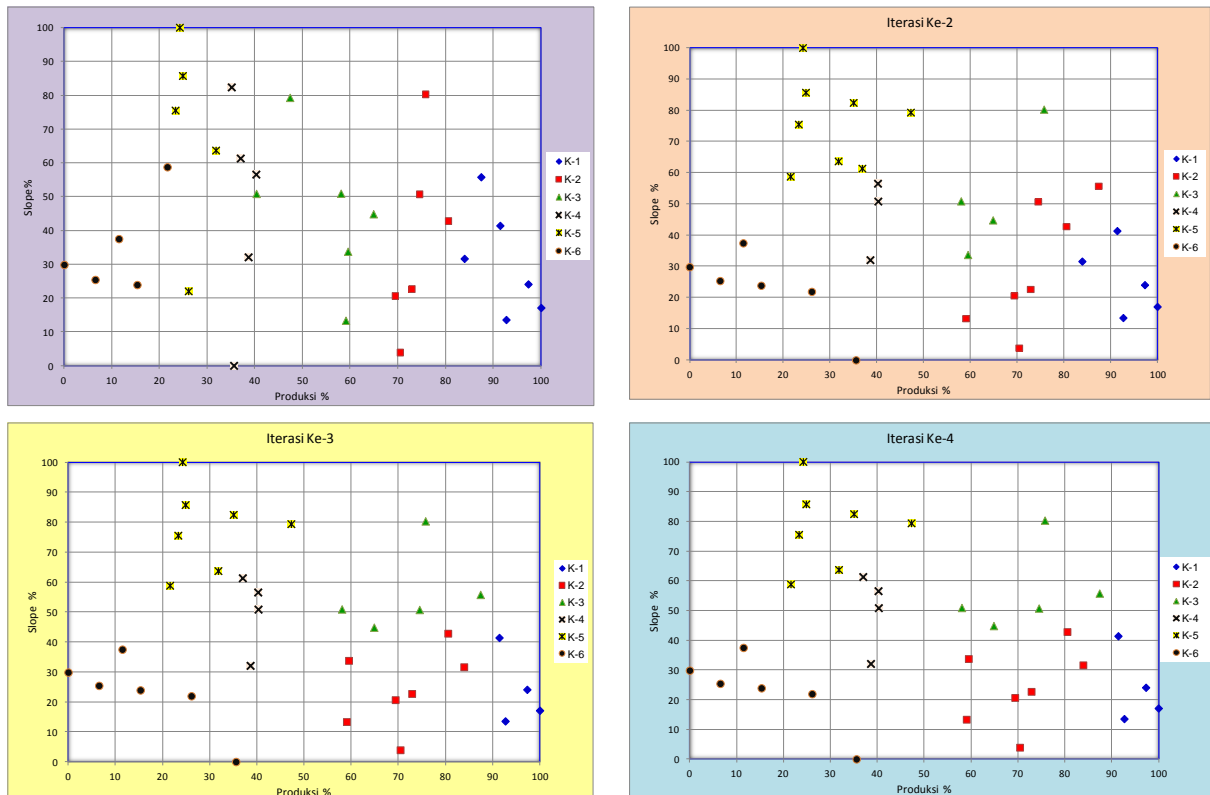
Ulangi lagi simulasi dengan langkah awal perubahan anggota kelompok atau kalau pada Tabel 5 ini anggota K pada kolom pertama diganti dengan anggota K baru (kolom 12), lakukan penentuan centroid, hitung jarak, hitung

FO, hitung apakah perubahan FO kurang dari batas ambang, hitung jarak pada semua centroid kelompok, lihat perubahan anggota kelompok, ganti anggota kelompok, berikan kode perubahan,

seterusnya hingga diperoleh nilai perubahan FO kurang dari ambang batas sebesar 10^{-2} .

Simulasi untuk kasus ini telah dilaksanakan dengan hasil optimasi diperoleh setelah

melakukan empat iterasi, perhatikan perpindahan titik pada kelompoknya seperti lebih jelas digambarkan hasil seluruh simulasi pada Gambar 8.



