

## PEMODELAN *STATISTICAL DOWNSCALING* REGRESI KUANTIL LASSO DAN ANALISIS KOMPONEN UTAMA UNTUK PENDUGAAN CURAH HUJAN EKSTRIM

### *STATISTICAL DOWNSCALING MODELING OF LASSO QUANTIL REGRESSION AND ANALYSIS OF MAIN COMPONENTS FOR EXTREME RAINFALL ESTIMATION*

Dewi Santri<sup>1§</sup>, Yusrianti Hanike<sup>2</sup>

<sup>1</sup>IAIN Kendari, Indonesia [Email: [dewisantri@iainkendari.ac.id](mailto:dewisantri@iainkendari.ac.id)]

<sup>2</sup>IAIN Ambon, Indonesia [Email: [yusriantihanike@iainambon.ac.id](mailto:yusriantihanike@iainambon.ac.id)]

<sup>§</sup>Corresponding Author

Received Mei 2020; Accepted Juni 2020; Published Juni 2020;

#### Abstrak

Curah hujan ekstrem yang sering terjadi di Indonesia menimbulkan berbagai dampak negatif bagi masyarakat. Terdapat banyak pemodelan curah hujan yang telah dilakukan untuk meminimumkan dampak yang terjadi. *Global circulation model* (GCM) diyakini menjadi metode terbaik untuk meramalkan data curah hujan ekstrem. Kelemahan dari data GCM adalah masih bersifat global sehingga akan sulit untuk menjelaskan keragaman dalam skala lokal yang lebih rinci. *Statistical Downscaling* (SD) hadir untuk menangani permasalahan tersebut. SD menghubungkan antara data luaran GCM dan curah hujan untuk menduga perubahan pada skala lokal dengan menggunakan metode regresi. Untuk mengakap nilai ekstrem dari curah hujan maka digunakan metode regresi kuantil. Data luaran GCM yang memiliki multikolinearitas tidak dapat langsung diterapkan dalam model SD. Metode-metode yang dapat digunakan untuk mengatasi masalah multikolinearitas dalam SD antara lain metode analisis komponen utama (AKU) dan metode *shrinkage* seperti *Least Absolute Shrinkage and Selection Operator* (LASSO). Metode AKU paling sering digunakan dalam mereduksi dimensi data luaran GCM dan menangani masalah multikolinearitas. Metode *shrinkage* selain dapat menghilangkan multikolinearitas juga dapat meminimumkan ragam penduga parameter dari model regresi. Tujuan penelitian ini adalah menentukan model curah hujan ekstrem di Kabupaten Indramayu dengan pendekatan SD menggunakan metode regresi kuantil dengan LASSO dan AKU serta memilih model SD terbaik dari kedua metode yang digunakan tersebut. Hasil penelitian menunjukkan bahwa dugaan curah hujan ekstrem di kabupaten Indramayu dengan model SD menggunakan regresi kuantil dengan LASSO menghasilkan prediksi yang lebih konsisten terhadap berbagai selang waktu dugaan dibandingkan model yang menggunakan metode AKU.

**Kata Kunci:** GCM, SD, Regresi Kuantil, AKU, LASSO

#### Abstract

*Extreme rainfall that frequently occurs in Indonesia has negative impact to society. there are several methods that required to minimize the damage that may occur. So far, Global circulation models (GCM) are the best method to forecast global climate changes include extreme rainfall. GCM data has global scale and unable to provide reliable information at local scale. Statistical Downscaling (SD) has been developed in an attempt to bridge this scale gap. SD uses regression models to represent the link between GCM data and local rainfall. Quantile regression is used to catch the extreme rainfall. GCM data which*

*has multicollinearity can not be directly applied in SD model. The methods that can be used to overcome multicollinearity are principal component analysis (PCA) and shrinkage methods such as Least Absolute Shrinkage and Selection Operator (LASSO) and ridge. PCA is the most commonly used in SD modeling. PCA can reduce the dimension of GCM data and multicollinearity. Shrinkage method can eliminate multicollinearity and minimize variance. The objectives of this study are modeling SD using quantile regression with LASSO and PCA to predict extreme rainfall in Indramayu and to choose the best SD model of both methods. The result shows that the prediction of extreme rainfall in Indramayu with SD models using quantile regression with LASSO is more consistent at any time prediction compared to models using PCA.*

**Keywords:** *GCM, SD, Quantile Regression, PCA, LASSO*

---

## 1. Pendahuluan

Curah hujan ekstrem yang sering terjadi di Indonesia memiliki dampak negatif bagi sektor pertanian. Salah satu contohnya pada tahun 2010 produksi padi mengalami penurunan akibat iklim ekstrem yaitu dari 1.58 juta ton di tahun 2009 menjadi sebesar 1.55 juta ton di tahun 2010 [2]. Hal ini menjadikan studi tentang perubahan iklim sangat diperlukan untuk meminimumkan kerugian yang mungkin terjadi.

Curah hujan merupakan kejadian kompleks yang melibatkan topografi dan interaksi antara laut, darat dan atmosfer sehingga mempersulit prediksi curah hujan itu sendiri. Oleh karena itu diperlukan model peramalan curah hujan yang akurat pada skala lokal dengan mempertimbangkan informasi tentang sirkulasi atmosfer global yang dapat diperoleh dari data luaran GCM (*global circulation model*) [6].

