

# Modeling Statistical Downscaling for Prediction Precipitation Dry Season in Bireuen District Province Aceh

<sup>1</sup>Juniana Husna, <sup>2</sup>Sanusi

<sup>1,2</sup>Departement of Information System, Faculty of Engineering, Abulyatama University

Email: <sup>1</sup>Juniana@abulyatama.ac.id, <sup>2</sup>sanusi@abulyatama.ac.id

## Article Info

### Article history:

Received Aug 04<sup>th</sup>, 2019

Revised Aug 15<sup>th</sup>, 2019

Accepted Sep 14<sup>th</sup>, 2019

### Keyword:

General Circulation Model

Prediction

Rainfall Prediction

Statistical Downscaling

Support Vector Regression

## ABSTRACT

The Asian-Australian monsoon circulation specifically causes the Indonesian region to go through climate changeability that impacts on rainfall variability in different Indonesia's zone. Local climate conditions such as rainfall data are commonly simulated using General Circulation Model (GCM). This study tries to model the statistical downscaling of GCM in the form of 7x7 matrix using Support Vector Regression (SVR) for rainfall forecasting during drought in Bireuen Regency, Aceh. The output yields optimal result using certain parameter i.e.  $C = 0.5$ ,  $\gamma = 0.8$ ,  $d = 1$ , and  $\xi = 0.01$ . The duration of computation during training and testing are  $\pm 45$  seconds for linear kernels and  $\pm 2$  minutes for polynomials. The correlation degree and RMSE values of GCM and the actually observed data at Gandapura weather station are 0.672 and 21.106. The RSME value obtained in that region is the lowest compared to the Juli station which is equal to 31,428. However, the Juli station has the highest correlation value that is 0.677. On the other hand, the polynomial kernel has a correlation degree and RMSE value equal to 0.577 and 29,895 respectively. To summary, the best GCM using SVR kernel is the one at Gandapura weather station in consideration of having the lowest RMSE value with a high correlation degree.

Copyright © 2019 Puzzle Research Data Technology

## Corresponding Author:

Juniana Husna

Departement of Information System

Faculty of Engineering

Abulyatama University

Jl. Blang bintang Lama KM. 8.5, Lampoh Keude, Aceh Besar, Aceh Indonesia

Email: Juniana@abulyatama.ac.id

DOI: <http://dx.doi.org/10.24014/ijaidm.v2i2.7518>

## 1. PENDAHULUAN

Pengaruh fenomena regional seperti sirkulasi monsun Asia-Australia merupakan daerah pertemuan angin antar tropis yang merupakan daerah pertumbuhan awan, serta terjadi perubahan suhu permukaan laut di sekitar wilayah Indonesia [1][2]. Berbagai fenomena tersebut secara spesifik menyebabkan wilayah Indonesia terjadinya keragaman iklim, pengaruh dari faktor keragaman iklim tersebut dapat dijadikan sebagai indikator dalam kaitannya terhadap tanaman adalah curah hujan bagi sektor pertanian khususnya padi di Provinsi Aceh.

Pada tahun 2018 sebanyak 451 hektar sawah di Provnsi Aceh yang berisi tanaman padi dilanda kekeringan, delapan hektar diantaranya termasuk kategori berat, 39 hektar termasuk kategori sedang, dan 403 hektar termasuk kategori ringan, dengan fenomena seperti ini bahkan satu hektar diantaranya sudah puso atau gagal panen. Daerah yang kerap sekali terjadinya kekeringan yaitu Aceh Besar 179 hektar, Gayo Lues 12 hektar dan sementara ini daerah paling luas terkena bencana kekeringan yaitu Kabupaten Bireuen seluas 260 hektar, Fenomena kekeringan tersebut dapat mempengaruhi penurunan hasil produksi tanaman pangan khususnya padi di Kabupaten Bireuen Provinsi Aceh [3].

Dari berbagai macam data model *time series* dapat digunakan untuk menduga curah hujan dimusim kemarau, salah satunya adalah luaran *General Circulation Model* (GCM) sebagai variabel peubah bebas dan dapat digunakan untuk menghasilkan kondisi iklim lokal seperti curah hujan dan temperatur [4]. GCM merupakan model yang dirancang untuk mensimulasikan *time series* tertentu dari variabel iklim secara global

dan bisa digunakan untuk prediksi iklim dan cuaca [5][6]. GCM simulasi model berbasis komputer yang terdiri dari berbagai persamaan numerik dan deterministik yang terpadu dan mengikuti kaidah-kaidah fisika yang dapat mengkaji keragaman iklim baik di masa lampau, sekarang maupun di masa yang akan datang, resolusi model GCM terletak pada kisaran 2.50 atau 300 km<sup>2</sup> pada setiap lapisan di atmosfer [7]. Oleh karena itu untuk menjembatani gap antara hasil global dengan regional maupun lokal dalam menurunkan skala spasial GCM maka perlu diselesaikan dengan teknik *downscaling*.

Teknik *downscaling* adalah suatu teknik untuk memperoleh informasi skala lokal dari skala global luaran GCM atau hubungan yang dapat dikembangkan antara *grid* skala global menjadi *grid* skala lokal [8], [9][10]. Ada dua pendekatan *downscaling* yang sering digunakan yaitu *Dynamic Downscaling* dan *Statistical Downscaling*. *Dynamic Downscaling* mencakup model iklim secara regional yang bersarang ke dalam sebuah model iklim global [11]. Metode *Dynamic Downscaling* paling kompleks karena membutuhkan persamaan dinamika atmosfer yang memerlukan perangkat keras dengan kecepatan tinggi dan kapasitas memori yang besar serta membutuhkan waktu komputasi yang cukup lama [12]. Sedangkan bila dilihat berdasarkan letak wilayah Indonesia yang beriklim tropis dengan ketersediaan data yang terbatas dengan wilayah yang cukup luas dan banyak pulau-pulau, untuk menurunkan *grid* ke skala lokal dalam penelitian ini sangat cocok bila dipadukan dengan metode *Statistical Downscaling*.