Gambar 8 Perubahan Pengelompokan Optimasi K-Means, Variabel bebas Produksi dan Variabel Tidak Bebas Kemiringan (Slope)

Perubahan Pengelompokan Optimasi K-Means, Variabel bebas Produksi dan Variabel Tidak Bebas Kemiringan (Slope), dan seluruh perubahan perhitungan simulasi disajikan pada lampiran 2. Simulasi Optimasi Kasus Rata-rata dan Slope (1993-2012), dimana nilai FO berubah mereduksi

mulai dari 622.68, menjadi 477.20, berkurang lagi 454.30, dan akhirnya 454.30, dengan begitu nilai perubahan FO terakhir adalah 0 yang berarti kurang dari satu perseratus atau 10^{-2} , dan diperoleh pengelompokan seperti disajikan pada Tabel 6.

Tabel 6 Hasil Pengelompokan Simulasi Optimasi K-Means Rata-rata dan Slope 1993-2012

KELOMPOK					
1	2	3	4	5	6
Bali	DKI Jakarta	DI Yogyakarta	Sulawesi Tenggara	Sulawesi Tengah	Maluku Utara
Jawa Timur	Banten	Gorontalo	Sumatera Selatan	Papua Barat	Riau
Jawa Tengah	Sulawesi Selatan	Sulawesi Barat	Bengkulu	Kalimantan Selatan	Nusa Tenggara Timur
Jawa Barat	Sumatera barat	Sulawesi Utara	Jambi	Kalimantan Timur	Kalimantan Barat
	Nusa Tenggara Barat	Lampung		Maluku	Bangka Belitung
	Sumatera Utara			Papua	Kalimantan Tengah
	Aceh			Kepulauan Riau	
4	7	5	4	7	6

4.4. Kasus Rata-rata dan Slope, skenario data 2001-2012

Simulasi untuk Kasus yang sama tetapi dengan skenario data yang digunakan adalah rekaman data mulai dari 2001 hingga data 2012 selanjutnya disebut simulasi Kasus-02, dilakukan untuk merepresentasikan simulasi pengelompokan kondisi saat ini, dimana rata-rata produksi dan kemiringan relatif lebih besar dari kondisi 20 tahun data, adapun perbedaannya adalah rata-rata produksi 1993-2012 sebesar 4.00 Ton/Ha sedangkan rata-rata produksi 2001-2012 meningkat menjadi 4.168 Ton/Ha, demikian juga nilai Slopenya meningkat dari 0.0556 menjadi 0.0703.