GCM tersusun atas rangkaian model-model numerik yang merepresentasikan sejumlah komponen subsistem dari iklim bumi. GCM memiliki kemampuan untuk melakukan simulasi iklim secara skala besar. Model ini diyakini sebagai model penting dalam upaya memahami iklim di masa lampau, sekarang dan masa yang akan datang [15].

Data luaran GCM masih berskala global sehingga akan sulit untuk memperoleh informasi yang berskala lokal (kabupaten atau kota). Agar dapat menangani masalah tersebut diperlukan suatu metode untuk mentransformasi hasil simulasi GCM pada skala global ke skala lokal. Pendekatan *statistical downscaling* (SD) diyakini dapat menangani permasalahan rendahnya akurasi prediksi curah hujan. Metode ini menghubungkan antara data luaran GCM dan curah hujan untuk menduga perubahan pada skala lokal dengan menggunakan model regresi. Informasi dari skala global dalam data luaran GCM akan diproyeksikan terhadap informasi skala lokal stasiun cuaca.

Metode regresi yang dapat digunakan untuk menganalisis data yang mengandung nilai ekstrem adalah regresi kuantil. Regresi kuantil tidak membutuhkan asumsi parametrik dan bermanfaat untuk menganalisis bagian tertentu dari suatu sebaran bersyarat [3]. Kelebihan lain dari regresi kuantil adalah efisien jika sisaan tidak menyebar normal dan kekar terhadap pencilan. Hal tersebut membuat regresi kuantil banyak digunakan untuk memodelkan data yang mengandung nilai ekstrem.

Pada data luaran GCM terdapat multikolinearitas atau antar peubah GCM saling berkorelasi tinggi. Hal ini akan menyebabkan nilai dugaan parameter model regresi menjadi tidak tepat. Oleh karena itu langkah pertama yang harus dilakukan adalah mengatasi masalah multikolinearitas dalam data GCM tersebut.

Metode-metode yang dapat digunakan untuk mengatasi masalah multikolinearitas dalam SD antara lain metode analisis komponen utama (AKU), metode *shrinkage* ridge dan *Least Absolute Shrinkage and Selection Operator* (LASSO). Metode AKU merupakan metode yang paling sering digunakan. Metode AKU dapat mereduksi dimensi dari GCM dan menangani masalah multikolinearitas. Metode *shrinkage* selain dapat menghilangkan multikolinearitas juga dapat meminimumkan ragam penduga parameter dari model regresi. Kelebihan LASSO dibandingkan metode ridge adalah LASSO dapat menyusutkan koefisien penduga tepat nol sehingga dapat melakukan seleksi peubah sehingga model lebih sederhana dan saling bebas [7].

Penelitian mengenai model SD dengan regresi kuantil sebelumnya telah banyak dilakukan, antara lain [5] menggunakan regresi kuantil untuk mengeksplorasi curah hujan di Kabupaten Indramayu pada data yang mengandung pencilan. Penelitian tersebut menghasilkan kesimpulan bahwa regresi kuantil dapat digunakan untuk mendeteksi kondisi-kondisi ekstrim, baik ekstrim kering (kuantil ke-5) maupun ekstrim basah (kuantil ke-95). [10] mengkaji pemodelan curah hujan ekstrim dengan

menggunakan regresi kuantil dengan reduksi peubah GCM menggunakan analisis komponen utama, namun model yang dihasilkan belum dapat mengakomodasi dengan baik pengaruh nonlinear. Keseluruhan dari penelitian tersebut menggunakan metode AKU dalam pemodelan SD.

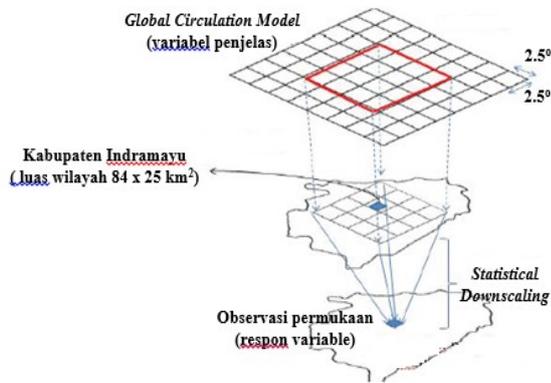
Pemodelan SD pada penelitian ini menggunakan regresi kuantil dengan LASSO. Penambahan penalti LASSO dalam model regresi kuantil membuat analisis SD menjadi lebih sederhana karena pemodelan SD dan penanganan multikolinearitas data luaran GCM dapat dilakukan dalam satu kali tahapan. Selain itu LASSO juga dapat secara simultan melakukan reduksi dimensi data luaran GCM dan mengontrol ragam penduga koefisien pada regresi kuantil.