Usaha pelaksanaan pengelolaan dampak keragaman iklim menggunakan *Statistical Downscaling* (SD) yang telah dilakukan oleh peneliti sebelumnya seperti Wigena menggunakan pemodelan SD dengan *Regression Projection Pursuit* (RPP) untuk peramalan curah hujan [13], penelitian tersebut melakukan prediksi curah hujan bulanan di wilayah Indramayu Jawa Barat menggunakan skala *grid* 8x8. Hasil yang diperoleh dengan domain berukuran 8x8 *grid* pendugaannya lebih stabil dan tidak terlalu sensitif terhadap data pencilan. Selanjutnya Agmaloro melakukan pemodelan SD data GCM menggunakan *Support Vector Regression* (SVR) untuk memprediksi curah hujan bulanan Indramayu [14]. Penelitian tersebut menggunakan domain *grid* berukuran 5x5 untuk prediksi curah hujan bulanan di wilayah Indramayu Jawa Barat. Hasil optimasi parameter dengan *grid search* kernel RBF menunjukkan *performance* paling baik diantara kernel lainnya. Dari kedua penelitian tersebut dapat disimpulkan bahwa pemanfaatan luaran GCM hasilnya mampu memprediksi curah hujan observasi dan sangat cocok bila digunakan untuk mengembangkan model prediksi, sehingga dapat memberikan informasi tentang iklim yang memadai. Oleh karena itu penelitian ini mencoba memodelkan *Statistical Downscaling* menggunakan *Support Vector Regression* (SVR) dengan memanfaatkan data global luaran GCM berupa matriks 7x7 untuk memprediksi curah hujan musim kemarau di Kabupaten Bireuen Provinsi Aceh. Pemilihan algoritma SVR dalam penelitian ini karena SVR merupakan salah satu algoritma yang digunakan untuk kasus prediksi atau pengambilan keputusan. Algoritma ini pengembangan dari regresi linear. Jika regresi linear hanya membentuk atau memperoleh sebuah fungsi regresi dengan hasil yang linear (berbentuk garis lurus), maka SVR dapat menghasilkan sebuah fungsi polinomial yang mengikuti sebaran data yang terbentuk sehingga hasil prediksi lebih akurat.

## 2. DATA DAN METODE

Data yang digunakan pada penelitian ini terdiri dari data curah hujan observasi pada 7 stasiun hujan di Kabupaten Bireuen Provinsi Aceh, diperoleh dari Badan Meteorologi Klimatologi dan Geofisika (BMKG) Indrapuri Aceh Besar Provinsi Aceh. Data *hindcast* GCM sebanyak 5 model dan diakses melalui *website* CLIK APCC ([klik.apcc21.org](http://klik.apcc21.org)). Adapun luaran GCM beserta negara pembuatnya dapat dilihat pada Tabel 1.

Tabel 1. Luaran GCM dan Negara Pembuat

No	Nama luaran GCM	Negara	Bulan	Tahun
1	NASA-GSFC L34	USA	Mei-Juni-Juli	1990-2013
2	METRI AGCM L17	Korea	Mei-Juni-Juli	1990-2013
3	PNU	Korea	Mei-Juni-Juli	1990-2013
4	MGO	Columbia	Mei-Juni-Juli	1990-2013
5	BCC	USA	Mei-Juni-Juli	1990-2013

Tabel 2. Nama dan lokasi Stasiun pengamatan curah hujan di Kabupaten Bireuen

Y	Nama Stasiun	Latitude	Longitude	Bulan	Tahun
Y1	Juli	5.073	96.530	Mei-Juni-Juli	1990-2013
Y2	Peudada	5.051	96.425	Mei-Juni-Juli	1990-2013
Y3	Peulimbang	5.047	96.366	Mei-Juni-Juli	1990-2013
Y4	Gandapura	5.226	96.846	Mei-Juni-Juli	1990-2013
Y5	Jeumpa	5.145	96.583	Mei-Juni-Juli	1990-2013
Y6	Jeunieb	5.064	96.338	Mei-Juni-Juli	1990-2013
Y7	Peusangan	5.196	96.738	Mei-Juni-Juli	1990-2013

## 2.1. Praproses

Data luaran GCM di ambil berdasarkan bujur dan lintang pada setiap stasiun pengamatan untuk prediksi curah hujan di musim kemarau. *Grid* pada model GCM menjelaskan struktur matriks dengan titik pusatnya merupakan lokasi titik koordinat observasi. Penelitian ini menggunakan *grid* 7x7, maka setiap bulannya ada satu ukuran *grid* 7x7 dan setiap tahunnya ada 3 *grid* 7x7 yaitu pada bulan mei-juni-juli. Selanjutnya dilakukan praproses data yaitu dengan mengkonversi ukuran matriks *grid* 7x7 menjadi sebuah vector baris berukuran 1x(7x7) agar dapat digunakan sebagai masukkan proses regresi menggunakan SVR. Selanjutnya didapatkan untuk 1 bulan vektor *grid* GCM 1x49 dan matriks *grid* 3x49 dalam setahun. Sedangkan keseluruhan observasi menjadi 72x49. 72 menyatakan observasi sebanyak 3 bulan selama 24 tahun, sedangkan 49 merupakan vektor ciri yang dikonversi dari *grid* 7x7 hasil *downscaling* luaran GCM.