4.5. Aplikasi Pada Kasus Rata-rata dan Perkiraan Produksi tahun 2013

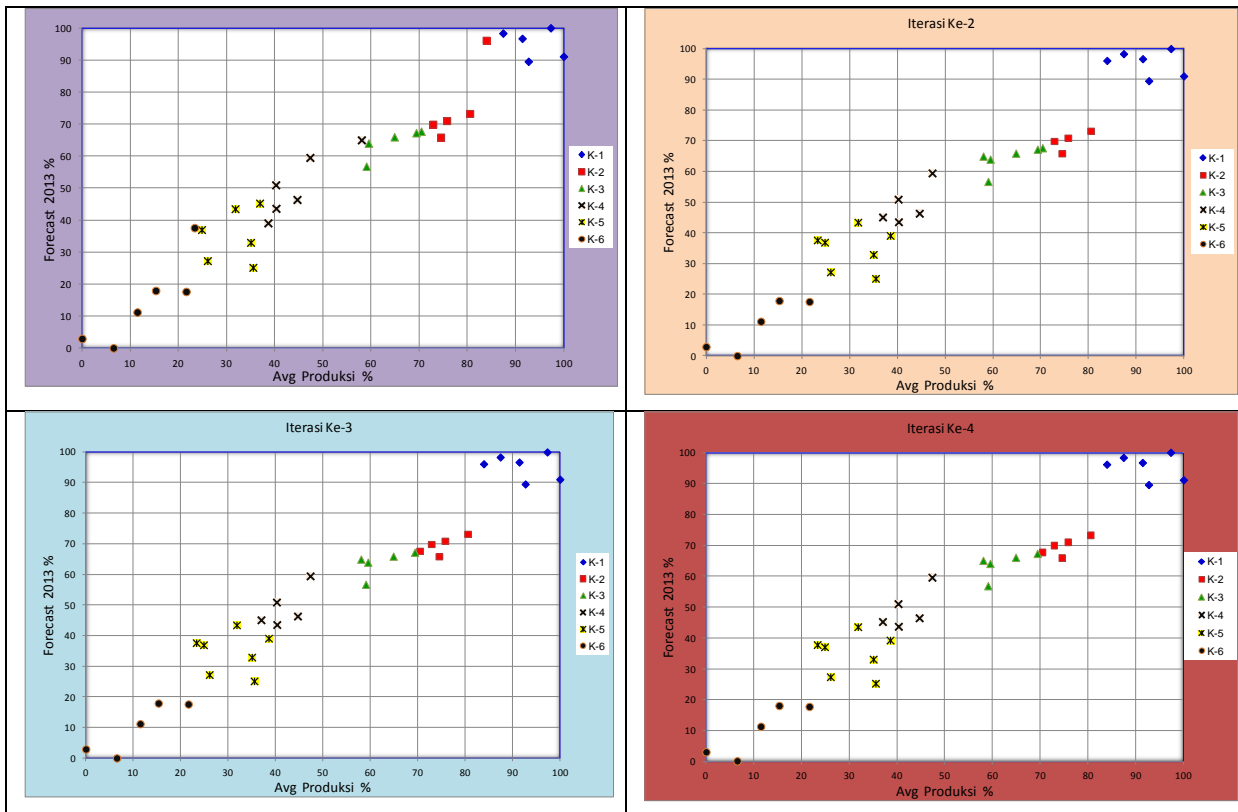
Simulasi optimasi ketiga (Kasus-03) dilakukan untuk pengelompokan 33 propinsi berdasarkan produksi padi menggunakan set data dua dimensi yaitu :

X: Rata-rata besaran produksi padi dari tahun 1993 sampai dengan 2012 dalam unit

(Ton/Ha); (kolom 3, pada Tabel statistik dasar produksi padi nasional)

Y: Perkiraan produksi padi dari tahun 2013, dalam unit (Ton/Ha) (kolom 12, pada Tabel statistik dasar produksi padi nasional)

Menggunakan tahapan persis sama seperti dilakukan pada dua simulasi terdahulu, yang membedakannya adalah variabel tidak bebas Y berupa produksi padi 2013 sehingga besarnya relatif seimbang dengan rata-rata yang menjadikan gambar mendekati garis lurus. Pada simulasi ini initial pengelompokan terbagi dalam 6 Kelompok dengan banyaknya anggota masing-masing Kelompok adalah 6, 6, 6, 5, 5, dan 5, dengan urutan berdasarkan besaran produksi, hasil simulasi optimasi sampai empat iterasi dengan FO keempat bernilai 204.03 dan perubahan $FO = 0 < 10^{-2}$, berikut ini gambar pengelompokan Optimasi K-Means dengan variabel tidak bebas berupa forecasting produksi tahun 2013 (Gambar 9), sementara hasil simulasi tiga kasus pengelompokan produksi padi nasional disajikan pada Tabel 7.



Gambar 9 Perubahan Pengelompokan Optimasi K-Means, Variabel bebas Produksi dan Variabel Tidak Bebas Perkiraan Produksi 2013

