

---

# Pemodelan *Statistical Downscaling* dengan Regresi Modifikasi *Jackknife Ridge Dummy* Berbasis *K-means* untuk Pendugaan Curah Hujan

Dewi Santika Upa P.<sup>1\*</sup>, Sitti Sahrinan<sup>2</sup>, dan Sri Astuti Thamrin<sup>3</sup>

<sup>1,2,3</sup>Departemen Statistika, Fakultas MIPA

Universitas Hasanuddin, Makassar, 90245, Indonesia

\* Corresponding author, email: dewisantikaupap@gmail.com

## Abstract

A Indonesia is a country with a tropical climate, which has two seasons throughout the year, namely the dry season and the rainy season. Thus, it is likely that rain will continue to fall during the dry season, which has serious impacts on various sectors of life. General Circulation Model (GCM) is used to overcome climate change, but GCM cannot simulate well for climate variables at local scale. Therefore, Statistical Downscaling (SD) is used to predict local scale rainfall in Pangkep Regency based on GCM data (CMIP5) in the form of an  $8 \times 8$  grid grid. Jackknife ridge modification regression (MJR) is used to solve the multicollinearity problem that occurs in GCM-lag data. Three dummy variables were added as predictor variables to overcome the heterogeneity of the residual variants. The SD MJR dummy model gives good results based on high correlation and lower root mean square error prediction.

**Keywords:** Dummy Variable, Global Circulation Model, Modified Jackknife Ridge Regression, Multicollinearity, Statistical Downscaling.

## Abstrak

Indonesia adalah negara beriklim tropis, yang memiliki dua musim sepanjang tahun yaitu musim kemarau dan musim penghujan. Sehingga, kemungkinan hujan akan tetap turun pada musim kemarau, yang berdampak serius terhadap berbagai sektor kehidupan. *General Circulation Model* (GCM) digunakan untuk mengatasi perubahan iklim tersebut, namun GCM tidak dapat melakukan simulasi dengan baik untuk peubah iklim yang berskala lokal. Oleh karena itu, *Statistical Downscaling* (SD) digunakan untuk meramalkan curah hujan skala lokal di Kabupaten pangkep berdasarkan data GCM (CMIP5) berbentuk persegi berukuran  $8 \times 8$  grid. Regresi modifikasi *jackknife ridge* (MJR) digunakan untuk mengatasi masalah multikolinearitas yang terjadi pada data GCM-lag. Tiga peubah *dummy* ditambahkan sebagai variabel prediktor untuk mengatasi keheterogenan ragam sisaan. Model SD MJR *dummy* memberikan hasil yang baik berdasarkan korelasi yang tinggi dan *root mean square error prediction* yang lebih rendah.

**Kata Kunci:** Global Circulation Model, Multikolinearitas, Peubah *Dummy*, Regresi Modifikasi *Jackknife Ridge*, *Statistical Downscaling*.

## 1. Pendahuluan

Curah hujan merupakan unsur iklim yang paling dominan dalam mencirikan kondisi iklim di Indonesia karena curah hujan di Indonesia memiliki keragaman dan fluktuasi yang tinggi dibandingkan suhu [1]. Proses pembentukan hujan merupakan proses yang paling sukar untuk disimulasikan, sehingga pendugaan model iklim di Indonesia masih

terus dikembangkan dengan memanfaatkan informasi iklim global yang salah satunya dapat diperoleh dari *Global Circulation Model* (GCM) [2].

GCM merupakan representasi numerik dari gambaran perilaku sistem iklim dan interaksi antar komponen-komponennya, yakni atmosfer, lautan, kriosfer, biosfer, dan kemoster. GCM adalah alat terpenting dalam upaya memahami sistem iklim, khususnya curah hujan. GCM menghasilkan data dalam bentuk *grid* atau petak wilayah dengan resolusi rendah ( $2,5^\circ$  atau  $\pm 300 \text{ km}^2$ ) yang merepresentasikan keadaan iklim global tapi tidak dalam skala regional [3]. Oleh karena itu diperlukan suatu teknik untuk menduga peubah iklim skala lokal dengan tingkat akurasi tinggi yaitu *statistical downscaling* (SD). SD adalah model statistika yang dapat menghubungkan peubah iklim luaran GCM yang berskala global (presipitasi) dengan peubah iklim yang berskala lokal (curah hujan)[4]. Pendekatan SD disusun berdasarkan adanya hubungan antara grid skala global (prediktor) dan grid skala lokal (respon) yang dinyatakan dengan model statistik yang dapat digunakan untuk menerjemahkan anomali-anomali skala global menjadi anomali dari beberapa peubah iklim lokal [5].

Dalam pemodelan SD data peubah GCM berskala besar dalam periode dan jangka waktu tertentu dan digunakan sebagai dasar untuk menentukan data pada *grid* berskala lebih kecil. Namun, pada umumnya data luaran GCM berkarakteristik dimensi besar dan kemungkinan berpotensi terjadinya multikolinearitas sehingga diperlukan metode statistik untuk mengatasi masalah tersebut [6]. Ada beberapa metode yang dapat dilakukan dalam mengatasi masalah multikolinearitas salah satunya adalah *Principal Component Analysis* (PCA), *Liu-type* dan *Ridge Regression* (RR). Dalam hal ini digunakan pengembangan dari metode RR, yaitu Regresi Modifikasi *Jackknife Ridge* (MJR). MJR merupakan metode pendugaan parameter yang menggabungkan ide-ide dari regresi GRR dan JRR, memperlihatkan bahwa MJR lebih baik dari GRR dan JRR berdasarkan nilai MSE.