Data luaran GCM memiliki ukuran dimensi data yang tinggi disebabkan GCM adalah suatu model berbasis komputer yang terdiri atas berbagai persamaan numerik dan deterministik dan mampu menduga perubahan unsur-unsur cuaca dalam bentuk luaran *grid-grid* yang berukuran km menurut lintang dan bujur. Untuk mengurangi masalah ukuran dan terjadi autokorelasi pada data luaran GCM yang berdimensi besar, maka dilakukan reduksi dimensi spasial dengan menggunakan *Principal Component Analysis* sebesar 95%.

Sebelum dilakukan penyusunan model SVR, terlebih dahulu dilakukan pembagian data latih dan data uji menggunakan teknik *k-fold cross validation*, penelitian ini ditentukan nilai *k* sebesar 8. Selanjutnya tahap pelatihan dan pengujian menggunakan 2 fungsi kernel SVR yaitu kernel linear dan polinomial.

## 2.2. Support Vector Regression (SVR)

SVR merupakan metode yang dapat mengatasi *overfitting*, sehingga akan menghasilkan performansi yang lebih bagus [15][18]. Model yang dihasilkan *Support Vector Regression* tergantung pada *subset* dari data pelatihan, karena fungsi biaya untuk membangun sebuah model dan mengabaikan data pelatihan yang dekat dalam kasus prediksi.

Misalkan kita mempunyai sebanyak  $\ell$  set data *training* yaitu:  $\{\chi = [x_i, y_i]$  dimana  $i = 1, \dots, \ell\}$ , dengan data *input*  $x = \{x_1, x_2, x_3, \dots, n\} \subseteq \mathfrak{R}^N$  dan *output* yang bersangkutan yaitu  $\{y = [y_1, \dots, y_\ell] \subseteq \mathfrak{R}\}$ . Manakala nilai  $\varepsilon$  sama dengan 0 maka kita dapatkan suatu regresi yang sempurna. Misalkan kita mempunyai fungsi sebagai garis regresi berikut [16]:

$$f(x) = w^T \cdot \phi(x) + b \quad (1)$$

Dimana  $\phi(x)$  menunjukkan suatu titik di dalam *feature space* hasil pemetaan  $x$  di dalam *input space*. Koefesien  $w$  dan  $b$  diestimasi dengan cara meminimalkan *risk function* yang didefinisikan dalam persamaan berikut:

$$\min \frac{1}{2} \|w\|^2 + C \frac{1}{\ell} \sum_{i=1}^{\ell} L_{\varepsilon}(y_i, f(x_i)) \quad (2)$$

tergantung

$$\begin{aligned} y_i - w\phi(x_i) - b &\leq \varepsilon \\ w\phi(x_i) + b - y_i &\leq \varepsilon, i = 1, 2, 3, \dots, \ell \end{aligned} \quad (3)$$

dimana

$$L_{\varepsilon}(y_i, f(x_i)) = \begin{cases} |y_i - f(x_i)| - \varepsilon, & |y_i - f(x_i)| \geq \varepsilon \\ 0, & \text{untuk yang lain} \end{cases} \quad (4)$$

Faktor  $\|w\|^2$  dinamakan *regularisasi*. meminimalkan  $\|w\|^2$  akan membuat fungsi setipis mungkin, sehingga bisa mengontrol kapasitas fungsi. Faktor kedua adalah kesalahan empirik yang diukur dengan  $\varepsilon$ -*insensitive loss function* yang diharuskan meminimalkan *norm* dari  $w$  agar mendapatkan *generalisasi* yang baik untuk fungsi regresi  $f(x)$ . Karena itu kita perlu meyelesaikan masalah optimasi berikut:

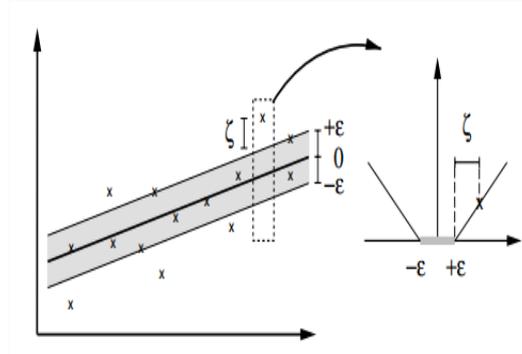
$$\min \frac{1}{2} \|w\|^2 \quad (5)$$

tergantung

$$y_i - w\phi(x_i) - b \leq \varepsilon \quad (6)$$

$$w\phi(x_i) + b - y_i \leq \varepsilon, i = 1, 2, 3, \dots, \ell \tag{7}$$

Kita asumsikan bahwa ada sebuah fungsi  $f(x)$  yang dapat mengaproksimasi semua titik  $(x_i, y_i)$ . Maka dengan SVR kita akan mendapatkan suatu tabung seperti pada Gambar 1.



Gambar 1.  $\varepsilon$ -insensitive loss function.

