

# PEMODELAN *STATISTICAL DOWNSCALING* UNTUK MENDUGA CURAH HUJAN BULANAN MENGGUNAKAN MODEL LINIER TERAMPAT SEBARAN GAMMA

## *Statistical Downscaling to Predict Monthly Rainfall Using Generalized Linear Model with Gamma Distribution*

Agus M Soleh, Aji H Wigena, Anik Djuraidah, Asep Saefuddin

Departemen Statistika, Institut Pertanian Bogor, Kampus IPB Darmaga Bogor 16680

Telp. (0251) 8624535, Fax. (0251) 8624535

E-mail: agusms@apps.ipb.ac.id

(Makalah diterima, 7 Juli 2015 – Disetujui, 4 Desember 2015)

### ABSTRAK

Pemodelan *statistical downscaling* (SDS) umumnya menggunakan kovariat terkondisi buruk (dimensi besar dan memiliki korelasi tinggi/multikolinieritas). Permasalahan ini dapat diatasi dengan teknik seleksi variabel menggunakan regularisasi L1/LASSO atau pereduksian dimensi menggunakan analisis komponen utama (AKU). Kedua teknik diaplikasikan ke dalam pemodelan linier terampat sebaran gamma dan diperbandingkan untuk pendugaan model curah hujan di 11 pos hujan Kabupaten Indramayu dan sekitarnya menggunakan pemodelan SDS. Model linier terampat sebaran gamma digunakan untuk mendapatkan dugaan curah hujan non-negatif dan diperbandingkan dengan pemodelan regresi komponen utama (RKU) sebagai metode umum dalam pemodelan SDS. Kovariat yang digunakan dalam pemodelan SDS adalah dua tipe kovariat terkondisi buruk (luaran CMIP5 dan data observasi GPCP versi 2.2) yang memiliki karakteristik berbeda. Berdasarkan nilai statistik Root Mean Square Error (RMSE) terkecil, ketiga pemodelan (RKU, Gamma-KU, dan Gamma-L1) memberikan nilai yang relatif tidak berbeda secara signifikan setelah penambahan peubah dummy bulan. Dalam kasus ini, pemodelan linier terampat sebaran gamma dapat dianggap sebagai model terbaik karena memberikan dugaan curah hujan non-negatif.

**Kata kunci:** curah hujan, *statistical downscaling*, model linier terampat, respons sebaran gamma

### ABSTRACT

*Statistical Downscaling (SDS) models might involve ill-conditioned covariates (large dimension and high correlation/multicollinear). This problem could be solved by a variable selection technique using L1 regularization/LASSO or a dimension reduction approach using principal component analysis (PCA). In this paper, both methods were applied to generalized linear modeling with gamma distribution and compared to predict rainfall models at 11 rain posts in Indramayu. More over, generalized linear model with gamma distribution was used to obtain non-negative rainfall prediction and compared with principal component regression (PCR). Two types of ill-conditioned data with different characteristics (CMIP5 and GPCP version 2.2) were used as covariates in SDS modeling. The results show that three methods (PCR, Gamma-PC, and Gamma-L1) did not demonstrate significant difference in term of Root Mean Square Error (RMSE) after addition of dummy variables (month) in the models. However, a generalized linear modeling with gamma distribution could be considered as the best methods since it provided non-negative rainfall predictions.*

**Key words:** rainfall, *statistical downscaling*, generalized linear model, gamma distribution response

## PENDAHULUAN

Curah hujan adalah variabel yang sangat penting dalam proses pertanian. Beberapa metode untuk menduga curah hujan telah banyak dilakukan di Indonesia, salah satunya menggunakan teknik pemodelan *Statistical Downscaling* (SDS) (Wigena 2006; Sutikno 2008; Vimont *et al.*, 2010; Sutikno *et al.*, 2010). Pemodelan SDS merupakan salah satu penerapan pemodelan dengan kovariat yang umumnya besar dan tidak saling bebas. *Downscaling* (Benestad *et al.*, 2008) merupakan proses untuk membuat hubungan antara beberapa "skala besar" dengan peubah yang merepresentasikan "skala kecil". Data berskala besar umumnya merupakan luaran *General Circulation Model* (GCM) dan data berskala kecil berupa iklim lokal seperti curah hujan di wilayah tertentu.