## 2. Landasan Teori

### 2.1 Global Circulation Model (GCM) dan Statistical Downscaling (SD)

*Global circulation model* (GCM) merupakan alat terpenting dalam memahami sistem iklim karena mampu memberikan informasi tentang pergeseran iklim dari masa lampau sekarang dan di masa yang akan datang. GCM berskala besar (global) atau memiliki resolusi yang rendah sehingga belum memperhitungkan fenomena pada skala kecil (lokal). GCM membuat simulasi peubah-peubah iklim global pada setiap grid (berukuran  $\pm 2,5^\circ$  atau  $\pm 300 \text{ km}^2$ ) setiap lapisan (*layer*) atmosfer, yang selanjutnya digunakan untuk memprediksi pola-pola iklim dalam jangka waktu tahunan [13].

Kelemahan model GCM dalam melakukan pendugaan adalah tingkat akurasi prediksinya yang rendah. Kelemahan ini disebabkan oleh curah hujan yang ditentukan oleh suatu proses fisika yang sangat rumit, sensitif, dan nonlinier yang tidak dapat dimodelkan oleh GCM [11].



Gambar 1. Ilustrasi proses *Statistical Downscaling* [12]

Model SD adalah suatu fungsi transfer yang menggambarkan hubungan fungsional sirkulasi atmosfer global (hasil GCM) dengan unsur-unsur iklim lokal. Pemilihan peubah-peubah prediktor dan penentuan domain (lokasi dan jumlah grid) merupakan faktor kritis yang akan mempengaruhi kestabilan peramalan [14]. Model ini juga memerlukan data deret waktu yang homogen dalam berbagai perubahan iklim. Model SD memberikan hasil yang baik dengan syarat berikut: (1) Hubungan erat antara respon dengan prediktor yang menjelaskan keragaman iklim lokal dengan baik, (2) Peubah prediktor disimulasi baik oleh GCM, dan (3) Hubungan antara respon dengan prediktor tidak berubah dengan perubahan waktu dan tetap sama meskipun ada perubahan iklim. Bentuk umum model SD, yaitu :

$$y_t = f(X_{t \times g})$$

dengan,

- $y_t$  : peubah iklim lokal (curah hujan)
- $X_{t \times g}$  : peubah luaran GCM (presipitasi)
- $t$  : banyaknya waktu (bulanan)
- $g$  : banyaknya *grid* domain GCM

## 2.2 Regresi Kuantil dengan Lasso

Regresi kuantil yang dikemukakan oleh Koenker dan Bassett pada tahun 1978. Regresi kuantil merupakan suatu pendekatan analisis regresi untuk menduga fungsi regresi pada kuantil tertentu yang berguna jika distribusi data tidak homogen (heterogenous), yang bila ditinjau dari segi kurva, kurva tidak berbentuk standar atau tidak simetris, dan terdapat ekor pada sebaran (*truncated distribution*). Metode ini merupakan suatu metode regresi dengan pendekatan memisahkan atau membagi data menjadi kuantil-kuantil tertentu yang kemungkinan memiliki nilai dugaan yang berbeda.

Regresi kuantil meminimumkan galat mutlak terboboti dan menduga model dengan menggunakan fungsi kuantil bersyarat pada suatu sebaran data. Metode regresi kuantil tidak membutuhkan asumsi parametrik dan regresi kuantil sangat bermanfaat untuk menganalisis bagian tertentu dari suatu sebaran bersyarat [3]. Keuntungan utama dari regresi kuantil adalah efisien jika sisaan tidak menyebar normal dan kekar terhadap adanya pencilan.

Untuk peubah acak  $Y$  dengan fungsi sebaran peluang

$$F(y) = P(Y \leq y)$$

kuantil ke- $\tau$  dari  $Y$  didefinisikan sebagai fungsi invers

$$Q(\tau) = \inf\{y, F(y) \geq \tau\}$$

dengan  $\tau \in (0,1)$ , sebagai contoh median adalah

$Q(0.5)$

Untuk contoh acak berukuran  $n$  dari peubah acak  $Y$ , yaitu  $(y_1, \dots, y_n)$ , median contoh adalah penduga yang meminimumkan jumlah mutlak galat yaitu

$$\min_{\xi \in R} \sum_{i=1}^n |y_i - \xi|$$

seperti halnya median contoh, metode ini bisa dikembangkan untuk model regresi kuantil

$$y = X'\beta + \varepsilon$$

dengan  $y = (y_1, \dots, y_n)$  adalah vektor respon berukuran  $(n \times 1)$ ,  $X = (x_1, \dots, x_n)'$  adalah matriks peubah penjelas berukuran  $(n \times p)$ ,  $\beta = (\beta_1, \dots, \beta_p)'$  adalah vektor parameter berukuran  $(p \times 1)$ , dan  $\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)$  adalah vektor galat berukuran  $(n \times 1)$ . Regresi  $L_1$  disebut sebagai regresi median yang merupakan perluasan dari median contoh. Penduga koefisien pada model regresi  $L_1$  merupakan solusi dari minimisasi fungsi

$$\min_{\beta \in R} \sum_{i=1}^n |y_i - x_i'\beta|$$

Hubungan fungsional antara peubah respon dengan peubah penjelas pada regresi kuantil merupakan hubungan fungsional antara kuantil bersyarat peubah respon dengan peubah penjelas yang membentuk fungsi linier yaitu  $Q(\tau|X = x) = x'\beta(\tau)$ . Menurut [9], penduga regresi kuantil ke- $\tau$  untuk  $\tau \in (0,1)$  merupakan solusi dari masalah minimisasi fungsi

$$\min_{\beta \in R} \left[ \sum_{i \in \{i|y_i \geq x_i'\beta\}} \tau |y_i - x_i'\beta| + \sum_{i \in \{i|y_i < x_i'\beta\}} (1 - \tau) |y_i - x_i'\beta| \right] \tag{1}$$

