

# SELEKSI PREDIKTOR DATA *GLOBAL CLIMATE MODEL* DENGAN TEKNIK *SINGULAR VALUE DECOMPOSITION* UNTUK PREDIKSI CURAH HUJAN DI PANTAI UTARA JAWA BARAT

## *PREDICTOR SELECTION OF GLOBAL CLIMATE MODEL DATA USING SINGULAR VALUE DECOMPOSITION TECHNIQUE FOR MONTHLY RAINFALL PREDICTION IN NORTHERN COASTAL AREA OF WEST JAVA*

Trinah Wati<sup>1\*</sup>, Aji Hamim Wigena<sup>2</sup>

<sup>1</sup>Pusat Perubahan Iklim dan Kualitas Udara, BMKG, Jl. Angkasa 1 No. 2 Kemayoran Jakarta Pusat

<sup>2</sup>Departemen Statistika, FMIPA, Institut Pertanian Bogor

\*E-mail: trinah.wati@gmail.com

Naskah masuk: 27 Februari 2014; Naskah diperbaiki: 14 Desember 2014; Naskah diterima: 15 Desember 2014

### ABSTRAK

Pemilihan prediktor terbaik untuk curah hujan di 18 pos hujan periode 1980-2005 di Wilayah Pantai Utara di kabupaten Karawang, Indramayu dan Subang dilakukan menggunakan Teknik Singular Value Decomposition (SVD) menggunakan data presipitasi bulanan luaran GPCP dan CMAP, serta data tekanan udara, sea level pressure, precipitable water, suhu udara, kelembaban udara, kecepatan angin luaran NCEP/NCAR reanalisis sebagai input. Prediktor terbaik untuk memprediksi rata-rata curah hujan bulanan di 18 stasiun hujan di wilayah pantai utara adalah kecepatan angin luaran NCEP/NCAR reanalisis. Selanjutnya prediktor terbaik digunakan untuk memprediksi curah hujan bulanan tahun 2006-2007 menggunakan pemodelan Statistical Downscaling Multiple Linear Regression (MLR) dan Partial Least Square Regression (PLS). Hasil validasi curah hujan prediksi dengan aktual menggunakan PLS memiliki korelasi yang lebih tinggi dan RMSEP yang lebih kecil dibandingkan MLR.

**Kata kunci:** statistical downscaling, GCM, singular value decomposition, multiple linear regression, partial least square regression.

### ABSTRACT

*Selection for best predictor of 18 rain gauge stations from 1980 to 2005 in Northern Coastal Area in Karawang, Indramayu and Subang District has been investigated with The Singular Value Decomposition technique using monthly rainfall data from GPCP and CMAP, and air pressure, sea level pressure, precipitable water, relative humidity temperature, wind speed from NCEP/NCAR reanalysis as the input. The best predictor for monthly rainfall prediction of 18 rain gauges is wind speed from NCEP/NCAR reanalysis. Further, this best predictor was used to predict monthly rainfall 2006 to 2007 using Multiple Linear Regression (MLR) Statistical Downscaling model and Partial Least Square (PLS) Regression. These methods showed that the validation results using PLS have higher correlation and smaller RMSEP than MLR method.*

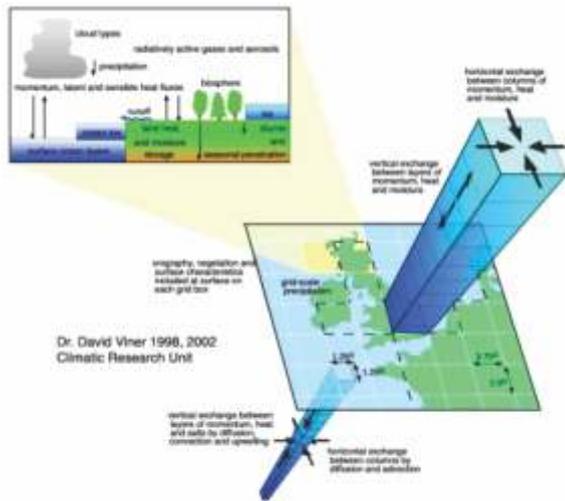
**Keywords:** statistical downscaling, GCM, singular value decomposition, multiple linear regression, partial least square regression.

## 1. Pendahuluan

*General Circulation Model* (GCM) merupakan data model numerik yang menggambarkan proses fisik atmosfer, lautan, kriosfir dan permukaan daratan sebagai alat paling terdepan dalam mensimulasi pengaruh sistem iklim global terhadap peningkatan gas rumah kaca akibat perubahan iklim [1] (Gambar 1). GCM menggambarkan iklim dalam grid tiga dimensi bola bumi dengan resolusi horizontal antara 250 dan 600 km, dan antara 10 hingga 20 lapisan vertikal di atmosfer dan 30 lapisan di dalam lautan. Skala resolusi tersebut cukup rendah sehingga

diperlukan suatu teknik untuk pengamatan peubah iklim secara lokal.

Secara umum, ada dua pendekatan yang dapat dilakukan untuk memanfaatkan data GCM menjadi informasi skala lokal yaitu *Dynamical Downscaling* dan *Statistical Downscaling* [2]. *Statistical Downscaling* (SD) merupakan suatu teknik yang menggunakan model statistika untuk melihat hubungan antara suatu data yang berskala besar dengan data yang berskala lokal. Pada SD pola hubungan antara kedua peubah dapat diasumsikan linier dan nonlinier [3].



**Gambar 1. Ilustrasi data GCM**

Metode-metode yang dapat digunakan untuk pola hubungan linier, seperti Regresi Linier Peubah Ganda, Analisis Korelasi Kanonik, Regresi Komponen Utama, dan *Singular Value Decomposition* (SVD). Metode-metode nonlinier umumnya menggunakan Jaringan Saraf Tiruan dan Metode Analog [4]. Beberapa permasalahan dalam SD adalah penentuan domain, penentuan prediktor, dan mendapatkan model yang tepat sesuai dengan karakteristik data sehingga bisa menggambarkan hubungan antara gugus data GCM dan data stasiun cuaca. Banyaknya variabel X dan variabel Y dapat meningkatkan kompleksitas model yang memungkinkan terjadinya kasus kolinearitas antar variabel X dan antarvariabel Y serta kasus autokorelasi.

Penggunaan teknik SVD merupakan pra pemrosesan statistical downscaling dalam upaya mereduksi dimensi dari GCM. Prediktor yang paling dominan yang kemudian akan digunakan untuk memprediksi curah hujan. Teknik statistical downscaling regresi berganda yang digunakan adalah Regresi Linier Peubah Ganda (*Multiple Linier regression / MLR*) dan *Partial Least Square Regression (PLS)*. Proses validasi model dilakukan dengan menghitung koefisien korelasi dan RMSEP antara data prediksi dan data observasi.