Nilai curah hujan merupakan peubah yang dicatat dalam rentang nilai  $\geq 0$ . Pemodelan curah hujan dengan teknik regresi biasa yang mengasumsikan respons berasal dari sebaran normal tidak secara alami memperoleh dugaan selalu  $\geq 0$ . Oleh karena itu pemodelan curah hujan dengan respons non-normal perlu dipertimbangkan untuk digunakan. Salah satu sebaran dengan rentang nilai  $\geq 0$  adalah sebaran gamma dan pemodelannya termasuk kelas pemodelan linier terampat (*Generalized Linear Model/ GLM*). Pemodelan dengan respons sebaran gamma ditujukan untuk pemodelan dengan respons kontinu dan menjulur (*skewed*) (Faraway, 2006). Dalam hal ini curah hujan dimodelkan sebagai fungsi kovariat presipitasi "skala besar" (luaran GCM CMIP5 dan observasi GPCP versi 2.2) dengan teknik pemodelan SDS.

GCM merupakan model numerik yang menghasilkan sejumlah data dari berbagai parameter iklim seperti presipitasi, temperatur, dan kelembaban untuk pendugaan iklim. Model GCM merepresentasikan kemungkinan cara terbaik mensimulasi kondisi iklim "skala besar" dan memproyeksikan perubahan iklim ke depan akibat pengaruh kekuatan yang diketahui (seperti pengaruh gas rumah kaca). Kemampuan pendugaan skala kecil menggunakan GCM sangat terbatas karena resolusi spasial dalam GCM umumnya kasar berukuran  $2.5^0 \times 2.5^0$  ( $\pm 300 \text{ km} \times 300 \text{ km}$ ). Pemodelan SDS menggunakan kovariat "skala besar" dapat dilakukan menggunakan teknik seleksi peubah dan analisis komponen utama (AKU).

Seleksi peubah merupakan isu penting dalam pemodelan regresi, terutama ketika jumlah kovariat sangat besar dan tidak saling bebas seperti dalam luaran GCM. Tibshirani (1996) mengusulkan LASSO (*least absolute shrinkage and selection operator*), metode baru untuk seleksi peubah dan penyusutan (*shrinkage*) koefisien kovariat. LASSO menambahkan penalti  $L_1$  ( $\sum |\beta_i| \leq t$  atau disebut regularisasi  $L_1$ ) terhadap fungsi tujuan pendugaan model yang menghasilkan dua keuntungan: seleksi dan

penyusutan. Kegunaan dari penyusutan adalah untuk mencegah timbulnya *overfit* akibat kolinieritas kovariat atau dimensional tinggi (Hastie *et al.*, 2008). Penerapan LASSO dalam pemodelan SDS menunjukkan hasil yang lebih baik dibanding metode regresi stepwise (Hammami *et al.*, 2012; Gao *et al.*, 2014).

Pengkajian ini bertujuan untuk mendapatkan dugaan curah hujan selalu non-negatif ( $\geq 0$ ) secara alami sesuai karakteristiknya, dan penggunaan teknik LASSO sebagai solusi mengatasi masalah data terkondisi buruk dibandingkan dengan teknik AKU yang selama ini sering digunakan dalam pemodelan SDS.

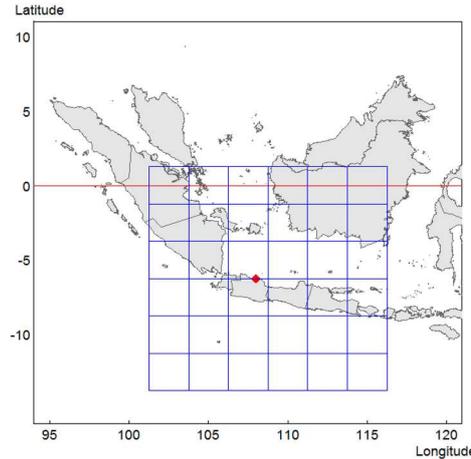
## BAHAN DAN METODE

### Data

Kajian pada pemodelan SDS menggunakan dua tipe data, yaitu data presipitasi bulanan dari dua skala besar sebagai kovariat dan data curah hujan bulanan dari 11 pos hujan di Kabupaten Indramayu dan sekitarnya sebagai respons. Berdasarkan peta prakiraan musim, wilayah Kabupaten Indramayu terbagi menjadi empat zona musim (ZOM) yaitu ZOM 77, 78, 79, dan 80. Pada penelitian ini digunakan curah hujan bulanan dari pos hujan yang mewakili tiga ZOM, yaitu ZOM 77 (Kr Anyar, Pusakanegara dan Tulang Kacang), ZOM 78 (Dempet, Indramayu, Juntinyuat dan Losarang), dan ZOM 79 (Gegesik, Karangendal, Krangken dan Sukadana).