Presisi  $\varepsilon$  dalam kasus ini kita asumsikan bahwa semua titik ada dalam rentang  $f \pm \varepsilon$  (*feasible*). Dalam hal ketidaklayakan, dimana ada beberapa titik yang mungkin keluar dari rentang  $f \pm \varepsilon$ , maka perlu kita tambahkan variabel *slack*  $\xi, \xi^*$  untuk mengatasi masalah pembatas yang tidak layak dalam masalah optimasi. Selanjutnya masalah optimasi tersebut bisa menggunakan rumus berikut:

$$\min = \frac{1}{2} \|w\|^2 + C \sum_{\ell=1}^{\ell} (\xi_i + \xi_i^*) \tag{8}$$

tergantung

$$y_i - w^T \phi(x_i) - \xi_i - b \leq \varepsilon, i = 1, 2, 3, \dots, \ell \tag{9}$$

$$w\phi(x_i) - y_i - \xi_i^* + b \leq \varepsilon, i = 1, 2, 3, \dots, \ell \text{ dimana } \xi, \xi^* \geq 0 \tag{10}$$

Konstanta  $C > 0$  menentukan tawar-menawar antara ketipisan fungsi  $f$  dan batas atas deviasi lebih dari  $\varepsilon$  masih ditoleransi. Semua deviasi lebih besar dari  $\varepsilon$  akan dikenakan pinalti sebesar  $C$ . Dalam SVR,  $\varepsilon$  sebanding dengan akurasi dari aproksimasi terhadap data *training*. Nilai  $\varepsilon$  yang kecil terkait dengan nilai yang tinggi pada variabel *slack*  $\xi_i^*$  dan akurasi aproksimasi yang tinggi. Sebaliknya, nilai yang tinggi untuk  $\varepsilon$  berkaitan dengan nilai  $\xi_i^*$  yang kecil dan akurasi aproksimasi yang rendah. Nilai yang tinggi untuk variabel *slack* akan membuat kesalahan empirik mempunyai pengaruh yang besar terhadap faktor *regularisasi*. Dalam SVR *support vector* adalah data *training* yang terletak diluar batas  $f$  dari fungsi keputusan, karena itu jumlah *support vector* menurun dengan naiknya nilai  $\varepsilon$ . Dalam formulasi dual, masalah optimasi dari SVR adalah sebagai berikut [16]:

$$\max \sum_{\ell=1}^{\ell} y_i (\alpha_i - \alpha_i^*) - \varepsilon \sum_{\ell=1}^{\ell} (\alpha_i + \alpha_i^*) - \frac{1}{2} \sum_{i=1}^{\ell} \sum_{j=1}^{\ell} (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) \tag{11}$$

tergantung

$$\sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) = 0 \text{ dimana } 0 \leq \alpha_i, \alpha_i^* \leq C, i = 1, 2, 3, \dots, \ell \tag{12}$$

Dimana  $C$  ditentukan sendiri oleh pengguna,  $K(x_i, x_j)$  adalah *dot-product* kernel yang didefinisikan sebagai  $K(x_i, x_j) = \phi^T(x_i) \phi^T(x_j)$ . Dengan menggunakan *Lagrange multipliers* dan kondisi optimalitas, fungsi regresi secara eksplisit dirumuskan sebagai berikut:

$$f(x) = \sum_{i=1}^{\ell} (\alpha_i - \alpha_i^*) K(x, x_i) + b \tag{13}$$

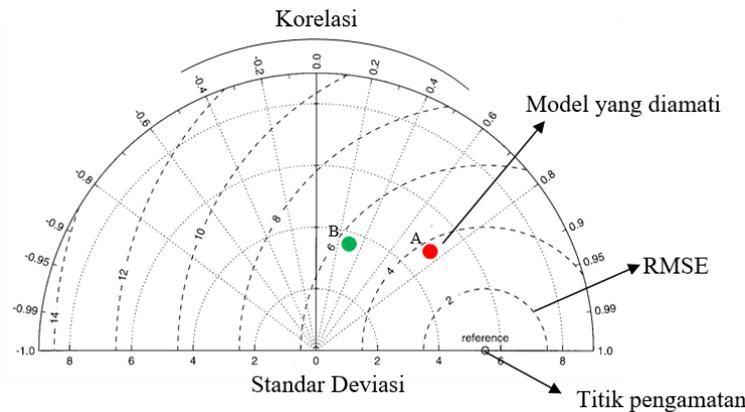
Ada 2 fungsi kernel SVR yang digunakan dalam penelitian ini sebagai berikut:

1. Persamaan fungsi kernel linear adalah:  $k(x, y) = x^T y + C$  (14)
2. Persamaan fungsi kernel polynomial adalah:  $k(x, y) = (ax^T y + C)^d$  (15)

Sebelum pelatihan dan pengujian SVR ditentukan terlebih dahulu nilai parameter  $C$ ,  $\gamma$  untuk fungsi kernel linear, nilai parameter  $C$ ,  $\gamma$ ,  $d$  dan  $\varepsilon$  untuk kernel polinomial.