Persamaan (1) memberikan bobot  $\tau$  untuk seluruh

pengamatan yang lebih besar dari nilai optimum yang belum diketahui dan memberikan bobot  $1 - \tau$  terhadap seluruh pengamatan yang lebih kecil dari nilai optimum. Persamaan (1) dapat diringkas menjadi persamaan dengan ekspresi tunggal seperti pada persamaan (2).

$$\min_{\beta \in R} \sum_{i=1}^n \rho_\tau(y_i - x_i'\beta) \tag{2}$$

dengan  $\rho_\tau(u)$  adalah fungsi kerugian yang tidak simetrik. Fungsi kerugian dapat dijabarkan sebagai berikut:

$$\rho_\tau(u) = u(\tau - I(u < 0)) \quad , \quad 0 < \tau < 1$$

dengan  $I(\cdot)$  merupakan fungsi indikator,  $I(A) = 1$  jika A benar dan  $I(A) = 0$  selainnya.

Penduga regresi kuantil ke- $\tau$  untuk  $\tau \in (0,1)$  dengan penalti lasso dapat ditulis sebagai berikut

$$\min_{\beta \in R} \sum_{i=1}^n \rho_\tau(y_i - x_i'\beta) + \lambda \sum_{j=1}^p |\beta_j|$$

Pendugaan dalam regresi kuantil diperoleh dengan menyelesaikan masalah pemrograman linier. Salah satu metode yang dapat digunakan untuk pendugaan parameter regresi kuantil adalah metode simpleks.

Penilaian kebaikan model dilakukan dengan menghitung *Quantile Verification Skill Score* (QVSS). QVSS didefinisikan sebagai berikut :

$$QVSS = 1 - \frac{\sum_{i=1}^n \rho_\tau |y_i - \hat{\beta}_\tau^T x_i|}{\sum_{i=1}^n \rho_\tau |y_i - Q_\tau(y)|}$$

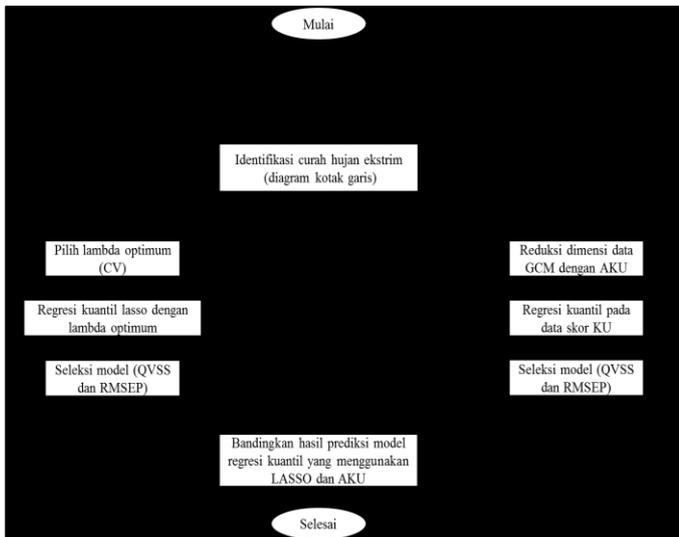
dengan  $Q_\tau(y)$  merupakan kuantil ke- $\tau$  dari  $y$ .

### 2.3 Data dan Digram Alir Metode Analisis

Data luaran GCM yang digunakan adalah data presipitasi bulanan *Climate Model Intercomparison Project* (CMIP5) yang dikeluarkan oleh KNMI. Belanda dari situs web

<http://www.climatexp.knmi.nl/> pada tahun 1979 sampai dengan 2008 dengan posisi wilayah – 18.75° – 1.25°LS dan 101.25° – 116.25°BT. Data curah hujan lokal yang digunakan adalah data rata-rata curah hujan di Kabupaten Indramayu Propinsi Jawa Barat.

Diagram alir metode analisis sebagai berikut:



Gambar 2. Diagram alir metode analisis

Jumlah model yang akan dibangun pada penelitian ini sebanyak 6 model yang terdiri dari:

1. M1 adalah model regresi kuantil linear
2. M2 adalah model regresi kuantil linear dengan tambahan peubah dummy pada setiap bulan
3. M3 adalah model regresi kuantil linear dengan tambahan peubah dummy untuk empat bulanan. Model ini didasarkan pada penelitian [12] yang mengatakan bahwa curah hujan dibagi dalam 4 kelompok bulan yakni bulan basah (Januari, Februari, Desember), kelompok bulan peralihan bulan basah ke bulan kering (Maret, April dan Mei), kelompok bulan kering (Juni, Juli dan Agustus) dan kelompok bulan peralihan bulan kering ke bulan basah (September, Oktober