Pemanfaatan data GCM untuk memprediksi curah hujan di Indonesia sangat diperlukan untuk meningkatkan ketepatan prediksi curah hujan bulanan. Penelitian SD sebelumnya yang telah dilakukan menggunakan metode Regresi Projection Pursuit [5], SVD [6] [7], PLS [8], Pricpal Component Regression dan PLS [9] dengan studi kasus wilayah Indramayu.

Dalam penelitian ini bertujuan untuk menentukan prediktor terbaik dari data GCM menggunakan metode SVD dan pemodelan SD regresi berganda MLR dibandingkan dengan metode PLS antara GCM

dengan curah hujan bulanan di wilayah jalur pantura yang meliputi kabupaten Subang, Karawang dan Indramayu propinsi Jawa Barat.

## 1. Metode Penelitian

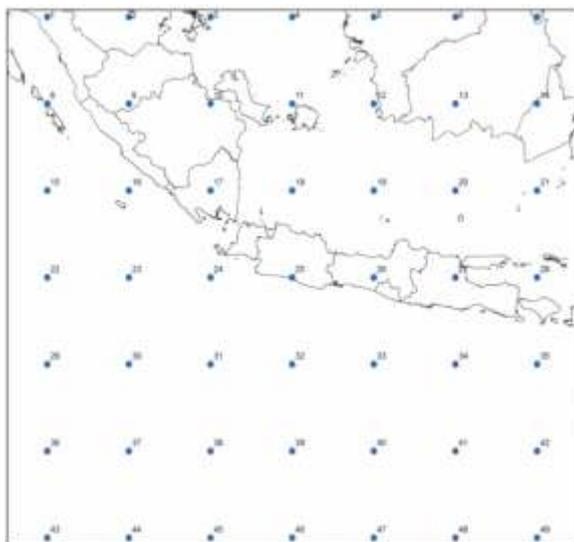
**Data dan Lokasi.** Data yang digunakan dalam penelitian ini adalah data curah hujan bulanan di wilayah jalur pantura Jawa Barat meliputi kabupaten Indramayu, Subang dan Karawang periode 1980-2007 diperoleh dari BMKG (Badan Meteorologi Klimatologi dan Geofisika) yang terdiri dari 18 pos hujan, yaitu : Krangkeng, Cidempet, Juntinyuat, Sukadana, Indramayu dan Losarang (Indramayu), Karanganyar, Salam Darma, Pamanukan, Barugbug, Ciasem dan Pusakanegara (Subang) dan Pasirukem, Petaruman, Rawamerta, Batujaya, Pedes dan Teluk Buyung (Karawang) (Gambar 2).

Data GCM yang digunakan yaitu presipitasi CMAP (*CPC Merged Analysis of Precipitation*) (sumber <http://www.esrl.noaa.gov/psd/data/gridded/data.cmap.html>) , presipitasi GPCP (*Global Precipitation Climatological Project*) (sumber <http://www.esrl.noaa.gov/psd/data/gridded/data.gpcp.html>) dan unsur iklim seperti suhu udara, kecepatan angin, tekanan udara, *sea level pressure* dan *precipitable water* dari *NCEP/NCAR Reanalysis derived data* (sumber <http://www.esrl.noaa.gov/psd/data/gridded/data.ncep.reanalysis.derived.html>) dengan skala  $2.5^{\circ} \times 2.5^{\circ}$  atau  $300 \text{ km} \times 300 \text{ km}$  dengan periode yang sama. Pengolahan data untuk pemodelan menggunakan data 1980 sampai 2005, sedangkan untuk validasi model data tahun 2006 dan 2007.

Ukuran domain untuk downscaling adalah sejumlah grid berukuran  $8 \times 8$  untuk data presipitasi CMAP dan GPCP sedangkan NCEP/NCAR  $7 \times 7$  pada  $1.4^{\circ} \text{ LU} - 18.1^{\circ} \text{ LS}$  dan  $98.4^{\circ} - 118.1^{\circ} \text{ BT}$ , yaitu di atas wilayah kajian. Grid NCEP/NCAR  $7 \times 7$  dapat dilihat pada Gambar 3.



**Gambar 2. Lokasi stasiun hujan wilayah pantura Jawa Barat**



Gambar 3. Domain grid GCM NCEP/NCAR

Tahapan dalam penelitian ini adalah menentukan prediktor dari data GCM menggunakan metode SVD sebagai pra pemrosesan dan kemudian pemodelan SD dilakukan menggunakan metode MLR dan PLS.

**Metode Singular Value Decomposition.** Penerapan SVD pada klimatologi digunakan untuk mengetahui hubungan linier dua gugus data, seperti data GCM dan data curah hujan stasiun cuaca. Metode ini menggunakan matriks koragam silang dari dua gugus data. Misal, terdapat dua buah matriks X dan Y, masing-masing berukuran  $t \times p$  dan  $t \times q$  [10]: (persamaan 1)

$$\mathbf{X} = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{t1} & \dots & X_{tp} \end{pmatrix}$$

$$\mathbf{Y} = \begin{pmatrix} Y_{11} & \dots & Y_{1q} \\ \vdots & \ddots & \vdots \\ Y_{t1} & \dots & Y_{tq} \end{pmatrix} \quad (1)$$

$X_{t \times p}$  adalah nilai prediktor pada waktu ke-t di grid ke-p.  $Y_{t \times q}$  adalah nilai curah hujan pada waktu ke-t di stasiun ke-q. Matriks koragam silang ( $C_{xy}$ ) dapat dituliskan dengan rumus persamaan (2) [10]:

$$C_{xy} = \mathbf{Z}^T \mathbf{S} \quad (2)$$

dengan:

$\mathbf{Z}^T$  = matriks transpose dari matriks Z.

$$\mathbf{Z} = \begin{pmatrix} X_{11} & \dots & X_{1p} \\ \vdots & \ddots & \vdots \\ X_{t1} & \dots & X_{tp} \end{pmatrix} - \begin{pmatrix} \bar{X}_1 & \dots & \bar{X}_p \\ \vdots & \ddots & \vdots \\ \bar{X}_1 & \dots & \bar{X}_p \end{pmatrix}$$

$$\mathbf{S} = \begin{pmatrix} Y_{11} & \dots & Y_{1q} \\ \vdots & \ddots & \vdots \\ Y_{t1} & \dots & Y_{tq} \end{pmatrix} - \begin{pmatrix} \bar{Y}_1 & \dots & \bar{Y}_q \\ \vdots & \ddots & \vdots \\ \bar{Y}_1 & \dots & \bar{Y}_q \end{pmatrix} \quad (3)$$

X = rata-rata kolom ke-i matriks X untuk  $i=1, 2, 3, \dots, p$

Y = rata-rata kolom ke-j matriks Y untuk  $j=1, 2, 3, \dots, q$

Tujuan dari analisis SVD adalah untuk mencari kombinasi linier p peubah prediktor yang mempunyai nilai koragam maksimum dengan kombinasi linier q peubah respon.