Presipitasi dua "skala besar" yang digunakan dan diperbandingkan sebagai kovariat, yaitu luaran GCM dari *ensemble multi-model Phase 5 Couple Model Intercomparisson Project* (CMIP5) dengan skenario perubahan iklim moderat RCP (*Representative Concentration Pathways*) 4.5 (Taylor *et al.*, 2012) dan data observasi *Global Precipitation Climatology Project* (GPCP) versi 2.2 (Adler *et al.*, 2003). Luaran GCM CMIP5 merupakan luaran dari *ensemble* beberapa model proyeksi iklim sampai tahun 2100 pada skenario perubahan iklim tertentu, sedangkan data observasi GPCP versi 2.2 merupakan kombinasi data observasi permukaan dan satelit dalam bentuk grid.

Data presipitasi luaran CMIP5 dapat diunduh melalui <http://pcmdi9.llnl.gov/esgf-web-fe/> dan data observasi presipitasi GPCP diberikan oleh "the NOAA/OAR/ESRL PSD, Boulder, Colorado, USA", dari web sitenya di <http://www.esrl.noaa.gov/psd/>. Kedua kovariat "skala besar" diambil dalam domain  $7 \times 7$  grid (49 kovariat) pada sistem koordinat  $101,25^0$ - $116,25^0$  BT dan  $13,75^0$  LS -  $1,25^0$  LU dengan lebar grid sebesar  $2.5^0 \times 2.5^0$ . Posisi Kabupaten Indramayu berada di bawah grid tengah domain yang dipilih (Gambar 1).



Gambar 1. Domain grid kovariat yang dipilih dalam penelitian

## Metode

Pemodelan SDS digunakan untuk mendapatkan model yang menghubungkan data presipitasi luaran GCM dari CMIP5 dan data observasi GPCP versi 2.2 dengan curah hujan di 11 pos hujan di Kabupaten Indramayu dan sekitarnya. Pemodelan dilakukan dengan teknik pemodelan linier terampat menggunakan respons yang diasumsikan berasal dari sebaran gamma. Teknik seleksi menggunakan LASSO untuk mengatasi masalah kovariat yang dibandingkan dengan teknik pereduksian dimensi menggunakan AKU. Langkah pendugaan model adalah sebagai berikut:

### 1. Penyiapan data

Penyiapan data menggunakan software komputasi statistik R dengan paket tambahan `ncdf4` pada sistem operasi Linux untuk membaca data kovariat “skala besar” dalam format `netCDF`. Data kovariat “skala besar” selanjutnya direduksi dimensinya menggunakan metode AKU untuk mendapatkan peubah laten yang ortogonal dalam bentuk skor komponen utama. Banyaknya peubah laten ditentukan oleh grafik `screplot`, proporsi keragaman, dan nilai akar ciri.

### 2. Pendugaan model linier rata-rata curah hujan

Pendugaan model linier dilakukan terhadap respons curah hujan dari 11 pos hujan dengan presipitasi kovariat “skala besar” (luaran GCM CMIP5 dan observasi GPCP versi 2.2).

#### a. Pendugaan menggunakan pemodelan linier terampat sebaran gamma dengan pra pemrosesan AKU.

Sebaran gamma termasuk ke dalam keluarga sebaran eksponensial dengan dua parameter, yaitu, *shape* ( $\xi$ ) dan *rate* ( $\theta$ ), dalam hal ini fungsi

kepekatan peluang dari sebaran berada pada range  $(0, \infty)$ :

$$f_Y(y|\theta, \xi) = \frac{\theta^\xi}{\Gamma(\xi)} y^{\xi-1} \exp(-\theta y).$$

Pemodelan linier dengan respons dari keluarga eksponensial termasuk metode model linier terampat (GLM) untuk menduga nilai tengah respons. Oleh karena itu, untuk keperluan pendugaan respons dari sebaran gamma diperlukan parameterisasi ulang terhadap parameter *rate* yaitu  $\theta = \xi/\mu$ . Pendugaan parameter model dilakukan dengan kovariat hasil pereduksian dimensi menggunakan AKU. Fungsi hubung yang digunakan dalam pemodelan adalah fungsi hubung kanonik yaitu *inverse/reciprocal*.