SD dengan Regresi kuantil menggunakan Lasso dan AKU

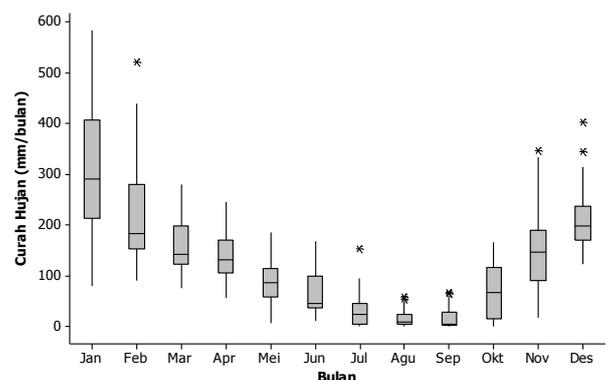
dan November) sehingga untuk menampung informasi tersebut bisa ditambahkan peubah dummy dalam model sesuai dengan kelompok bulan yang ada.

4. M4 adalah model regresi kuantil kuadratik, model ini didasari oleh penelitian [5] yang mengatakan curah hujan bulanan di Indramayu memiliki pola kuadratik.
5. M5 adalah model regresi kuantil kuadratik dengan penambahan dummy untuk setiap bulan.
6. M2 adalah model regresi kuantil kuadratik dengan penambahan peubah dummy untuk empat bulanan.

### 3. Hasil Dan Pembahasan

#### 3.1 Deskripsi Data

Gambar 3 merupakan diagram kotak garis dari data curah hujan di Kabupaten Indramayu menunjukkan bahwa pada bulan Februari, Juli, Agustus, September, November dan Desember terdapat curah curah hujan bulanan yang lebih tinggi dari kondisi normalnya hal ini membuktikan bahwa curah hujan di Indramayu terindikasi terjadi curah hujan ekstrim.



Gambar 3. Deskripsi Data curah Hujan Kabupaten Indramayu

#### 3.2 Regresi Kuantil dengan LASSO

Korelasi tertinggi antar peubah GCM adalah sebesar 0.99 yang terletak pada grid 75 dan grid 65 di mana kedua grid ini saling berdekatan. Semakin dekat letak antar grid maka akan memiliki korelasi yang semakin tinggi dan sebaliknya.

Kendala yang sering dihadapi oleh peneliti adalah tentang penentuan luasan grid dari data GCM. Terlalu sempit luasan grid/domain yang digunakan akan mengurangi informasi pengaruh global/regional. Sebaliknya, luasan grid yang terlalu luas menyebabkan informasi lokal akan berkurang [12]. Pada penelitian ini akan dicobakan jumlah grid grid  $8 \times 8$  dan  $6 \times 6$  grid. Hasil prediksi dari kedua jumlah grid tersebut dibandingkan untuk mengetahui jumlah grid mana yang memiliki tingkat ketepatan prediksi yang lebih baik.

Nilai lambda optimum yang diperoleh untuk Model M1 yaitu sebesar 0.35 dengan nilai CV sebesar 38.75. Tabel 2 adalah nilai lambda optimum dari LASSO pada masing-masing model pada kedua ukuran luasan grid.

Tabel 1. Nilai lambda optimum untuk masing-masing model

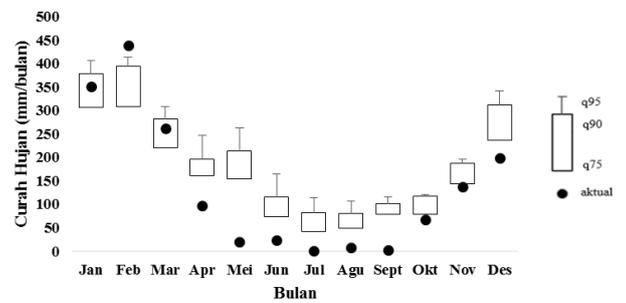
Model	Lambda optimum	
	Grid $8 \times 8$	Grid $6 \times 6$
M1	0.35	21.14
M2	0.22	0.41
M3	1.10	21.63
M4	94.70	0.45
M5	6.92	0.49
M6	94.70	0.45

Masing-masing model akan dihitung nilai QVSS dan RMSEP. Gambar 4 menunjukkan bahwa untuk grid  $8 \times 8$  nilai QVSS tertinggi ada pada Model M4 (kuantil kuadrat), model