Tabel 7 Hasil Simulasi tiga kasus Pengelompokan Produksi Padi Nasional

SIMULASI KASUS-01		SIMULASI KASUS-02		SIMULASI KASUS-03	
K	Propinsi	K	Propinsi	K	Propinsi
1	Bali	1	Bali	1	Bali
1	Jawa Timur	1	Jawa Timur	1	Jawa Timur
1	Jawa Tengah	1	Jawa Barat	1	Jawa Tengah
1	Jawa Barat	1	DI Yogyakarta	1	Jawa Barat
2	DKI Jakarta	1	Jawa Tengah	1	DI Yogyakarta
2	Banten	1	DKI Jakarta	1	DKI Jakarta
2	Sulawesi Selatan	2	Banten	2	Banten
2	Sumatera barat	2	Sulawesi Selatan	2	Gorontalo
2	Nusa Tenggara Barat	2	Gorontalo	2	Sulawesi Barat
2	Sumatera Utara	2	Sulawesi Barat	2	Sulawesi Selatan
2	Aceh	2	Nusa Tenggara Barat	2	Sumatera barat
3	DI Yogyakarta	2	Sulawesi Utara	2	Nusa Tenggara Barat
3	Gorontalo	2	Sumatera barat	2	Sulawesi Utara
3	Sulawesi Barat	3	Lampung	3	Sumatera Utara
3	Sulawesi Utara	3	Sumatera Utara	3	Aceh
3	Lampung	3	Aceh	3	Lampung
4	Sulawesi Tenggara	3	Sulawesi Tenggara	3	Sulawesi Tengah
4	Sumatera Selatan	4	Sulawesi Tengah	4	Maluku
4	Bengkulu	4	Sumatera Selatan	4	Sulawesi Tenggara
4	Jambi	4	Kalimantan Selatan	4	Sumatera Selatan
5	Sulawesi Tengah	4	Maluku	4	Bengkulu
5	Papua Barat	4	Kalimantan Timur	4	Jambi
5	Kalimantan Selatan	4	Papua Barat	4	Kalimantan Selatan
5	Kalimantan Timur	4	Papua	5	Maluku Utara
5	Maluku	5	Jambi	5	Papua Barat
5	Papua	5	Bengkulu	5	Riau
5	Kepulauan Riau	5	Riau	5	Kalimantan Timur
6	Maluku Utara	5	Kepulauan Riau	5	Papua
6	Riau	6	Maluku Utara	6	Kepulauan Riau
6	Nusa Tenggara Timur	6	Nusa Tenggara Timur	6	Nusa Tenggara Timur
6	Kalimantan Barat	6	Kalimantan Barat	6	Kalimantan Barat
6	Bangka Belitung	6	Bangka Belitung	6	Bangka Belitung
6	Kalimantan Tengah	6	Kalimantan Tengah	6	Kalimantan Tengah

4.6. Usulan Pengelompokan Propinsi berdasarkan Produksi Padi

Telah dapat dibuktikan bahwa perbedaan hasil optimasi dikarenakan perbedaan pendekatan yang diambil dalam hal ini variabel tidak bebas yang berbeda walaupun dengan variabel bebas yang sama yaitu produksi padi. Simulasi telah mencoba mengakomodir tiga kondisi yaitu (1) dimensi waktu seluruhnya (Kasus-01, 1993-2012) yang diharapkan dapat mengcover kejadian masa lampau hingga saat ini, (2) dimensi kondisi saat ini (Kasus-02 2001-2012) mewakili realitas kondisi saat ini dimana peningkatan produksi sepuluh tahun terakhir berbeda signifikan dibanding kondisi masa lampau 20-10 tahun yang lalu, dan (3) Simulasi optimasi kondisi masa mendatang (Kasus-03, Forecast 2013) yang menyajikan perkiraan produksi satu tahun kedepan berdasarkan peningkatan yang terjadi.

Hasil simulasi ketiga kasus ini (lihat Tabel 7) menunjukkan terdapat propinsi-propinsi yang memang benar-benar konsisten dikelompoknya kecuali untuk Kelompok 5 tidak ditemukan satu pun propinsi yang konsisten mulai dari pengamatan masa lampau hingga prediksi kedepan ada pada Kelompok ini. Tetapi juga ditemukan propinsi yang berubah Kelompok pada tiap Kasus Simulasi dan juga terdapat beberapa yang berada pada batas akhir kelompok.

Satu catatan penting diperoleh juga nampaknya Kasus-02 dan Kasus-03 menghasilkan pengelompokan yang hampir sama ini berarti ada kesetaraan kecenderungan antara kejadian produksi padi sepuluh terakhir (Kasus-02) dan forecast 2013 (Kasus-03).