- b. Pendugaan menggunakan pemodelan linier terampat sebaran gamma dengan regularisasi  $L_1$ . Pendugaan parameter menggunakan optimisasi fungsi log kemungkinan sebaran gamma. Secara umum pendugaan parameter GLM dengan regularisasi  $L_1$  dalam bentuk lagrange yang dituliskan sebagai ( $\lambda \geq 0$ ):

$$\underset{\beta_k}{\operatorname{argmin}} \left\{ -\log[L(\beta_k; \mathbf{y})] + \lambda \sum_{k=1}^p |\beta_k| \right\}$$

Fungsi log kemungkinan untuk observasi tunggal sebaran gamma adalah:

$$l(\mu; y) = \xi \log\left(\frac{\xi}{\mu}\right) - \log \Gamma(\xi) + (\xi - 1) \log y - \left(\frac{\xi}{\mu} y\right)$$

Banyaknya kovariat yang terselksi dalam pendugaan model ditentukan oleh nilai validasi silang *10-fold* terkecil.

### 3. Pemilihan metode pendugaan terbaik

Nilai *Root Means Square Error* (RMSE) terkecil digunakan untuk menentukan metode terbaik. Nilai statistik RMSE merupakan metode untuk mengukur perbedaan antara nilai prediksi dengan nilai aktual yang didefinisikan sebagai berikut:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{n}}$$

$Y_i$  adalah nilai observasi/aktual dan  $\hat{Y}_i$  adalah nilai dugaan ke- $i$ .

## HASIL DAN PEMBAHASAN

### Karakteristik Curah Hujan 11 Pos Hujan di Kabupaten Indramayu dan Sekitarnya

Secara umum Kabupaten Indramayu merupakan wilayah yang datar dan tidak terdapat dataran tinggi atau pegunungan dengan kemiringan tanah rata-rata 0%-2%. Sebagian besar wilayah Indramayu bagian utara mempunyai klasifikasi agroklimatik *seasonally dry* (kering musiman) 5-8 bulan per tahun dengan rata-rata curah hujan kurang dari 100 mm/bulan, sedangkan bagian selatan *permanently moist* (tetap basah) selama 0-4 bulan dengan rata-rata curah hujan kurang dari 100 mm berdasarkan tinjauan sumber daya lahan oleh Bakosurtanal tahun 1990 (Haryoko, 2015).

Pola sebaran curah hujan bulanan di 11 pos hujan yang berada pada wilayah ZOM 77, 78 dan 79 di Kabupaten Indramayu dan sekitarnya menunjukkan pola musonal, yaitu berbentuk U (Gambar 2). Curah hujan tertinggi terjadi pada bulan Januari dan terendah antara bulan Juli-September. Rata-rata curah hujan terendah di wilayah ZOM 77 dimulai sekitar bulan Juli, dan wilayah ZOM 78 dan ZOM 79 sekitar bulan Agustus, sedangkan awal musim hujan ketiga ZOM dimulai pada bulan Oktober. Secara umum curah hujan pada bulan tertentu untuk semua pos hujan menunjukkan pola yang menjulur ke kanan, sehingga pemodelan menggunakan model linier terampat sebaran gamma dengan penambahan peubah *dummy* bulan diprediksi akan meningkatkan ketepatan dalam dugaan curah hujan.

### Karakteristik Kovariat Skala Besar: Luaran CMIP5 dan Observasi GPCP versi 2.2

Terdapat perbedaan karakteristik dua kovariat “skala besar” yang digunakan dalam penelitian ini, yaitu perbedaan berdasarkan nilai koefisien korelasi pasangan kovariat dan hasil pereduksian dimensi menggunakan teknik analisis komponen utama. Banyaknya pasangan yang memiliki nilai korelasi  $> 0,6$  dari data luaran CMIP5

sebanyak 91,24% (1073 dari 1176 pasangan), sedangkan dari data observasi GPCP versi 2,2 sebanyak 23,55% (277 dari 1176 pasangan). Koefisien korelasi semua pasangan dari luaran CMIP5 memiliki tanda positif, sedangkan observasi GPCP versi 2,2 beberapa diantaranya memiliki tanda negatif, yaitu 308 pasangan (26,19%).