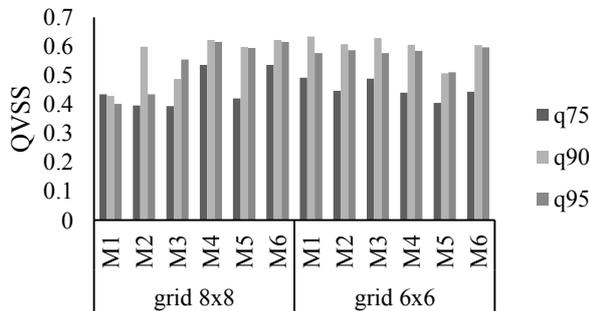
dengan nilai QVSS tertinggi adalah model terbaik. Nilai RMSEP dari Model M4 dapat dilihat pada gambar 5 yaitu 71.85 untuk kuantil ke 75, 103.88 untuk kuantil ke 90 dan 135.56 untuk kuantil 95. Nilai RMSEP ini relatif kecil jika dibandingkan dengan model lainnya. Sedangkan pada grid  $6 \times 6$  model M1 memiliki nilai QVSS tertinggi (Gambar 4) dan memiliki nilai RMSEP terendah, sehingga diperoleh model terbaik adalah Model M4 pada grid  $8 \times 8$  dan M1 pada grid  $6 \times 6$ . Diketahui kedua model tersebut merupakan model yang tidak memiliki peubah boneka. Hal ini mengindikasikan bahwa penambahan peubah boneka atau pengelompokan bulan tidak memberikan pengaruh yang signifikan terhadap kebaikan hasil prediksi. Hal ini juga terjadi pada penelitian [10] dimana model terbaik yang didapatkan adalah model tanpa peubah boneka. Secara grafis juga dapat dilihat pada gambar 5 bahwa nilai RMSEP model pada grid  $6 \times 6$  cenderung lebih kecil dibandingkan model pada grid  $8 \times 8$  sehingga dapat dikatakan bahwa grid  $6 \times 6$  memiliki ketepatan prediksi yang lebih baik dibandingkan dengan model pada grid  $8 \times 8$ .

Prediksi curah hujan untuk tahun 2008 dilakukan dengan menggunakan model terbaik. Gambar 6 adalah hasil prediksi curah hujan oleh model M4 dengan jumlah grid  $8 \times 8$ . Pada Model M4 dapat dilihat bahwa curah hujan ekstrim pada bulan februari dapat diprediksi dengan baik oleh Model M4 pada kuantil ke 95. Bulan Februari merupakan intensitas curah hujan tertinggi yang terjadi di tahun 2008 dengan nilai 439 mm/bulan. Nilai ini dapat diestimasi dengan baik oleh

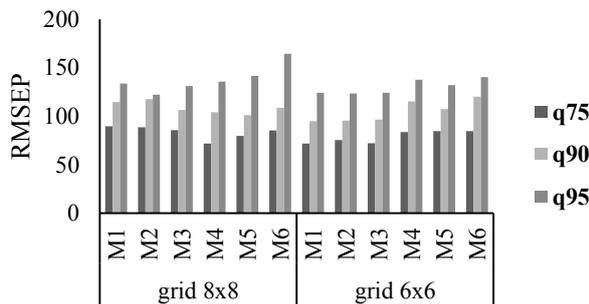
prediksi pada kuantil ke-90 yakni 429.177 mm/bulan. Model M1 juga dapat menduga curah hujan pada bulan Januari dan Maret dengan baik. Secara umum, untuk bulan-bulan yang berada di musim kemarau (April-September), nilai prediksi pada kuantil ke-75, ke-90, dan ke-95 lebih tinggi dari nilai aktual, namun mampu mengikuti pola curah hujan aktual dengan baik. Hal yang sama terjadi pada model terbaik yang menggunakan grid 6×6 (Model M1) pada gambar 7 dimana curah hujan pada bulan Januari dan Maret dapat diprediksi dengan baik oleh Model M1 pada kuantil ke 90.



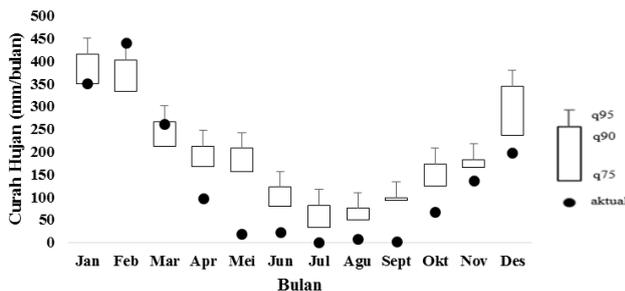
Gambar 7. Nilai prediksi dari Model M1 untuk metode LASSO pada grid 6×6



Gambar 4. Nilai QVSS dari keenam model yang menggunakan LASSO untuk grid 8×8 dan grid 6×6



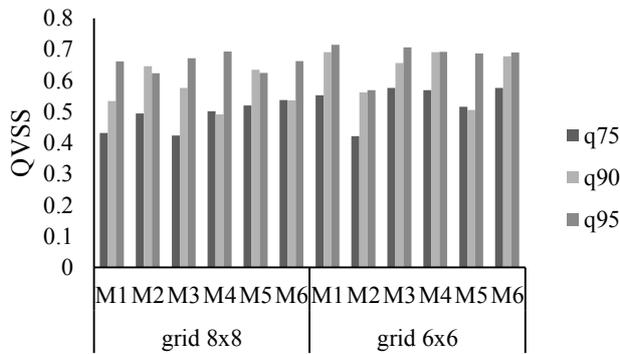
Gambar 5. Nilai RMSEP keenam model yang menggunakan LASSO untuk grid 8×8 dan grid 6×6