$$\mathbf{A} = u_1 x_1 + u_2 x_2 + \dots + u_p x_p = \mathbf{xu}$$

$$\mathbf{B} = v_1 y_1 + v_2 y_2 + \dots + v_q y_q = \mathbf{Yv} \quad (4)$$

Pasangan kombinasi linier A dan B disebut *expansion coefficient*. *Expansion coefficient* analog dengan skor komponen pada analisis komponen utama. Dengan SVD, matriks koragam silang akan diuraikan menjadi dua gugus *spatial pattern* yang orthogonal dan berpasangan dengan nilai singular. *Spatial pattern* analog dengan vektor ciri pada analisis komponen utama. Matriks  $C_{xy}$  diuraikan dengan SVD menjadi persamaan (5) [10]:

$$C_{xy} = \mathbf{UDV}^T \quad (5)$$

dengan:

U = matriks singular berukuran  $p \times m$  dari matriks  $C_{xy}$   
V = matriks singular berukuran  $m \times q$  dari matriks  $C_{xy}$   
 $m = \min(p, q)$

Matriks D merupakan matriks diagonal tak negatif yang unsur-unsurnya  $\lambda^1 > \lambda^2 > \dots > \lambda^m$ . Matriks U dan V masing-masing secara berurutan dapat ditentukan dengan mencari vektor ciri dari  $C_{xy} C_{xy}^T$  dan  $C_{xy}^T C_{xy}$  yaitu vektor ciri  $u_1, u_2, u_3, \dots, u_p$  dan  $v_1, v_2, v_3, \dots, v_q$  yang berpadanan dengan akar ciri  $\lambda_1^2 > \lambda_2^2 > \dots > \lambda_m^2$ . Pasangan *spatial pattern* ke-i merupakan pasangan dari vektor pada kolom matriks U ke-i dan vektor pada kolom matriks V ke-i.

Keeratan hubungan pasangan *spatial pattern* ke-i ditunjukkan oleh koefisien korelasi Pearson dari pasangan *expansion coefficient* ke-i. Tiap pasang *spatial pattern* ini menggambarkan *square covariance fraction* (SCF) dari gugus data X dan Y. SCF ke-i dapat dianalogkan dengan proporsi keragaman peubah X dan Y yang diterangkan oleh pasangan *spatial pattern* ke-i. SCF ke-i, ditunjukkan dengan persamaan (6):

$$SCF_i = \frac{\lambda_i^2}{\sum_{i=1}^m \lambda_i^2} \quad (6)$$

Ada dua cara untuk menampilkan pola spasial agar lebih mudah dipahami, yaitu dengan membuat peta korelasi homogen dan heterogen [10]. Peta korelasi homogen ke-i adalah peta yang menggambarkan koefisien korelasi antara *expansion coefficient* A ke-i dengan X atau korelasi antara *expansion coefficient* B ke-i dengan Y, sedangkan peta korelasi heterogen ke-i adalah peta yang menggambarkan koefisien korelasi antara *expansion coefficient* A ke-i dengan Y atau

korelasi antara expansion coefficient B ke-i dengan X. Peta korelasi homogen berguna sebagai indikator area geografis yang mencakup bagian dari stasiun stasiun cuaca, sedangkan peta korelasi heterogen menunjukkan seberapa baik stasiun diprediksi oleh *expansion coefficient*.

**Metode Multi Linear Regression.** MLR digunakan sebagai proses pemodelan SD untuk memprediksi curah hujan bulanan, MLR adalah analisis regresi yang terdiri dari dua atau lebih peubah respon dengan peubah bebas yang sama. Misal, terdapat q peubah respon dan p peubah bebas akan dibuat q model liniertunggal dengan persamaan sebagai berikut:

$$\begin{aligned} y_1 &= \beta_{01}1_t + \beta_{11}x_1 + \beta_{21}x_2 + \dots + \beta_{p1}x_p + \varepsilon_1 \\ y_2 &= \beta_{02}1_t + \beta_{12}x_1 + \beta_{22}x_2 + \dots + \beta_{p2}x_p + \varepsilon_2 \\ &\vdots \\ y_q &= \beta_{0q}1_t + \beta_{1q}x_1 + \beta_{2q}x_2 + \dots + \beta_{pq}x_p + \varepsilon_q \end{aligned} \quad (7)$$

Persamaan (7) dapat dinotasikan dengan matriks sebagai berikut [11]:

$$Y = X\beta + \varepsilon \quad (8)$$

dimana:

$$\begin{aligned} Y_{t \times q} &= \begin{pmatrix} Y_{t1} & Y_{t2} & \dots & Y_{tq} \\ \vdots & \vdots & \ddots & \vdots \\ Y_{t1} & Y_{t2} & \dots & Y_{tq} \end{pmatrix} \\ X_{t \times p} &= \begin{pmatrix} 1 & X_{t1} & X_{t2} & \dots & X_{tp} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{t1} & X_{t2} & \dots & X_{tp} \end{pmatrix} \\ \beta &= \begin{pmatrix} \beta_{01} & \beta_{02} & \dots & \beta_{0q} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{p1} & \beta_{p2} & \dots & \beta_{pq} \end{pmatrix} \\ \varepsilon_{t \times q} &= \begin{pmatrix} \varepsilon_{t1} & \varepsilon_{t2} & \dots & \varepsilon_{tq} \\ \vdots & \vdots & \ddots & \vdots \\ \varepsilon_{t1} & \varepsilon_{t2} & \dots & \varepsilon_{tq} \end{pmatrix} \end{aligned} \quad (9)$$

Dengan pendekatan kuadrat terkecil, dapat diperoleh matriks penduga bagi  $\beta$  pada persamaan (10):

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (10)$$

Pada penelitian ini peubah bebas yang digunakan adalah *expansion coefficient* Adan data pos hujan sebagai peubah respon [12]. Prediksi curah hujan dihitung dengan persamaan (11):

$$\begin{aligned} \hat{\beta} &= (A^T A)^{-1} A^T Y \\ \hat{Y} &= A \hat{\beta} \end{aligned} \quad (11)$$

dimana:

$\hat{Y}$  = prediksi curah hujan

$$A = \begin{pmatrix} 1 & A_{11} & A_{12} & \dots & A_{1p} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & A_{t1} & A_{t2} & \dots & A_{tp} \end{pmatrix}$$

t = panjang tahun validasi

p = banyaknya *expansion coefficient* yang digunakan

**Metode Partial Least Square Regression.** *Partial least square* (PLS) atau Regresi Kuadrat Terkecil Parsial digunakan untuk melakukan ekstraksi atas sejumlah komponen yang disebut sebagai peubah laten. Peubah laten merupakan peubah prediktor (X) yang dipilih sebagai sejumlah komponen yang relevan dengan sejumlah peubah respon (Y) melalui proses dekomposisi peubah X dan peubah Y secara simultan dengan batasan bahwa komponen-komponen tersebut dapat menjelaskan sebanyak mungkin keragaman antara peubah prediktor X dan peubah respon Y. Dalam metode PLS telah mencakup teknik *Principle Component Analysis* (PCA) yang bertujuan mereduksi dimensi dan autokolerasi dan Regresi Berganda.