### 1.3. Analisis dan Evaluasi

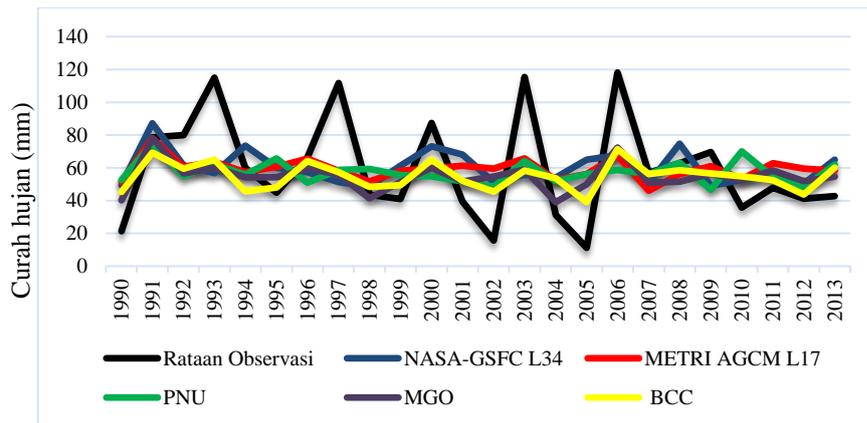
Diagram *Taylor* sangat baik untuk melakukan evaluasi banyak aspek dari model yang kompleks ataupun menilai kehandalan dari beberapa model sekaligus [17]. Diagram *Taylor* dibangun dari *Root Mean Square Error* (RMSE) *Standart Deviation* (STD), dan *Correlation* ( $r$ ) antara prediksi terhadap observasi. Diagram *Taylor* secara lengkap dapat dilihat pada Gambar 2.



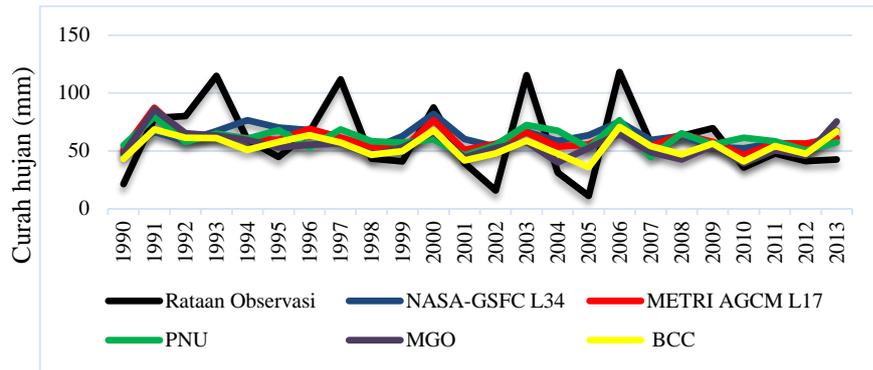
Gambar 2. Diagram Taylor

## 2. HASIL DAN PEMBAHASAN

Penelitian ini melakukan percobaan terhadap data luaran GCM dengan menggunakan SVR. Berdasarkan hasil estimasi yang telah dilakukan terhadap keseluruhan data luaran GCM maupun data observasi yang digunakan pada 7 pos stasiun hujan di Kabupaten Bireuen diperoleh hasil bahwa model estimasi menghasilkan nilai yang kurang baik pada fungsi kernel linear dan polinomial dengan melakukan perbandingan dari nilai rata-rata hasil estimasi, diagram garis pada semua model GCM terhadap rata-rata curah hujan observasi, dapat dilihat pada Gambar 3 dan Gambar 4.



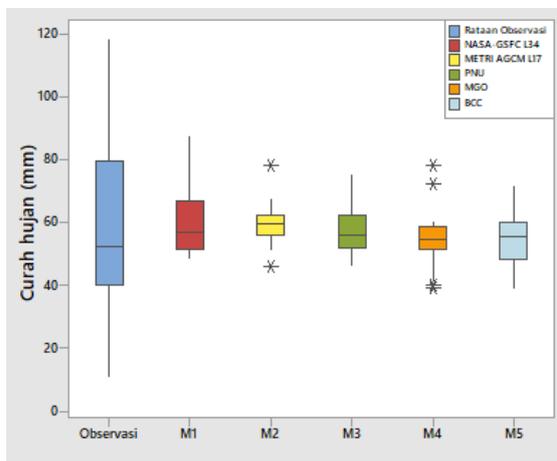
Gambar 3. Rataan hasil prediksi pada kernel linear



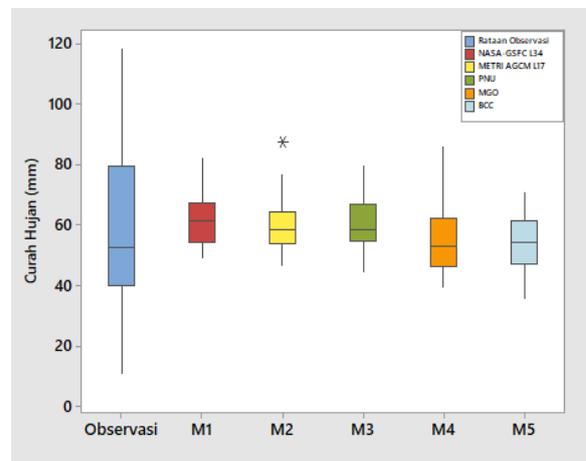
Gambar 4. Rataan hasil prediksi pada kernel polinomial

Pada Gambar 3 dan Gambar 4 dapat diketahui estimasi pada kernel liner dan polinomial model luaran GCM hasilnya belum sepenuhnya mengikuti pola data curah hujan observasi, karena model GCM cenderung hasil prediksinya lebih rendah dibandingkan dengan data curah hujan observasi. Adapun pada titik tertentu curah hujan observasi terdapat beberapa jumlah curah hujan yang ekstrim cenderung maksimum, seperti berkisar antara tahun 1992-1994, antara tahun 1996-1998, antara tahun 2002-2004, dan tahun 2006. dan beberapa titik ekstrim lainnya yang berada pada titik minimum dari curah hujan observasi, seperti tahun 2002 dan 2005. Oleh karena itu, pendugaan pada titik ekstrim tersebut secara rata-rata model luaran GCM belum menghasilkan nilai prediksi yang optimal, karena hasil prediksi yang dihasilkan oleh model belum sensitif menangkap pola curah hujan yang ekstrim tersebut, maka dalam hal ini secara rata-rata model luaran GCM yang diestimasi dengan fungsi kernel linear dan polinomial belum mampu sepenuhnya dapat menjelaskan data curah hujan observasi.