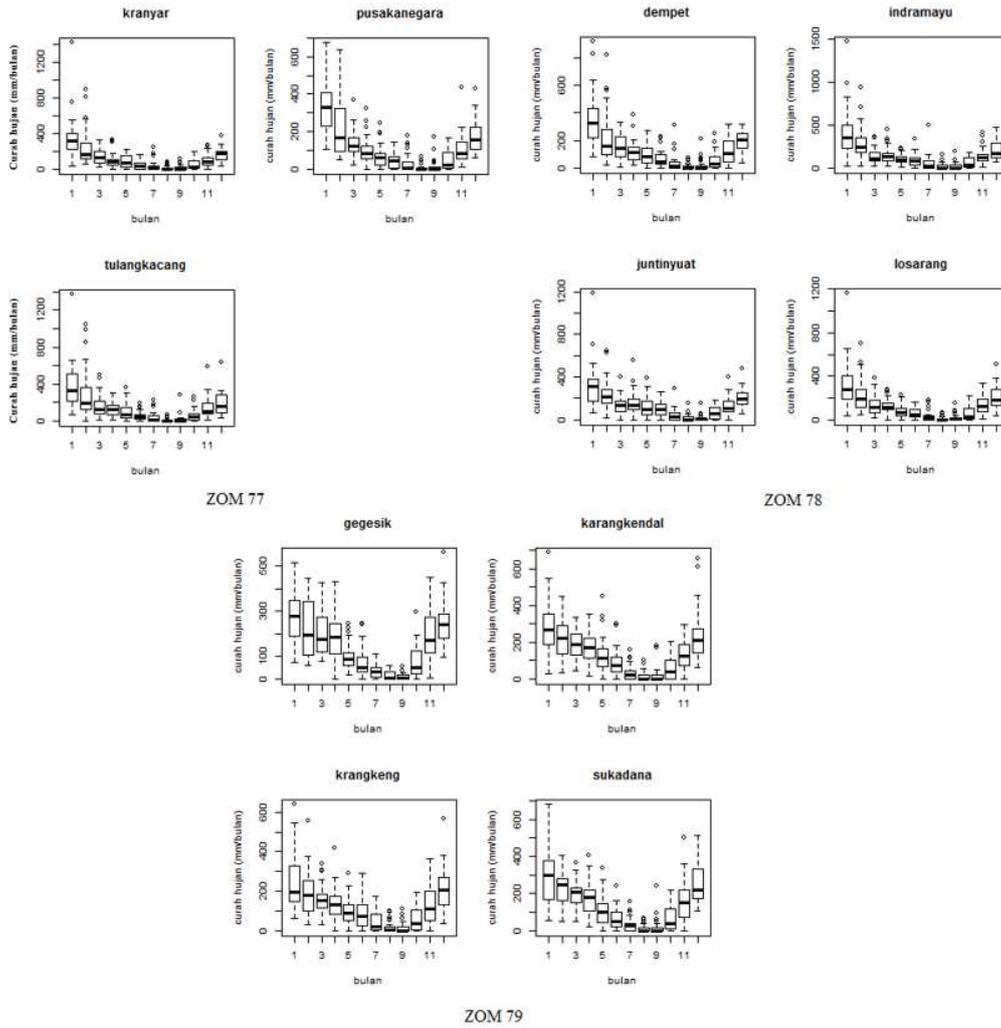
Analisis komponen utama digunakan sebagai teknik prapemrosesan untuk mendapatkan peubah-peubah laten hasil kombinasi linier dari kovariat yang saling ortogonal. Banyaknya peubah laten/komponen utama yang digunakan untuk analisis lanjut ditentukan oleh tiga hal, yaitu grafik screeplot, proporsi keragaman kumulatif, dan nilai ragam yang ditunjukkan oleh nilai akar ciri. Grafik screeplot yang diperoleh dari luaran CMIP5 dan data observasi GPCP disajikan pada Gambar 3. Grafik screeplot untuk kedua kovariat “skala besar” menyarankan untuk mengambil tiga komponen utama sebagai peubah laten, yang diindikasikan oleh grafik batang yang relatif konstan. Hal ini memberikan informasi bahwa keragaman masing-masing komponen utama relatif tidak berbeda nyata.

Proporsi keragaman kumulatif untuk tiga komponen utama dari luaran CMIP5 memberikan nilai di atas 90% (97,30%), tetapi dari data observasi GPCP versi 2,2 untuk mendapatkan proporsi keragaman yang relatif sama ( $> 90\%$ ) diberikan oleh delapan komponen utama sebesar 90,08% (Tabel 1). Secara umum tidak ada metode pasti untuk menetapkan batasan proporsi keragaman yang digunakan (Mattjik dan Sumertajaya, 2011). Dalam kajian ini proporsi di atas 90% digunakan dengan tujuan kedua kovariat “skala besar” dapat diperbandingkan secara seimbang.