Model PLS akan mendapatkan sejumlah komponen baru yang akan memodelkan X dan Y sedemikian sehingga X dan Y memiliki hubungan. Komponen baru tersebut selanjutnya disebut skor X, yang dicatat dengan  $t_a$ , dengan  $a = 1, 2, 3, \dots, A$ .

Skor  $X(t_a)$  merupakan kombinasi linier peubah-peubah asal  $X_k$  dengan koefisien yang disebut sebagai pembobot yang dicatat sebagai  $w_{ka}$  ( $a = 1, 2, 3, \dots, A$ ), dengan formulasi sebagai berikut: [13]

$$\begin{cases} t_{ia} = \sum_k X_{ik} W_{ka} & i = 1, 2, 3, \dots, N \\ T = XW \end{cases} \quad (12)$$

Skor X digunakan prediktor untuk X dan juga Y, sebagai berikut:

a. Skor X sebagai prediktor bagi X:

$$\begin{cases} t_{ia} = \sum_k X_{ik} W_{ka} & i = 1, 2, 3, \dots, N \\ T = XW \end{cases} \quad (13)$$

b. Skor Y sebagai prediktor bagi Y:

$$\begin{cases} Y_{im} = \sum_a t_{ia} r_{am} + f_{im} \\ X = TR' + F \end{cases} \quad (14)$$

Residu Y yakni  $f_{im}$  menyatakan deviasi antara respon pengamatan dengan respon dugaan. Berdasarkan persamaan (12) dan (14) maka model regresi berganda dapat ditulis sebagai berikut:

$$\begin{cases} Y_{im} = \sum_a r_{ma} \sum_k W_{ka} x_{ik} + f_{im} \\ = \sum_k b_{mk} x_{ik} + f_{im} \\ Y = XWR' + F = XB + F \end{cases} \quad (15)$$

Koefisien model PLS,  $b_{mk}(B)$  dapat ditulis sebagai berikut:

$$\begin{cases} b_{mk} = \sum_a r_{ma} W_{ka} \\ T = WR' \end{cases} \quad (16)$$

Prediksi bagi data pengamatan yang baru dapat diperoleh berdasarkan pada data X dan matriks koefisien B.

**Validasi Model.** Validasi model dilakukan dengan menghitung RMSEP dan korelasi 18 pos hujan menggunakan data curah hujan tahun 2006 dan 2007 untuk masing-masing predictor. Dalam suatu pemodelan nilai RMSEP diharapkan kecil. RMSEP merupakan salah satu ukuran keakuratan suatu model. Nilai RMSEP dapat dituliskan dengan persamaan (17):

$$RMSEP = \sqrt{\sum_{i=1}^{n_p} (y_i - \hat{y}_i)^2 / n_p} \quad (17)$$

dimana:  $y_i$  = data aktual,  $\hat{y}_i$  = nilai dugaan,  $n_p$  = banyaknya data untuk validasi model. Performa model yang baik ditunjukkan dengan nilai RMSEP yang lebih kecil dan nilai korelasi yang lebih tinggi.

Hubungan linear antara dua peubah kuantitatif dapat dilihat dengan menggunakan analisis korelasi Pearson. Koefisien korelasi ini mengukur keeratan antara dua peubah yang bernilai -1 sampai dengan 1, yang diduga dengan koefisien korelasi contoh r, yaitu:

$$r_{xy} = \frac{S_{xy}}{\sqrt{S_{xx}} \sqrt{S_{yy}}} \quad (18)$$

Dengan :

- $r_{xy}$  = koefisien korelasi antara peubah X dan peubah Y
- $S_{xy}$  = koragam peubah X dan peubah Y
- $S_{xx}$  = ragam peubah X
- $S_{yy}$  = ragam peubah Y

Uji nyata hasil korelasi dilakukan dengan uji t hitung dengan rumus:

$$t = \frac{r_{xy} \sqrt{n-k}}{\sqrt{1-r^2}} \quad (19)$$

Dengan :

- $r_{xy}$  = koefisien korelasi antara peubah X dan peubah Y
- n = banyaknya data
- k = banyaknya peubah bebas

Hasil uji t hitung dibandingkan dengan t tabel pada tingkat keyakinan 99% ( $\alpha = 0.01$ ), dengan derajat bebas (db) = n-k-1.

### 3. Hasil Dan Pembahasan

**Penentuan Prediktor Terbaik.** Grid GCM sebagai domain prediktor menggunakan domain prediktor yang berlokasi di atas stasiun Indramayu[5] (Gambar 2) karena lokasi yang berdekatan dengan wilayah jalur pantura di Jawa Barat yaitu kabupaten

Karawang, Indramayu dan Subang dan masih tercakup dalam domain tersebut. Daftar prediktor yang digunakan dalam penelitian dapat dilihat pada tabel 1. Presipitasi dalam tabel 1 merupakan segala bentuk hasil proses kondensasi uap air di atmosfer yang jatuh akibat gravitasi (curah hujan, salju, drizzle, sleet, es, graupel).

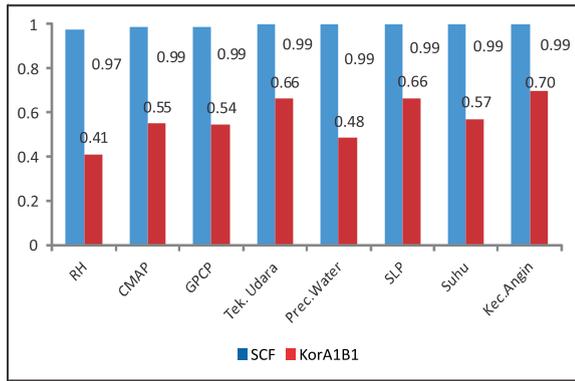
Nilai korelasi pasangan antara *expansion coefficient* pertama dan SCF terlihat pada Gambar 4, kisaran korelasi 0.41 hingga 0.7. Prediktor yang memiliki nilai korelasi *expansion coefficient* terbesar adalah kecepatan angin sebesar 0.7 sedangkan yang terkecil adalah kelembaban udara dengan nilai 0.41. Penelitian sebelumnya [6, 7] menunjukkan prediktor terbaik wilayah kajian Indramayu adalah data GCM Presipitasi CMAP dan angin zonal U-wind yang memiliki nilai korelasi pasangan *expansion coefficient* 0.78 dan 0.77.