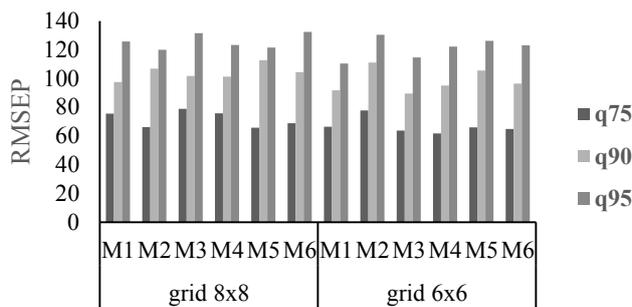
Gambar 6. Nilai prediksi dari Model M4 untuk metode LASSO pada grid 8×8

### 3.3 Regresi Kuantil dengan AKU

Langkah pertama yang dilakukan sebelum memodelkan regresi kuantil adalah melakukan reduksi dimensi GCM dengan menggunakan AKU. Pemilihan jumlah komponen utama yang digunakan dengan menggunakan kriteria ragam kumulatif dari komponen utama (KU). Jumlah KU yang terpilih adalah KU dengan ragam kumulatif yang lebih dari 95%. Sama halnya dengan metode LASSO kriteria pemilihan model terbaik pada metode AKU juga menggunakan nilai QVSS, model dengan nilai QVSS tertinggi merupakan model terbaik. Gambar 8 menunjukkan nilai QVSS masing-masing model untuk jumlah grid 8×8 dan grid 6×6. Pada jumlah grid 8×8 model terbaik ada pada Model M5 sedangkan pada jumlah grid 6×6 model terbaik ada pada Model M1. Gambar 9 juga menunjukkan bahwa RMSEP antara model yang menggunakan grid 8×8 dan grid 6×6 tidak menunjukkan perbedaan yang signifikan sehingga tidak bisa diketahui jumlah grid yang paling baik untuk model yang menggunakan metode AKU.

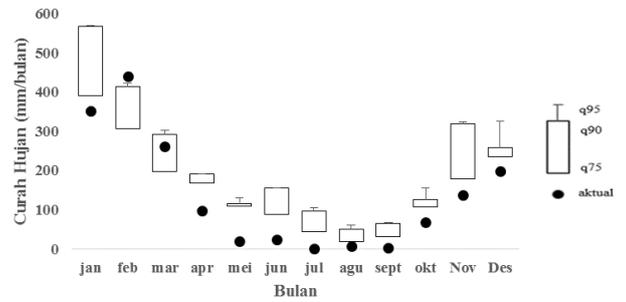


Gambar 8. Nilai QVSS dari enam model yang menggunakan AKU untuk grid 8x8 dan grid 6x6

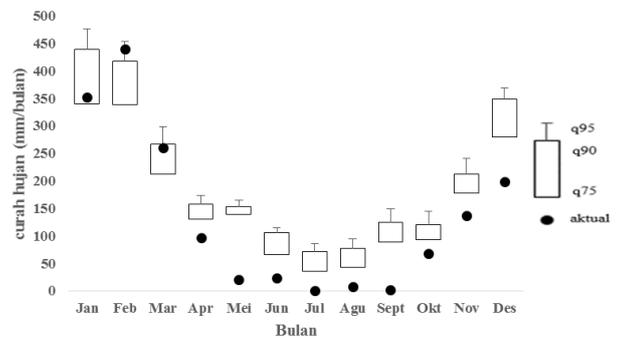


Gambar 9. Nilai RMSEP dari enam model yang menggunakan AKU untuk grid 8x8 dan grid 6x6

Prediksi curah hujan untuk tahun 2008 dengan menggunakan model terbaik. Gambar 10 adalah hasil prediksi curah hujan untuk model dengan jumlah grid 8x8 (Model M2) dan gambar 11 untuk hasil prediksi curah hujan model yang menggunakan grid 6x6 (Model M1). Dapat dilihat pada gambar 10 bahwa prediksi M2 menunjukkan pola yang sedikit berbeda dari curah hujan sebenarnya dimana prediksi bulan Februari lebih rendah dari bulan Januari. M1 pada grid 6x6 menunjukkan pola prediksi yang lebih baik dari M2 pada grid 8 x 8 dimana M1 dapat mengikuti pola data aktual dengan baik. M1 juga dapat menduga curah hujan pada bulan Januari, Februari dan Maret dengan baik.



Gambar 10. Nilai prediksi dari Model M2 untuk metode AKU pada grid 8x8



Gambar 11. Nilai prediksi dari Model M1 untuk metode AKU pada grid 6x6

### 3.4 Perbandingan SD Regresi Kuantil menggunakan LASSO dan Analisis Komponen Utama

Perbandingkan hasil prediksi dari model SD yang menggunakan metode LASSO dan yang menggunakan metode AKU dapat dilihat dari nilai RMSEP. Menurut [7], pada dasarnya ketepatan hasil prediksi yang diperoleh dari regresi LASSO dan AKU saling berkompetisi. Salah satu kelebihan metode LASSO dibandingkan AKU terletak pada model yang dihasilkan dimana model yang menggunakan LASSO lebih sederhana dan mudah diinterpretasi. [1] juga membandingkan regresi LASSO dan AKU dengan menggunakan kriteria RMSEP dimana diketahui bahwa regresi LASSO memiliki nilai RMSEP yang lebih rendah

dibandingkan dengan regresi AKU. Pada kasus ini diketahui bahwa nilai RMSEP model yang menggunakan metode LASSO dan AKU tidak berbeda signifikan atau cenderung sama. Tabel 3 adalah perbandingan nilai RMSEP dari kedua metode untuk panjang data yang berbeda. Dapat diketahui bahwa nilai RMSEP dari LASSO khususnya pada kuantil 90 memiliki standar deviasi yang kecil, hal ini menunjukkan bahwa model LASSO khususnya model pada kuantil ke 90 memiliki hasil prediksi yang lebih konsisten dibandingkan dengan model yang menggunakan metode AKU.