Analisis *boxplot* pada penelitian ini menggunakan *software* Minitab V.17, untuk mengetahui ukuran pemusatan data curah hujan observasi dan penyebaran data luaran GCM dari nilai tengahnya, curah hujan ekstrim atau pencilan dan beberapa pengukuran lainnya dapat diamati secara langsung, baik pada kernel linear maupun kernel polynomial dapat dilihat pada Gambar 5 dan Gambar 6.



Gambar 5. Kernel Linear



Gambar 6. Kernel Polinomial

Tabel 3. *Boxplot* hasil prediksi curah hujan di musim kemarau selama 24 tahun di Kabupaten Bireuen

Variabel	Kernel Linear				Kernel Polinomial			
	Mean	SD	Min	Mak	Mean	SD	Min	Mak
Observasi	59.00	31.74	11.10	118.10	59.90	31.74	11.10	118.10
NASA-GSFC L34	60.01	10.05	48.66	87.17	62.30	8.43	49.25	81.97
METRI AGCM L17	59.37	6.42	46.02	78.13	60.15	9.02	46.47	87.27
PNU	57.37	6.98	46.65	75.03	60.24	8.75	44.64	79.58
MGO	54.75	8.57	39.07	78.11	55.01	11.08	39.37	85.91
BCC	54.94	8.43	38.95	71.33	54.05	9.63	35.85	70.78

Berdasarkan Gambar 5, Gambar 6, dan Tabel 3, analisis *boxplot* mampu menjelaskan ukuran pemusatan dan penyebaran data curah hujan dimusim kemarau. Dari *boxplot* data curah hujan observasi memperlihatkan bahwa data berasal dari distribusi normal dengan rataaan 59.00 dan standar deviasi 31.74.

Model luaran GCM METRI AGCM L17 dan MGO pada fungsi kernel linear terdapat data yang mengandung amatan pencilan (nilai ekstrim) yang menyebar jauh dari amatan lain yang umumnya menyebar normal serta memiliki keragaman yang kecil, sehingga model ini kurang layak digunakan untuk memprediksi curah hujan dimusim kemarau. Data curah hujan model yang memiliki keragaman yang besar merupakan model yang paling baik untuk menjelaskan data curah hujan observasi sehingga nilai prediksi curah hujan model tidak terlalu jauh dari nilai curah hujan observasinya. Sedangkan kernel polinomial hanya terdapat satu model amatan pencilan yaitu model luaran GCM METRI AGCM L17, sehingga model tersebut kurang layak digunakan untuk prediksi curah hujan di musim kemarau. Pada kernel polinomial Model MGO menunjukkan bahwa model ini merupakan model yang cenderung lebih baik untuk menjelaskan hubungannya dengan data curah hujan observasi, karena model ini memiliki ukuran penyebaran data seperti standar deviasi, ukuran pemusatan data, dan rata-rata yang paling mendekati dari data curah hujan observasi, sehingga untuk melakukan prediksi curah hujan di musim kemarau wilayah kabupaten Bireuen, maka model MGO lebih tepat digunakan.

### 3.1. Kinerja Model SVR

Penentuan nilai parameter pada fungsi kernel SVR sangat berpengaruh terhadap hasil yang diperoleh, semakin cocok nilai parameter maka semakin bagus model yang dihasilkan. Ada beberapa faktor yang dapat mempengaruhi kinerja kernel SVR untuk menghasilkan ketepatan dan kecepatan yang baik dalam melakukan prediksi curah hujan dimusim kemarau. Salah satunya adalah pemilihan parameter dengan nilai yang optimal dari setiap fungsi kernel. Penentuan nilai parameter kernel linear dan kernel polinomial dengan cara kombinasi nilai parameter  $C=0.01:100$ ,  $\gamma=0.5:100$ ,  $d=0.01:100$  dan  $\xi=0.0001:100$  dengan jumlah iterasi 6000. Parameter optimal dari  $C=0.5$ ,  $\gamma=0.8$ ,  $d=1$ , dan  $\xi=0.01$ . Lamanya komputasi pada saat pelatihan dan pengujian untuk kernel linear  $\pm 45$  detik, polinomial  $\pm 2$  menit. Pemilihan nilai parameter sangat berpengaruh terhadap waktu komputasi dan korelasi yang dihasilkan. Pengukuran kinerja fungsi kernel SVR dapat dihasilkan dengan melakukan pengamatan terhadap hasil yang diperoleh pada masing-masing fungsi kernel dengan cara mengamati ukuran korelasi dan RMSE pada 5 luaran GCM. Dapat dilihat pada Tabel 4.