Nilai ragam, yang dicirikan oleh nilai akar ciri untuk tiga komponen utama luaran CMIP5 dan delapan komponen utama observasi GPCP 2,2 sudah diperoleh lebih dari 1. Nilai ragam  $> 1$  untuk menentukan banyaknya komponen utama yang diambil dianggap sudah memenuhi asumsi yang digunakan dalam analisis komponen utama (Mattjik dan Sumertajaya, 2011). Oleh karena itu, berdasarkan ketiga kriteria tadi, pemodelan selanjutnya menggunakan tiga peubah laten untuk pendugaan model menggunakan luaran CMIP5 dan delapan peubah laten untuk data observasi GPCP versi 2,2 sebagai kovariat dalam pemodelan SDS pada kajian ini.

### Pendugaan Model Curah Hujan Bulanan

Pendugaan model rata-rata curah hujan bulanan di setiap pos hujan di Kabupaten Indramayu dan sekitarnya menggunakan model linier terampat sebaran gamma. Teknik regularisasi  $L_1$  dan analisis komponen utama digunakan dan diperbandingkan untuk mengatasi data terkondisi buruk dalam pemodelan linier. Sebagai pembanding model linier digunakan pendugaan



Gambar 2. Pola sebaran curah hujan bulanan 11 pos hujan pada wilayah ZOM 77, 78 dan 79

Tabel 1. Proporsi keragaman kumulatif 10 Komponen Utama untuk luaran model GCM CMIP5 dan data observasi GPCP versi 2.2

Komponen	CMIP5		GPCP versi 2,2	
	Akar ciri	Proporsi	Akar ciri	Proporsi
1	13,33	88,14%	16,82	45,87%
2	3,99	96,04%	12,19	69,97%
3	<b>1,60</b>	<b>97,31%</b>	6,44	76,70%
4	1,57	98,53%	5,30	81,24%
5	0,90	98,93%	4,67	84,77%
6	0,85	99,29%	3,71	87,00%
7	0,49	99,41%	3,10	88,56%
8	0,48	99,53%	<b>3,06</b>	<b>90,08%</b>
9	0,37	99,59%	2,78	91,33%
10	0,31	99,64%	2,40	92,26%

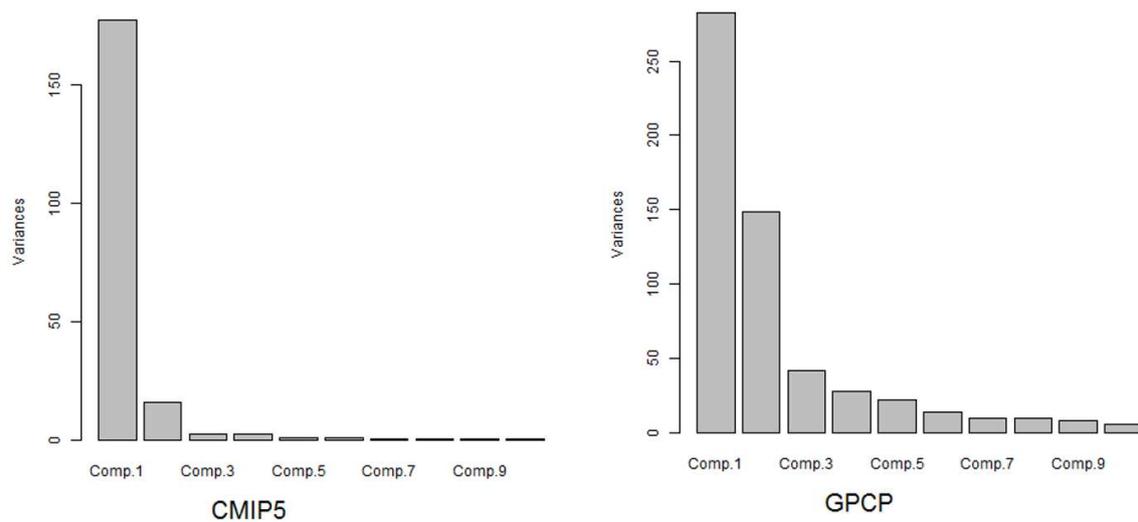
menggunakan metode regresi komponen utama (RKU). Pendugaan model dengan RKU umum digunakan dalam

pemodelan. Dalam kajian ini model curah hujan untuk setiap pos hujan menggunakan analisis komponen utama