**Tabel 1. Prediktor yang digunakan dalam SD**

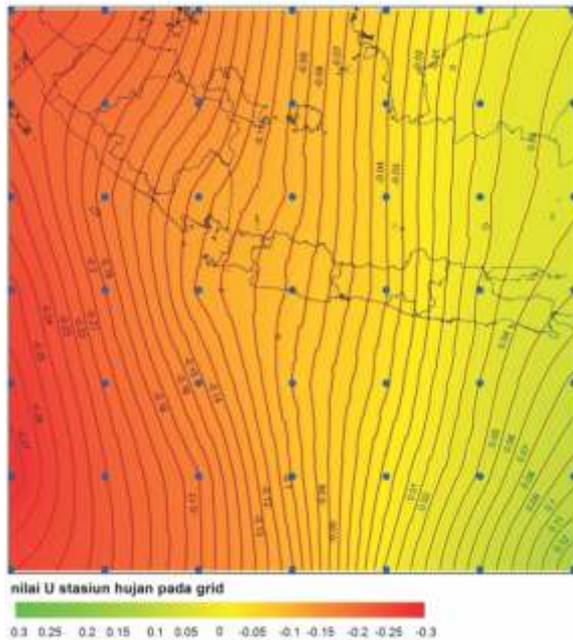
Prediktor	Keterangan
Presipitasi	GPCP (mm/hari)
Presipitasi	CMAP (mm/hari)
Tekanan udara permukaan	NCEP-NCAR (mbar)
Sea Level Pressure	NCEP-NCAR (mbar)
Kelembaban Udara	NCEP-NCAR (%)
Suhu Udara	NCEP-NCAR (C)
Kecepatan Angin	NCEP-NCAR (m/s)
Precipitable Water	NCEP-NCAR (kg/m <sup>2</sup> )

Setiap prediktor dalam penelitian ini memiliki nilai SCF lebih dari 0.97 menunjukkan lebih dari 97% keragaman data GCM dan stasiun hujan telah dijelaskan dengan baik oleh pasangan *expansion coefficient*. Besarnya SCF pertama menunjukkan bahwa pasangan *expansion coefficient* pertama mampu menjelaskan keragaman data rata-rata curah hujan bulanan pos hujan dan GCM secara baik.

Nilai mutlak *spasial patern U* pertama ( $U_1$ ) kecepatan angin menunjukkan besarnya kontribusi keragaman grid kecepatan angin terhadap A1. Gambar 5 menunjukkan lokasi grid kecepatan angin terhadap A1, kisaran nilai  $U_1$  antara -0.277 hingga 0.1311. Nilai positif dan negatif hanya menjelaskan hubungan saja. Warna merah pada Gambar 5 menunjukkan semakin besar keragaman yang diberikan lokasi (grid) tersebut terhadap model (meskipun negatif) berada di lokasi samudera Hindia.



Gambar 4. Korelasi pasangan *expansion coefficient* pertama dan nilai SCF pertama pada setiap prediktor



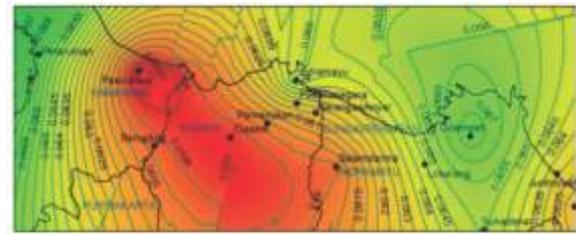
Gambar 5. Peta kontur *spatial pattern* U pertama ( $U_1$ ) kecepatan angin NCEP/NCAR

Nilai mutlak *spatial pattern* V pertama ( $V_1$ ) kecepatan angin menunjukkan kontribusi keragaman grid terhadap B1. Gambar 6 menunjukkan lokasi grid kecepatan angin terhadap B1 kisaran nilai  $V_1$  antara 0.1885 hingga 0.3154, pos hujan teluk buyung, Karawang memberikan kontribusi keragaman terhadap B1.

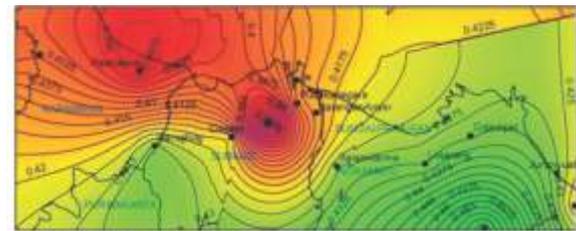
Penampilan pola spasial menggunakan peta korelasi heterogen yang digunakan untuk mengetahui hubungan antara lokasi grid atau stasiun dengan *expansion coefficient*.

Gambar 7 menunjukkan bahwa 5 stasiun pertama yang berkorelasi kuat dengan A1 kecepatan angin luaran NCEP/NCAR yaitu stasiun Sukadana, ciasem, pusakanegara, teluk buyung dan salamdarma. Stasiun yang memiliki korelasi terbesar dengan A1 adalah Sukadana sebesar 0.68.

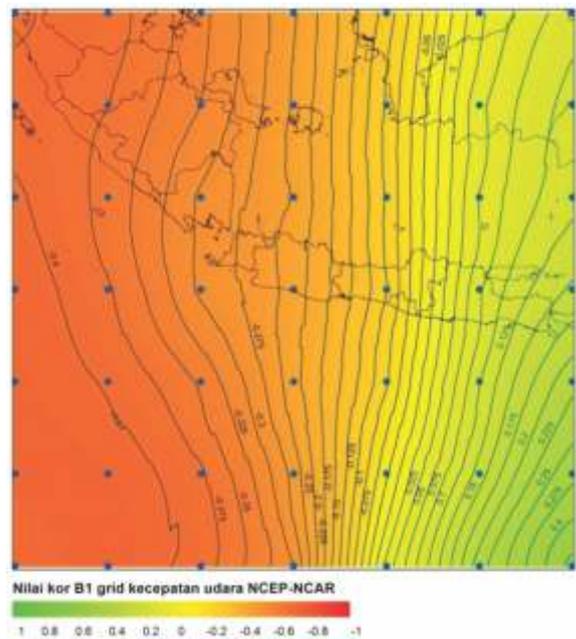
Gambar 8 menunjukkan bahwa 5 grid kecepatan angin luaran NCEP/NCAR pertama yang berkorelasi kuat dengan B1 yaitu grid X44, X36, X37, X29, dan X22. Stasiun yang memiliki korelasi terbesar dengan B1 adalah grid X44 sebesar -0.69.