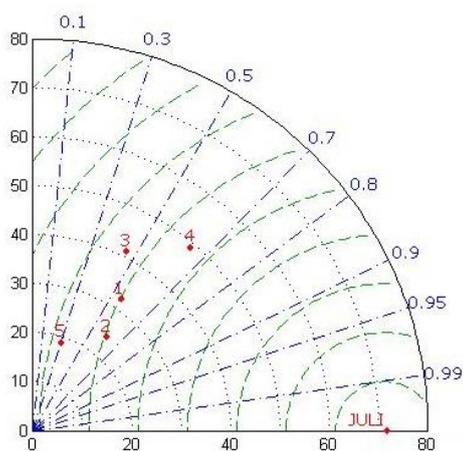
Tabel 4. Perbandingan kinerja kernel linear dan polinomial

Stasiun	Kernel Linear		Kernel Polinomial	
	Korelasi	RMSE	Korelasi	RMSE
Juli	0.580	67.600	0.677	31.428
Peudada	0.468	28.942	0.496	28.284
Peulimbang	0.538	35.483	0.487	35.713
Gandapura	0.668	40.193	0.672	21.106
Jeumpa	0.629	30.801	0.594	25.811
Jeunieb	0.529	27.795	0.475	31.407
Peusangan	0.413	45.566	0.642	35.521
Rata-rata	0.546	39.482	0.577	29.895

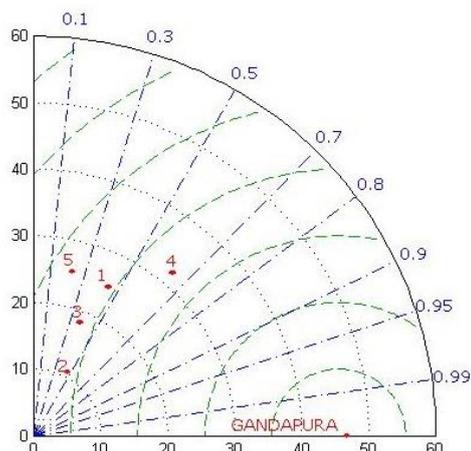
Berdasarkan Tabel 4, rata-rata hasil prediksi semua model pada setiap stasiun menunjukkan bahwa hasil prediksi dengan menggunakan fungsi kernel polinomial diperoleh hasil yang cukup baik bila dilihat dari ukuran korelasi dan RMSE pada 7 stasiun di Kabupaten Bireuen. Hasil rata-rata prediksi yang diperoleh dari model luaran GCM terhadap curah hujan observasi yang baik terdapat pada stasiun wilayah Gandapura yang memiliki ukuran korelasi 0.672 dan RMSE 21.106, nilai RSME yang diperoleh pada wilayah tersebut merupakan nilai RMSE paling rendah bila dibandingkan dengan stasiun wilayah Juli memiliki RMSE yang tinggi yaitu sebesar 31.428 dan walaupun ukuran korelasi paling tinggi 0.677. Kernel SVR yang baik bila memiliki nilai RMSE rendah dan korelasi tinggi. Dari hasil perbandingan kernel SVR, kernel polinomial memiliki nilai korelasi 0.577 dan RMSE 29.895. Lokasi stasiun hujan di wilayah tertentu dapat mempengaruhi hasil prediksi curah hujan di musim kemarau berdasarkan model GCM, hal tersebut menimbulkan dugaan bahwa besarnya tingkat ketepatan suatu model salah satunya tergantung pada hasil *statistical downscaling* pada suatu lokasi pengamatan.

### 3.2. Diagram Taylor

Plot diagram *Taylor* dilakukan pada kinerja kernel polinomial berdasarkan model luaran GCM pada setiap stasiun, karena secara rata-rata model luaran GCM pada kernel polinomial estimasinya cukup baik dan mendekati curah hujan observasi di Kabupaten Bireuen. Pada diagram *Taylor* terdapat 5 keterangan model GCM yaitu 1. NASA-GSFC L34, 2. METRI AGCM L17, 3. PNU, 4. MGO, 5. BCC. Plot Diagram *Taylor* yang dipilih 2 dari 7 stasiun pengamatan yang memiliki hubungan kedekatan hasil korelasi dan prediksi yang tinggi pada stasiun wilayah Gandapura dan yang memiliki hubungan kedekatan hasil prediksi yang baik namun RMSE tinggi pada stasiun wilayah Juli. Berikut model luaran GCM pada fungsi kernel polinomial, seperti ditunjukkan pada Gambar 7 dan Gambar 8.



Gambar 7. Stasiun Juli



Gambar 8. Stasiun Gandapura

Berdasarkan diagram *Taylor* di atas pada Gambar 7, stasiun Juli berada pada standar deviasi  $\pm 71$ . Luaran GCM terbaik pada stasiun ini adalah model 4 yaitu MGO, yang memiliki hubungan kedekatan terhadap curah hujan observasi, karena memiliki ukuran korelasi  $\pm 0.67$  dan RMSE  $\pm 53$  dan standar deviasi model mendekati stasiun pengamatan. Sedangkan pada stasiun Gandapura pada Gambar 8, model yang baik sama seperti stasiun Juli, akan tetapi model MGO jauh lebih baik dari pada stasiun Juli, karena memiliki hubungan yang sangat dekat terhadap stasiun pengamatan yaitu memiliki standar deviasi  $\pm 32$ , RMSE  $\pm 35$  dan besar korelasi  $\pm 0.67$ . Oleh karena itu model 4 lebih berpotensi untuk memprediksi curah hujan di musim kemarau wilayah Gandapura.