dimodelkan sebagai  $CH_i = \beta_0 + \beta_1 KU_1 + \beta_2 KU_2 + \beta_3 KU_3 + \beta_4 D_1 + \dots + \beta_{14} D_{11}$  untuk model dengan kovariat CMIP5, dan  $CH_i = \beta_0 + \beta_1 KU_1 + \beta_2 KU_2 + \dots + \beta_8 KU_8 + \beta_9 D_1 + \dots + \beta_{19} D_{11}$  untuk model dengan kovariat GPCP versi 2,2, sedangkan model dengan teknik regularisasi L1 untuk kedua kovariat adalah  $CH_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_{49} X_{49} + \beta_{50} D_1 + \dots + \beta_{60} D_{11}$ . Dalam hal ini,  $CH_i$

adalah curah hujan bulanan pos hujan ke- $i$ ,  $KU_j$  adalah skor komponen utama ke- $j$ ,  $X_k$  adalah presipitasi pada domain lokasi ke- $k$  dan  $D_l$  adalah peubah *dummy* bulan ke- $l$ .

Nilai statistik RMSE dari dugaan model disajikan pada Tabel 2. Hasil dugaan menunjukkan statistik RMSE untuk model RKU lebih kecil dibanding model linier



Gambar 3. Grafik screeplot hasil analisis komponen utama luaran model CMIP5 dan GPCP

Tabel 2. Nilai statistik RMSE untuk setiap hasil pendugaan model.

Model curah hujan	CMIP5					GPCP versi 2.2				
	RKU	Gamma-KU	RKU-dummy	Gamma-KU-dummy	Gamma-L <sub>1</sub> -dummy	RKU	Gamma-KU	RKU-dummy	Gamma-KU-dummy	Gamma-L <sub>1</sub> -dummy
Kr Anyar	118,52	128,27	115,67	115,48	115,14	121,19	131,57	110,81	108,51	100,32
Pusakanegara	86,78	107,45	83,69	83,72	87,56	95,35	96,81	83,57	83,24	85,76
Tulang										
Kacang	135,56	146,86	132,55	133,20	131,16	137,56	160,09	127,31	138,39	132,53
Dempet	102,48	116,93	100,11	100,17	102,97	108,31	110,15	99,17	98,45	96,77
Indramayu	125,01	135,19	121,75	121,22	123,88	131,07	130,95	120,33	121,01	126,17
Juntinyuat	101,62	112,03	99,55	99,33	100,30	103,83	112,85	97,61	102,00	99,24
Losarang	98,90	113,23	96,06	96,02	95,70	104,71	114,34	95,30	97,92	99,91
Gegesik	90,58	104,07	83,53	83,03	83,83	89,32	102,81	82,21	81,64	86,14
Karangkendal	91,91	106,21	88,84	88,63	92,12	93,28	107,67	88,12	90,35	91,71
Krangken	84,75	93,66	83,01	82,70	82,21	85,62	92,35	81,93	84,65	80,69
Sukadana	89,54	106,26	84,37	84,31	84,14	89,62	104,75	83,39	85,90	84,73

terampat respons sebaran gamma dengan prapemrosesan AKU. Dengan memperhatikan sebaran curah hujan (Gambar 2) pada bulan tertentu (Januari-Desember) yang menunjukkan pola menjulur ke kanan, maka penambahan kovariat peubah *dummy* bulan diduga akan memberikan pendugaan model linier terampat sebaran gamma yang lebih baik.

Penambahan peubah *dummy* bulan berdampak pada penurunan nilai statistik RMSE semua model seperti ditunjukkan pada Tabel 2. Penurunan nilai statistik RMSE yang sangat signifikan terjadi pada model dengan respons sebaran gamma dibanding penurunan pada RKU. Demikian juga pada model dengan kovariat data observasi GPCP versi 2,2 yang memberikan penurunan RMSE lebih signifikan dibanding model dengan kovariat luaran CMIP5.

Teknik regularisasi  $L_1$  yang diterapkan pada model linier terampat sebaran gamma memberikan nilai statistik RMSE yang relatif tidak berbeda nyata dibanding analisis komponen utama. Kedua teknik belum memberikan keunggulan mutlak (mendapatkan RMSE yang kecil) sebagai metode terbaik, sehingga keduanya sama-sama dapat dipertimbangkan dalam pemodelan SDS. Penggunaan model linier terampat sebaran gamma juga memberikan nilai RMSE yang relatif sama dibanding metode RKU yang umum digunakan dalam pemodelan SDS. Keunggulan yang diperoleh dengan menggunakan metode linier terampat sebaran gamma adalah nilai dugaan respons selalu  $\geq 0$ , sesuai dengan karakteristik curah hujan dibanding metode RKU.