Tabel 2. Perbandingan nilai RMSEP di setiap panjang data pendugaan model SD dengan regresi kuantil menggunakan LASSO dan AKU

Data Model	Data Validasi	RMSEP			
		LASSO		AKU	
		Q90	Q95	Q90	Q95
1979-2007	2008	90.71	118.92	91.70	110.46
1979-2006	2007-2008	89.48	119.99	115.36	121.73
1979-2005	2006-2008	95.30	133.62	102.93	126.45
1979-2004	2005-2008	98.95	147.50	110.68	133.21
1979-2003	2004-2008	106.32	141.30	116.37	136.53
Standar deviasi		6.82	11.35	10.26	10.28

#### 4. Kesimpulan Dan Saran

Berdasarkan hasil analisis curah hujan ekstrim di Kabupaten Indramayu, dapat disimpulkan bahwa :

1. Curah hujan ekstrim di Kabupaten Indramayu pada bulan basah Januari dan Maret dapat diprediksi baik dengan model yang menggunakan metode LASSO pada kuantil ke 90 sedangkan pada bulan-bulan lainnya nilai

prediksi curah hujan cenderung lebih tinggi dibandingkan dengan nilai aktualnya tetapi model yang didapatkan dapat mengikuti pola data curah hujan yang sebenarnya.

2. Berdasarkan kriteria RMSEP model yang menggunakan metode LASSO untuk jumlah grid 6×6 menghasilkan nilai prediksi yang lebih baik dibandingkan grid 8×8 sedangkan untuk model yang menggunakan metode AKU tidak ada perbedaan nilai prediksi antara kedua jumlah grid.

Pada penelitian ini, *statistical downscaling* dengan regresi kuantil menggunakan metode *shrinkage* LASSO untuk penelitian selanjutnya dapat dicobakan pemodelan *statistical downscaling* dengan regresi kuantil menggunakan metode *shrinkage* RIDGE untuk mengetahui perbandingan prediksi dari kedua metode tersebut.

#### 5. Ucapan Terima Kasih

Pada artikel ini penulis ucapkan terima kasih kepada pihak-pihak yang telah terlibat dalam membantu penulisan artikel ini.

#### Daftar Pustaka

- [1] Acharjee A, Finkers R, Visser RGF, Maliepaard C. 2013. Comparison of regularized regression methods for omics data. *Metabolomics* 3:126.
- [2] [BPS]. Badan Pusat Statistik. 2011. Indramayu dalam Angka 2011. Indramayu : BPS Kabupaten Indramayu.
- [3] Buhai S. 2004. Quantile regression overview and selected application. *Ad Astra*. 4:1-17.
- [4] Chen C, Wei Y. 2005. Computational Issues on Quantile Regression. *The Indian*

- Journal of Statistics*. 67:299-417.
- [5] Djuraidah A, Wigena AH. 2011. Regresi kuantil untuk eksplorasi pola curah hujan di Kabupaten Indramayu. *Jurnal Ilmu Dasar*. 12(1): 50 – 56.
- [6] Handayani L. 2014. Statistical downscaling dengan model aditif terampat untuk pendugaan curah hujan ekstrim [tesis]. Institut Pertanian Bogor.
- [7] Hastie T, Tibshirani R, Friedman J. 1990. *The Elements of Statistical Learning*. Ed ke-2. Springer, London (GB).
- [8] Koenker R, W Bassett, Gilbert Jr .1978. Regression quantiles econometric society. *Econometrica*. 46(1):33–50.
- [9] Koenker R. 2005. *Quantile Regression*. Cambridge University Press, New York.
- [10] Mondiana YQ. 2012. Pemodelan statistical downscaling dengan regresi kuantil untuk pendugaan curah hujan ekstrim [tesis]. Institut Pertanian Bogor.
- [11] Stockdale TN, Busalacchi AJ, Harrison DE, Seager R. 1998. Ocean modeling for ENSO. *Journal of Geophysical research*. 103:14325-14355.
- [12] Sutikno. 2008. Statistical downscaling luaran GCM dan pemanfaatannya untuk peramalan produksi padi, [disertasi]. Institut Pertanian Bogor.
- [13] Wigena AH. 2006. Pemodelan statistical downscaling dengan regresi projection pursuit untuk peramalan curah hujan bulanan (kasus curah hujan bulanan di Indramayu) [disertasi]. Bogor (ID): Institut Pertanian Bogor.
- [14] Wilby RL, Wigley TML. 2000. Precipitation predictors for downscaling: observed and general circulation model relationships. *Journal of Climatology*. 20: 641-661.
- [15] Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. 2009. A review of climate risk information for adaptation and development planning. *Journal of Climatology*. 29: 1193-1215.