### 3. KESIMPULAN

Secara keseluruhan model luaran GCM belum dapat sepenuhnya menjelaskan hasil estimasi terhadap curah hujan observasi, akan tetapi dari 5 model yang digunakan secara rata-rata hanya luaran GCM MGO yang dapat direkomendasikan untuk prediksi curah hujan dimusim kemarau pada kernel polinomial. Waktu yang dibutuhkan pada saat pelatihan dan pengujian untuk kernel linear  $\pm 45$  detik, polinomial  $\pm 2$  menit. Stasiun Gandapura model yang baik yaitu model MGO memiliki standar deviasi  $\pm 32$ , RMSE  $\pm 35$  dan besar korelasi  $\pm 0.67$ . Pemodelan *Statistical Downscaling* data luaran model GCM dapat digunakan untuk memprediksi curah hujan di musim kemarau dengan menggunakan Support Vector Regression, untuk penelitian selanjutnya pengambilan grid GCM dapat dilakukan dengan teknik yang lain seperti *Dynamical Downscaling* dan *Statistical Dynamical Downscaling*.

### UCAPAN TERIMAKASIH

Kami mengucapkan terima kasih kepada Direktorat Riset dan Pengabdian Masyarakat (DRPM) dari Kementerian Ristek Dikti yang mendanai penelitian ini melalui program “skim hibah penelitian dosen pemula tahun 2018 di danai tahun 2019”.

### REFERENSI

- [1] Analisis Musim Hujan 2013/2014 dan Prakiraan Musim Kemarau. Badan Meteorologi Klimatologi dan Geofisika Stasiun Klimatologi Klas II Pondok Betung, 2014.
- [2] Evana L, Effendy S, Hermawan E, “Pengembangan Model Prediksi Madden Julian Oscillation (MJO) Berbasis pada Hasil Analisis Data Real Time Multivariate MJO (RMM1 dan RMM2),” *J. Agromet* 22 (2): 144. Bogor Agricultural University, 2008.
- [3] <http://aceh.tribunnews.com/2018/07/26/451-hektare-padi-se-aceh-kekeringan>.
- [4] Tisseuil C, Vrac M, Lek S, Wade AJ, “Statistical downscaling of river flows,” *J of Hydrology*. 385: 279–291. ScienceDirect, 2010.
- [5] Salvi K, Kannan S, Ghosh, “*Statistical Downscaling and Bias Correction for Projections of Indian Rainfall and Temperature in Climate Change Studies*,” International Conference on Environmental and Computer Science IPCBEE vol.19. IACSIT Press, Singapore, 2011
- [6] Ghosh S, Mujamdar PP, “Future rainfall scenario over Orissa with GCM projections by statistical downscaling,” *J Current Science*. 90(3). India, 2006
- [7] Zhang P, Karori AM, “*Downscaling NCC CGCM output for Seasonal precipitation prediction over Islam abad – pakistan*,” *Pakistan Journal of Meteorology*, Vol. 4 Issue 8, 2008

- [8] Buono, A. et al, "A Neural Network Architecture for Statistical Downscaling Technique: A Case Study in Indramayu District," Dipublikasi dalam *International Conference, The Quality Information for Competitive Agricultural Based Production System and Commerce (AFITA)*. <http://repository.ipb.ac.id/handle/123456789/41728>, 2010.
- [9] Wilby RL, Wigley TM, "Precipitation Predictors For Downscaling: Observed and General Circulation Model Relationships," *International Journal Of Climatology*. Int. J. Climatol. 20: 641 – 661, 2000.
- [10] Aksornsingchai A, Srinilta C, "Statistical Downscaling for Rainfall and Temperature Prediction in Thailand," Proceedings of the International MultiConference of Engineers and Computer scientists, ISSN 2078-0966. Hong Kong, 2011
- [11] Devak M, Dhanya CT, "Downscaling of Precipitation in Mahanadi Basin," *India. J International of Civil Engineering*. 5(2); 111-120. India, 2014
- [12] Haryoko, U. "Pendekatan Reduksi Dimensi Luaran GCM untuk Penyusunan Model *Statistical Downscaling*". [Tesis]. Bogor: Sekolah Pascasarjana, Institut Pertanian Bogor, 2004.
- [13] Wigena HA, "Pemodelan *statistical downscaling* dengan *regression projection persuit* untuk peramalan curah hujan (kasus curah hujan bulanan di Indramayu)," [Disertasi]. Institut Pertanian Bogor. Bogor, 2006
- [14] Agmalaro AM 201, "Pemodelan *statistical downscaling* data GCM menggunakan *support vector regression* untuk memprediksi curah hujan bulanan Indramayu". [Tesis]. Bogor: Sekolah Pascasarjana, Institut Pertanian Bogor, 2011.
- [15] Santosa B, "Data Mining Teknik Pemanfaatan Data untuk Keperluan Bisnis," *Graha Ilmu*. Yogyakarta, 2007.
- [16] Smola A, Schölkopf, B, "A Tutorial on Support Vector Regression," *NeuroCOLT*, Technical Report NC-TR-98-030, Royal Holloway College, University of London, UK, 2003.
- [17] Xu Z, Hou Z, Han Y, Guo W, "A diagram for evaluating multiple aspects of model performance in simulating vector fields," Vol. 9, 4365–4380. *Geosci. Model Dev*, 2016.
- [18] M Mustakim, A Buono, I Hermadi. "Performance Comparison Between Support Vector Regression and Artificial Neural Network for Prediction of Oil Palm Production". *Jurnal Ilmu Komputer dan Informasi* 9 (1). 1-8. 2016.

#### BIBLIOGRAFI PENULIS



Juniana Husna adalah pengajar dan peneliti pada Program Studi Sistem Informasi, Universitas Abulyatama.



Sanusi adalah pengajar dan peneliti pada Program Studi Sistem Informasi, Universitas Abulyatama.