## KESIMPULAN

Penambahan peubah *dummy* bulan pada pemodelan SDS menggunakan model linier terampat sebaran gamma, baik dengan teknik regularisasi  $L_1$ /LASSO maupun analisis komponen utama, memberikan nilai RMSE yang sama baiknya dengan metode regresi komponen utama (RKU). Pada kajian ini, teknik LASSO juga memberikan nilai RMSE yang sama baiknya dengan analisis komponen utama (AKU). Keuntungan pendugaan menggunakan sebaran gamma pada pemodelan curah hujan adalah nilai dugaan respons dapat diperoleh dalam keadaan alami, yaitu bernilai non-negatif. Oleh karenanya, untuk pendugaan model curah hujan, metode ini dapat dianggap sebagai model terbaik. Teknik penanganan masalah kovariat menggunakan LASSO atau analisis komponen utama dapat dipertimbangkan sebagai teknik pendugaan pada pemodelan SDS karena sama-sama memberikan nilai RMSE yang tidak berbeda nyata.

## DAFTAR PUSTAKA

- Adler, R.F, G.J Huffman, A. Chang, R. Ferraro, P. Xie, J. Janowiak, B. Rudolf, U. Schneider, S. Curtis, D. Bolvin, A. Gruber, J. Susskind, and P. Arkin. 2003. The Version 2 Global Precipitation Climatology Project (GPCP) Monthly Precipitation Analysis (1979-Present). *J Hydrometeorol*, 4:1147-1167.
- Benestad, R.E, D. Chen, I. Hanssen-Bauer. 2008. *Empirical-Statistical Downscaling*. Singapore: World Scientific Publishing Co. Pte. Ltd.
- Draper, N.R and H. Smith. 1998. *Applied Regression Analysis*. Ed. ke-3. New York (USA): John Wiley & Sons Inc.
- Faraway, J.J. 2006. *Extending the Linear Model with R*. Chapman & Hall/CRC Taylor & Francis Group, London.
- Gao, L., K. Schulz, and M. Bernhardt. 2014. Statistical Downscaling of ERA-Interim Forecast Precipitation Data in Complex Terrain Using LASSO Algorithm. *Advances in Meteorology 2014*: <http://dx.doi.org/10.1155/2014/472741>.
- Hammami, D, T.S Lee, T.B.M.J Ouarda, and Lee J. 2012. Predictor selection for downscaling GCM data with LASSO. *J Geophys Res* 117. doi:10.1029/2012JD017864.
- Haryoko, U. 2015. *Pewilayahan Hujan untuk Menentukan Pola Hujan*. Contoh Kasus Kabupaten Indramayu. Badan Meteorologi Dan Geofisika. <http://www.staklimpondokbetung.net/publikasi/PengelompokanPolaHujan.pdf>. Diakses tanggal 11 Mei 2015.
- Hastie, T., R. Tibshirani, and J. Friedman. 2008. *The Elements of Statistical Learning*. Data mining, Inference, and Prediction. Ed. ke-2 [internet]. [diunduh 2014 Mar 14]; Springer. Tersedia pada: <http://www.stanford.edu/~hastie/pub.htm>.
- Mattjik, A.A dan I.M Sumertajaya. 2011. *Sidik Peubah Ganda*. Dengan Menggunakan SAS. IPB Press, Bogor.
- Sutikno. 2008. *Statistical Downscaling Luaran GCM dan Pemanfaatannya untuk Peramalan Produksi Padi [disertasi]*. Institut Pertanian Bogor, Bogor.
- Sutikno, Setiawan, dan Purnomoadi H. 2010. Statistical Downscaling Output GCM Modeling with Continuum Regression and Pre-Processing PCA Approach. *IPTEK J Tech Sci* 21(3): 109-117.
- Taylor, KE, Stouffer RJ, and Meehl GA. 2012. An Overview of CMIP5 and the Experiment Design. *Bull Amer Meteor Soc* 93: 485-498.
- Tibshirani, R. 1996. Regression Shrinkage and Selection via The LASSO. *J R Statist Soc (B)* 58: 267-288.

- Vimont D.J., D.S Battisti, and R.L Naylor. 2010. Downscaling Indonesian precipitation using large-scale meteorological fields. *Int J Climatol* 30:1706-1722.
- Wigena, A.H. 2006. *Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah Hujan*. Disertasi. Institut Pertanian Bogor, Bogor.