Apa yang diperoleh dari simulasi kasus-kasus ini adalah walaupun potensi pengembangan masih sedemikian luas (Gambar 4) perlu dipertimbangkan juga produksi yang akan

diperoleh karena muara pengembangan adalah produksi padi jadi mengapa tidak dibarengi dengan berfikir riwayat produksi.

Merujuk hasil ketiga simulasi yang sudah mengakomodir dimensi masa lampau, saat ini, dan mendatang maka pengelompokan propinsi berdasarkan produksi padi adalah seperti dituliskan pada Tabel 8. Resume Hasil Optimasi dan Usulan Pengelompokan dengan catatan khusus karena propinsi yang berada pada kelompok 1 nampaknya sudah tidak dapat dikembangkan luas irigasinya maka kelompok 2 kiranya dapat menjadi prioritas.

Tabel 8 Resume Hasil Optimasi dan usulan pengelompokan

No	Propinsi	KS-01	KS-02	KS-03	K
1	Bali	1	1	1	1.00
2	Jawa Timur	1	1	1	1.00
3	Jawa Tengah	1	1	1	1.00
4	Jawa Barat	1	1	1	1.00
5	DKI Jakarta	2	1	1	1.33
6	DI Yogyakarta	3	1	1	1.67
7	Banten	2	2	2	2.00
8	Sulawesi Selatan	2	2	2	2.00
9	Sumatera barat	2	2	2	2.00
10	Nusa Tenggara Barat	2	2	2	2.00
11	Gorontalo	3	2	2	2.33
12	Sulawesi Barat	3	2	2	2.33
13	Sulawesi Utara	3	2	2	2.33
14	Aceh	2	3	3	2.67
15	Sumatera Utara	3	3	3	3.00
16	Lampung	3	3	3	3.00
17	Sulawesi Tengah	4	4	3	3.67
18	Sulawesi Tenggara	4	3	4	3.67
19	Sumatera Selatan	4	4	4	4.00
20	Maluku	5	4	4	4.33
21	Bengkulu	4	5	4	4.33
22	Jambi	4	5	4	4.33
23	Kalimantan Selatan	5	4	4	4.33
24	Papua Barat	5	4	5	4.67
25	Kalimantan Timur	5	4	5	4.67
26	Papua	5	4	5	4.67
27	Riau	6	5	5	5.33
28	Kepulauan Riau	5	5	6	5.33
29	Maluku Utara	6	6	5	5.67
30	Nusa Tenggara Timur	6	6	6	6.00
31	Kalimantan Barat	6	6	6	6.00
32	Bangka Belitung	6	6	6	6.00
33	Kalimantan Tengah	6	6	6	6.00

Sulawesi Selatan (produksi rata-rata: 4.65, forecast 2013: 5.06 Ton/Ha), Sumatera Barat (4.57, 4.99), dan NTB (4.54, 4.98) adalah propinsi yang konsisten selalu berada pada Kelompok 2,

sementara Gorontalo (4.74, 5.10), Sulawesi Barat (4.70, 4.93), dan Sulawesi Utara (4.40, 4.93) adalah propinsi yang cukup progresif perkembangannya pada tahun-tahun terakhir, sebaliknya Aceh (4.22, 4.64) yang sebelumnya (10-20 tahun) kebelakang cukup bagus, turun menjadi Kelompok 3 pada Simulasi Optimasi Kasus-02 kondisi existing dan Kasus-03 Forecast 2013.

Sumatera Utara, dan Lampung dua propinsi penghasil beras perlu perhatian khusus karena selamanya (konsisten) selalu berada pada Kelompok 3.

Kelompok 4 dan Kelompok 5 anggotanya sering bertukar tempat catatan ada pada dua propinsi di Pulau Sulawesi, yaitu Sulawesi Tengah (3.86, 4.72) yang kemiringan kenaikannya cukup baik, sedangkan Sulawesi Tenggara (3.64, 4.21) Produksi 10 tahun terakhir cukup baik tetapi kenaikannya relatif datar.