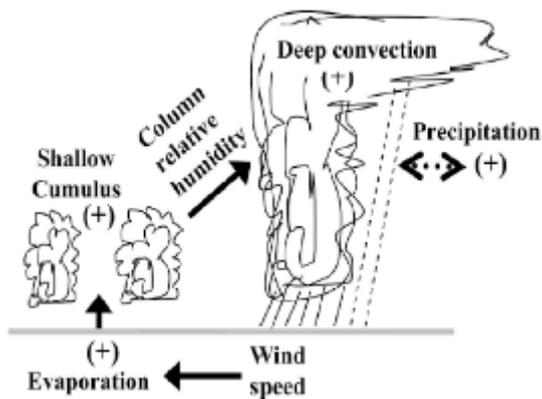
Gambar 6. Peta kontur *spatial pattern* V pertama ( $V_1$ ) kecepatan angin NCEP/NCAR



Gambar 7. Peta korelasi heterogen antara data stasiun dengan A1 data kecepatan angin luaran NCEP/NCAR



Gambar 8. Peta korelasi heterogen antara grid kecepatan angin luaran NCEP/NCAR dengan B1



**Gambar 9.** Skema inisiasi awan konvektif, evaporasi meningkatkan populasi awan cumulus dan kolom kelembapan udara meningkatkan proses konveksi awan yang menyebabkan terjadinya presipitasi [15]

Hubungan angin dengan curah hujan untuk wilayah tropis berkaitan dengan semakin meningkatnya kecepatan angin akan mengakibatkan peningkatan proses evaporasi dan fluks panas laten dari lautan ke atmosfer. Di wilayah tropikal konvektif, peristiwa konveksi awan mengikuti pergerakan kecepatan angin dimana dalam suatu wilayah pada periode peningkatan kecepatan rata-rata angin menyebabkan curah hujan semakin meningkat karena peningkatan fluks panas laten permukaan [14].

Hubungan antara kecepatan angin harian dengan presipitasi menggunakan data *Special Sensor Microwave Imager (SSM/I)* dan *Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI)* [15], secara fisis menunjukkan kecepatan angin yang semakin tinggi menyebabkan peningkatan evaporasi sehingga terjadi ketidak stabilan lapisan perbatas (*Boundary Layer*) dan menstimulasi kejadian konveksi awan yang semakin kuat. Peningkatan presipitasi yang diamati terlihat lebih besar dibandingkan dengan perubahan evaporasi terkait kecepatan angin akibat terjadinya feedback konvergensi (Gambar 9).

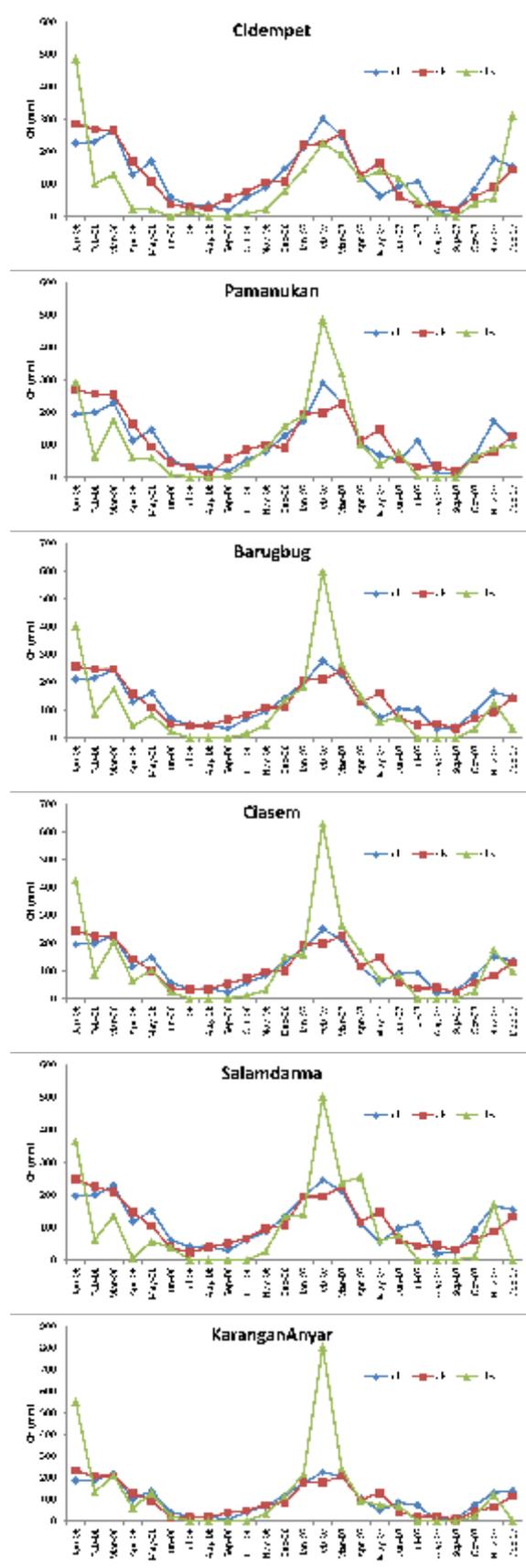
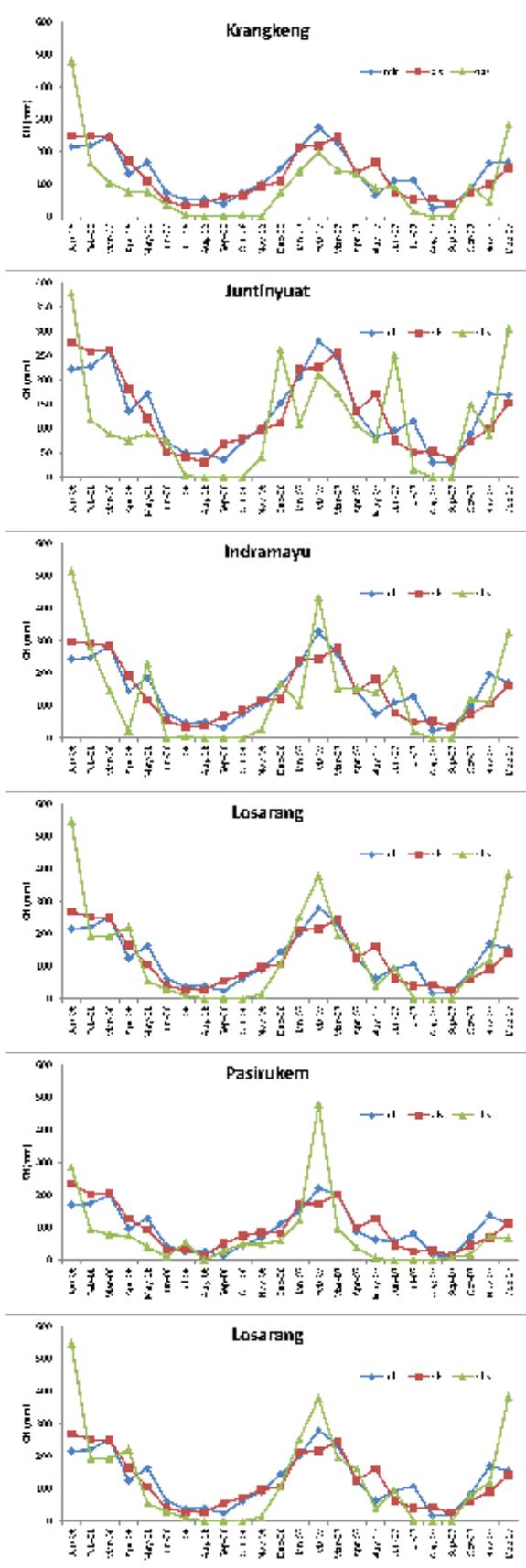
**Hasil Prediksi Curah Hujan Bulanan Tahun 2006 dan 2007.** Prediktor terbaik berdasarkan metode SVD yaitu kecepatan angin luaran NCEP/NCAR