Pengembangan di Papua dari segi lahan dan air sangat potensial tetapi berdasarkan produksi Nasional perlu mendapat perhatian khusus dan Iptek yang lebih maju, dua Propinsi disini ada di Kelompok 5 (3.48, 3.86) dan (3.11, 4.01).

Terakhir empat Propinsi yang konsisten berada dibawah yaitu NTT (2.86, 3.38), Kalimantan Barat (2.75, 3.16), Bangka Belitung (2.59, 2.80), dan Kalimantan Tengah (2.39, 2.89), disarankan alternatif pengembangan lain selain irigasi dipertimbangkan untuk propinsi-propinsi ini.

V. KESIMPULAN DAN SARAN

5.1. Kesimpulan

1. Pengembangan lahan irigasi tahun 2020 berdasarkan Studi FIDP 1993 yang dijadikan acuan Bappenas dan PU menempatkan Pulau Sumatera, Kalimantan, dan Maluku-Papua sangat potensial dari segi ketersediaan air dan kecocokan lahan untuk padi, akan tetapi berdasarkan rekaman data produksi padi selama 20 tahun (1993-2012) wilayah-wilayah tersebut bukan wilayah produksi tinggi (Ton/ha), artinya untuk menghasilkan satu Ton padi dibutuhkan lahan lebih luas pada wilayah ini.
2. Hasil Optimasi Data mining pendekatan K-Means pengelompokan data yang mengakomodir tiga dimensi waktu yaitu pengaruh masa lampau (Kasus-01), waktu saat ini atau kejadian 10 tahun terakhir

(Kasus-02), dan Kasus-03 (forecast 2013), menunjukkan bahwa produksi padi nasional di Indonesia tertinggi (kelompok 1) adalah Pulau Jawa dan Bali, artinya diperlukan upaya lebih untuk mempertahankan luas area dengan mencegah alih fungsi lahan, dan pengembangan pada wilayah ini memanfaatkan potensi seoptimal mungkin karena keterbatasan area pengembangan pada wilayah ini (hanya 62000 Ha).

3. Wilayah Sumatera umumnya menempati kelompok 3 yaitu Aceh yang secara rata-rata produksi relatif bagus tetapi peningkatan produksinya (*slope*) relatif rendah, sementara dua propinsi Sumatera penghasil padi yaitu Lampung dan Sumatera Utara konsisten ada dikelompok 3, Sumatera Selatan berada dibawahnya yaitu kelompok 4.
4. Wilayah Sumatera, Papua dan Kalimantan menjadi kelompok terakhir untuk pengembangan dari sisi produktifitas, apabila tidak ada perlakuan khusus seperti penerapan teknologi maju.
5. Usulan Pengembangan Irigasi berdasarkan Studi FIDEP tahun 1993 perlu direview ulang menggunakan data dua puluh tahun terakhir, berintegrasi dengan rencana dan pola pada studi pengembangan tiap wilayah sungai, kesesuaian dengan paradigma baru, dan kebijakan pemerintah mutakhir berdasarkan legal aspek terbaru turunan dari UU No. 7/2004 Tentang SDA.

5.2. Saran

1. Jawa dan Bali masuk dalam kategori kelompok 1 artinya mempunyai produktifitas tertinggi, namun potensi pengembangan pada wilayah ini sangat kecil, memanfaatkan potensi area seoptimal mungkin dan mempertahankan area irigasi dari alih fungsi lahan lebih diprioritaskan pada wilayah ini. Pengembangan daerah irigasi lebih rasional adalah pada kelompok 2 yaitu Wilayah Sulawesi, NTB, dan Sumatera Barat.
2. Terdapat empat Propinsi yang konsisten berada dikelompok paling bawah, artinya produksi rata-rata rendah dan peningkatan juga rendah yaitu NTT, Kalimantan Barat, Bangka Belitung, dan Kalimantan Tengah. Pada wilayah ini, alternatif pengembangan lain selain irigasi disarankan untuk

dipertimbangkan, kecuali perlakuan sangat khusus seperti penerapan teknologi maju.