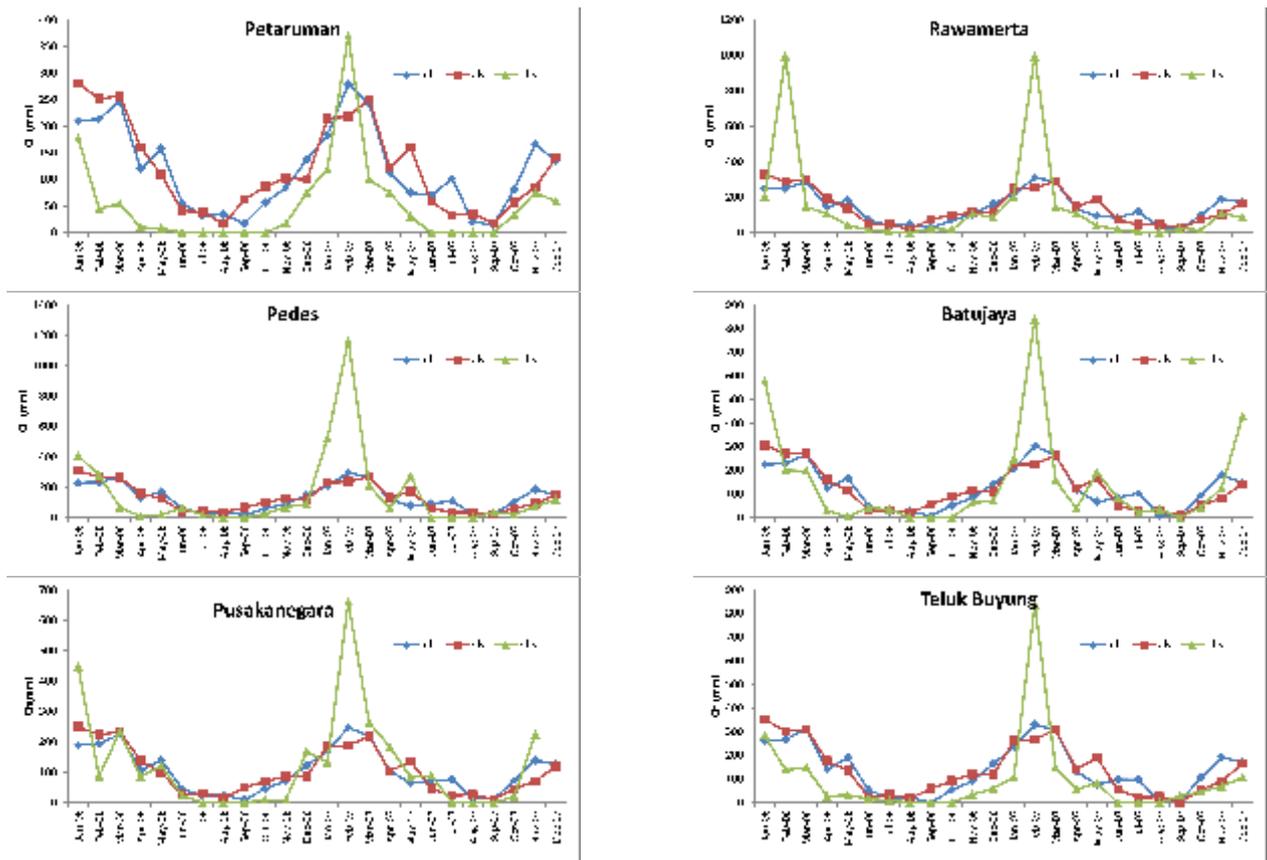
kemudian digunakan dalam pemodelan SD untuk memprediksi curah hujan bulanan 1 tahun dan 2 tahun kedepan yaitu tahun 2006 dan tahun 2006 - 2007 menggunakan MLR dan PLS. Ukuran kebaikan prediktor dalam memprediksi curah hujan dapat dilihat melalui koefisien korelasi dan RMSEP serta uji t (two-tails) antara data prediksi dan data observasi berdasarkan data validasi 1 tahun 2006 dan 2 tahun 2006–2007 (Tabel 2).

Gambar 10 memperlihatkan grafik antara hasil pemodelan menggunakan MLR dan PLS pada tahun 2006 dan 2007, secara umum pola hasil prediksi memiliki pola yang sama dengan observasi namun *underestimate* pada saat puncak musim penghujan yaitu pada bulan Januari dan Februari kecuali di pos hujan Krangkeng dan Cidempet. Sedangkan pada puncak musim kemarau yaitu bulan Juli, Agustus dan September curah hujan hasil prediksi *overestimate* dibandingkan dengan observasi karena nilainya selalu di atas nol. Hasil prediksi curah hujan menggunakan GCM kurang dapat menggambarkan nilai-nilai ekstrim pada musim penghujan.

Berdasarkan tabel 2 terlihat bahwa korelasi yang dihasilkan oleh metode PLS umumnya lebih besar dibandingkan dengan MLR untuk validasi tahun 2006 kecuali pos hujan Juntinyuat, Indramayu dan Pamanukan. Sedangkan nilai RMSEP, metode PLS umumnya memiliki nilai lebih kecil dibandingkan MLR kecuali pos hujan Juntinyuat, Indramayu, Pamanukan dan Teluk Buyung. Hasil uji t hitung dibandingkan dengan t tabel 2.576 (2 tails, db:166,  $\alpha = 0.01$ ) menunjukkan t hitung > t tabel sehingga korelasi signifikan.

Sedangkan prediksi untuk 2 tahun validasi tahun 2006-2007 menunjukkan metode PLS memiliki korelasi lebih tinggi dan RMSEP lebih kecil dibandingkan dengan regresi linear peubah ganda. Sehingga metode PLS terlihat dapat memprediksi curah hujan lebih baik dibandingkan metode MLR dan lebih konsisten untuk memprediksi periode curah hujan yang lebih panjang. Hasil uji t hitung dibandingkan dengan t tabel 2.576 (2 tails, db:166,  $\alpha = 0.01$ ) menunjukkan t hitung > t tabel sehingga korelasi signifikan.