3. Model Optimasi Data mining dengan pendekatan K-Means mudah dioperasikan dan dapat memartiskan suatu kumpulan data dalam kelompok-kelompok kecil sehingga data lebih memberikan informasi dan masing-masing kelompok menonjolkan kepusatannya. Pendekatan lain seperti Fuzzy logic, Neurol Network, perlu di uji cobakan untuk aplikasi permasalahan yang lain.
4. Penulis menyarankan kepada Ditjen Sumber Daya Air agar Studi FIDEP tahun 1993 data direview ulang untuk disesuaikan dengan fakta pengembangan selama 20 tahun terakhir dan disesuaikan dengan paradigma baru serta kebijakan pemerintah yang mutakhir sesuai UU No 7/2004 : SDA dan turunannya PP No 20/2006 : Tentang Air Irigasi, Permen PU No 30/PRT/M/2007 Tentang Pedoman Pengembangan dan Pengelolaan sistem irigasi partisipatif, Permen PU No 33/PRT/M/2007 Tentang Pedoman Perkumpulan Petani Pemakai Air , dan Permen PU No 32/PRT/M/2007 Tentang Operasi dan Pemeliharaan Jaringan Irigasi.

DAFTAR PUSTAKA

- BPS. 2013. *Pertanian dan Pertambangan, Tanaman Pangan Padi, seluruh propinsi 1993 - 2012*. (www.bps.go.id) diakses tanggal 27 Februari 2013.
- Departemen Pekerjaan Umum. 2007. Keputusan Menteri Pekerjaan Umum No. 390/KPTS/M/2007. *Tentang Penetapan Status Daerah Irigasi yang Pengelolaannya menjadi Tanggung Jawab Pemerintah, Pemerintah Provinsi dan Pemerintah Kabupaten/Kota*. Departemen Pekerjaan Umum. Jakarta
- Direktorat Irigasi dan Rawa. 2010. *Buku Pintar Irigasi*. Direktorat Irigasi dan Rawa, Direktorat Jenderal Sumber Daya Air, Kementerian PU.
- Hamdy, A. Taha. 2010. *Suatu Pengantar Riset Operasi*. Binarupa Pustaka. Jakarta.
- Irawan, Bambang. 2004. *Konversi Lahan Sawah di Jawa dan Dampaknya terhadap Produksi*

- Nasional, Ekonomi Padi dan Beras Indonesia*. Balittan, Departemen Pertanian.
- Kasryno, Faisal dan Pasandaran, Efendi. 2004. *Reposisi Padi dan Beras dalam Perekonomian Nasional, Ekonomi Padi dan Beras Indonesia*. Balittan, Departemen Pertanian.
- Makridakis, Spyros, Steven C. Wheelwright, dan Victor E. McGee. 1999. *Metode dan Aplikasi Peramalan*, Erlangga. Jakarta.
- Nippon Koei. 1993. *The Study fro formulation of irrigation development program un The Republic of Indonesia*. JICA-Directorate General of Water Resources Development Ministry of Public Works-Bureau of Water Resources and Iriigation. Bappenas.
- Prasetyo, Eko. 2012. *Data Mining – Konsep Aplikasi Menggunakan MATLAB*. ANDI Offset. Jogyakarta.
- Republik Indonesia. 2006. Peraturan Pemerintah. Nomor 20. 2006. *Tentang Irigasi*. Sekretariat Negara. Jakarta.
- Robert J. Vanderbei. 2008. “Linear Programming: Foundations and Extensions, 3rd ed.,2008”. *International Series in Operations Research & Management Science*, Vol. 114, Springer Verlag.
- Sastrohardjono, Sukrasno, dkk. 2011. *Pokok-pokok Pikiran Modernisasi irigasi di Indonesia: sebuah kajian akademik*. Direktorat Irigasi dan Rawa, Direktorat Jenderal Sumber Daya Air.
- Sujana. 2010. “Aplikasi Mining Data dengan Metode Decition Tree”. *Seminar Nasional Aplikasi teknologi Informasi 2010* (SNATI 2010).
- Tan, P, et al 2006. *Introduction to data Mining*. Pearson Education. Boston.