Gambar 10. Grafik curah hujan hasil prediksi dengan observasi

Tabel 2. Koefisien Korelasi dan RMSEP validasi 1 tahun dan 2 tahun

pos hujan	validasi tahun 2006						validasi tahun 2006-2007					
	metode PLS			Metode MLR			metode PLS			Metode MLR		
	korelasi	RMSEP	T hitung	korelasi	RMSEP	T hitung	korelasi	RMSEP	T hitung	korelasi	RMSEP	T hitung
krangkeng	0.74	95.36	14.23	0.67	105.11	11.74	0.71	117.64	13.12	0.60	135.35	9.63
cidempet	0.72	106.05	13.35	0.62	113.37	10.26	0.68	129.55	11.98	0.57	145.96	9.04
Juntinyuat	0.62	96.70	10.11	0.64	92.41	10.70	0.58	128.61	9.22	0.51	130.63	7.70
sukadana	0.75	96.89	14.70	0.74	101.13	14.36	0.74	132.77	14.35	0.67	166.67	11.62
indramayu	0.76	105.86	14.89	0.78	104.91	16.15	0.72	139.68	13.35	0.61	161.60	10.04
losarang	0.84	94.30	19.77	0.72	112.06	13.38	0.82	126.03	18.68	0.67	157.72	11.73
KaranganAnyar	0.80	100.18	17.36	0.74	109.36	14.02	0.75	201.87	14.48	0.60	220.23	9.61
salamdarma	0.72	82.99	13.42	0.67	86.56	11.79	0.71	134.54	13.05	0.58	148.71	9.25
pamanukan	0.75	74.82	14.47	0.76	62.03	14.92	0.79	105.88	16.60	0.69	120.80	12.45
barugbug	0.76	84.06	14.90	0.74	84.25	14.03	0.79	135.80	16.86	0.64	158.70	10.67
ciasem	0.78	79.23	16.15	0.76	81.91	15.08	0.80	141.57	17.08	0.64	162.12	10.65
pusakanegara	0.79	82.15	16.50	0.77	85.43	15.82	0.79	152.65	16.79	0.63	173.63	10.38
Pasirukem	0.78	57.69	16.32	0.61	63.68	9.83	0.73	109.19	13.72	0.55	142.69	8.49
petaruman	0.73	111.90	13.63	0.66	99.43	11.50	0.71	135.00	13.14	0.58	147.97	9.32
Rawamerta	0.60	217.97	9.68	0.56	225.85	8.85	0.62	230.51	10.26	0.52	243.26	7.92
Batujaya	0.80	103.61	17.52	0.67	120.72	11.58	0.72	212.38	13.23	0.59	238.34	9.32
Pedes	0.78	90.96	16.08	0.64	99.54	10.74	0.63	294.57	10.57	0.59	302.34	9.54
TelukBuyung	0.90	98.06	26.80	0.81	91.61	17.85	0.65	195.57	11.17	0.54	209.04	8.28

#### 4. Kesimpulan

Pemilihan predictor dengan metode SVD dapat digunakan untuk menentukan prediktor terbaik data GCM yang berskala global. Kecepatan angin luaran NCEP/NCAR merupakan prediktor terbaik untuk memprediksi rata-rata curah hujan bulanan di wilayah pantura Jawa Barat. Hasil validasi curah hujan prediksi model dibandingkan dengan curah hujan aktual memiliki korelasi yang nyata pada tingkat kepercayaan 99%. Pola curah hujan hasil

prediksi memiliki pola yang sama dengan hujan observasi namun *underestimate* di puncak musim hujan dan *overestimate* di puncak musim kemarau. Korelasi curah hujan hasil model dengan aktual menggunakan metode PLS umumnya memiliki korelasi yang lebih besar dan RMSEP yang lebih kecil dibandingkan dengan MLR untuk prediksi 1 tahun dan 2 tahun kedepan, sehingga metode PLS lebih konsisten untuk memprediksi data untuk periode yang lebih panjang.

## Daftar Pustaka

- [1] What is a GCM? (2013). ([http://www.ipcc-data.org/guidelines/pages/gcm\\_guide.html](http://www.ipcc-data.org/guidelines/pages/gcm_guide.html)), diakses tanggal 19 Januari 2014
- [2] Benestad RE. (2004). Empirical-Statistical Downscaling in Climate Modeling. *EOS*, 85(42), 417.
- [3] Zorita E & H von Storch. (1999). The Analog Method as a Simple Statistical Downscaling Technique. *Journal of Climate and Applied Meteorology*, 12, 2474-2489.
- [4] Bretherton SC, Smith C, & Wallace JM. (1992). An Intercomparison of Methods for Finding Coupled Pattern in Climate Data. *Journal of Climate and Applied Meteorology*, 5, 541-560.
- [5] Wigena AH. (2006). *Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah Hujan Bulanan: Kasus curah Hujan Bulanan di Indramayu*. Disertasi, Program Pascasarjana, Institut Pertanian Bogor.
- [6] Sanjaya, Imam. (2010). Penentuan Prediktor pada Statistical Downscaling dengan Singular Value Decomposition (Studi Kasus di Indramayu). Departemen Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Pertanian Bogor.
- [7] Gunawan, D. & Linarka, A.U. (2011). Penentuan Prediktor untuk Prediksi Curah Hujan Bulanan menggunakan Metode Statistical Dynamical Downscaling. *Jurnal Meteorologi dan Geofisika*, 12(1), 93 – 102.
- [8] Swarinoto YS. & Wigena AH. (2011). Statistical Downscaling Suhu Muka Laut Global untuk Prediksi Total Hujan Bulanan Menggunakan Teknik PLSR. *Jurnal Meteorologi dan Geofisika*, 12(1), 9-19
- [9] Estiningtyas W & Wigena AH. (2011). Teknik Statistical Downscaling dengan Regresi Komponen Utama dan Regresi Kuadrat Terkecil Parsial untuk Prediksi Curah Hujan pada kondisi El Niño, La Niña dan Normal. *Jurnal Meteorologi dan Geofisika*, 12(1), 65 – 72.
- [10] Bjornsson H & Venegas SA. (1997). *A Manual for EOF and SVD Analyses of Climatic Data*. Canada: McGill University.
- [11] Kattathee R & Naik DN. (1999). *Applied Multivariate Statistics with SAS® Software*, Second Edition. NC: SAS Institute Inc-Cary.
- [12] Uvo BC, et al. (2001). Statistical Atmospheric Downscaling for Rainfall Estimation in Kyushu Island, Japan. *Hydrology and Earth System Sciences*, 5(2), 259-271.
- [13] Wang J, Zhenzhen K, & Liang J. (2002). *Facial Feature Point Extraction by partial least square regression*. (<http://citeseer.ist.psu.edu/wang02facial.html>), diakses 22 Januari 2014
- [14] Raymond, (2003). Convective Forcing in the Intertropical Convergence Zone of The East Pacific. *J. Atmos. Sci*, 60, 2064-2082
- Back, LE & Bretherton, CS. (2005). The Relationship between Wind Speed and Precipitation in the Pacific ITCZ. *Journal of Climate*, 18, 4317-4328.