Batah dkk (2008) melakukan pendugaan parameter regresi MJR dengan hasil MJR lebih efisiensi digunakan disbanding regresi GRR dan JRR. Pada data curah hujan lokal dan GCM merupakan data *time series* dengan skala yang berbeda sehingga memungkinkan adanya pergeseran waktu (*time lag*) antara kedua variabel [7]. Dengan melihat *time lag* kedua variabel tersebut kita akan memperoleh data curah hujan yang sesuai pola curah hujan monsun.

Berdasarkan uraian tersebut maka tujuan penelitian ini yaitu mendapatkan nilai peramalan regresi modifikasi *jackknife ridge* (MJR) dengan peubah *dummy* berdasarkan kluster *K-means* pada data *time lag* GCM untuk peramalan curah hujan .

## **2. Material dan Metode**

Data yang digunakan dalam penelitian ini adalah data sekunder yaitu data curah hujan luaran GCM *Climate Model Intercomparison Project* (CMIP5) dalam satuan mm/bulan yang diperoleh dari web <http://www.climatexp.knmi.nl/> (dikeluarkan oleh

KNMI Belanda), selanjutnya data curah hujan rata-rata di Kabupaten Pangkep, Sulawesi Selatan periode 1997-2018 diperoleh dari stasiun BMKG wilayah IV Makassar. Domain GCM yang digunakan dalam penelitian ini adalah sejumlah grid berbentuk persegi berukuran  $8 \times 8$  grid ( $2.5^\circ \times 2.5^\circ$  untuk setiap grid) pada  $-14,75$  LS –  $5,25$  LU dan  $109,54$  BT –  $129,54$  BT. Variabel penelitian terdiri dari satu variabel respon dan 64 variabel prediktor sesuai ukuran domain GCM. variabel respon ( $y$ ) yaitu intensitas curah hujan (mm/bulan) dan variabel prediktor ( $x$ ) yang akan digunakan dalam penelitian ini yaitu data curah hujan GCM.

Data dianalisis menggunakan *modified jackknife ridge regression* (MJR). Batah dkk (2008) mengusulkan estimator baru yaitu regresi modifikasi jackknife ridge (*modified jackknife ridge*) yang menggabungkan ide-ide *generalized ridge regression* dan *jackknife ridge regression* yang diberikan oleh Singh dkk (1986), Mereka memperoleh kondisi untuk keunggulan MJR melalui JRR dan menunjukkan bahwa estimator MJR umumnya lebih unggul dari GRR dan JRR menggunakan kriteria *mean square error*. GRR dan JRR merupakan pengembangan dari metode *Ridge regression* yang digunakan untuk mengatasi masalah multikolinearitas melalui modifikasi terhadap metode kuadrat terkecil, dengan cara menambah tetapan bias  $k$  yang relatif kecil pada diagonal utama  $X'X$ , (Hoerl dan Kennard, 1970:235). Berdasarkan transformasi variabel  $Y_i^*$  dan  $X_{ij}^*$  yang didefinisikan dengan transformasi korelasi menghasilkan persamaan persamaan umum regresi yang digunakan pada persamaan 1:

$$Y_i^* = \beta_1^* X_{1i}^* + \beta_2^* X_{2i}^* + \dots + \beta_p^* X_{pi}^* + \varepsilon_i^* \quad (1)$$

Persamaan (1) disebut sebagai model regresi yang baku (*standardized regression model*). Diantara parameter  $\beta_1^*$ ,  $\beta_2^*$ , ...,  $\beta_p^*$  pada model regresi baku dengan parameter asli  $\beta_1, \beta_2, \dots, \beta_p$  pada model regresi linear berganda yang biasa terdapat suatu hubungan linear (Kutner dkk, 2005). Model regresi yang baku digunakan setelah mendeteksi multikolinearitas variabel prediktor GCM-lag.

Menentukan variabel prediktor yang orthogonal ( $Z = XT$ ), dimana  $X$  merupakan data yang telah ditransformasi dengan pemusatan dan penskalaan, dengan  $T$  adalah vektor eigen dengan data yang diperoleh dari transformasi korelasi menggunakan persamaan (2) dan (3);

$$Y_i^* = \frac{Y_i - \bar{Y}}{\sqrt{n-1} S_Y} \quad (2)$$

$$X_{ij}^* = \frac{X_{ij} - \bar{X}_j}{\sqrt{n-1} S_{x_j}} \quad (3)$$

Variabel prediktor orthogonal tersebut digunakan untuk menentukan penduga awal metode kuadrat terkecil sesuai persamaan (4);

$$\hat{\alpha}_{OLS} = (Z'Z)^{-1} Z'y \quad (4)$$

Setelah penduga awal MKT didapatkan, selanjutnya menentukan penduga awal regresi MJR sesuai persamaan (5) dan koefisien regresi MJR untuk memperoleh model persamaan regresi modifikasi *jackknife ridge* menggunakan persamaan (6) ;

$$\hat{\alpha}_{MJR} = [I - ((Z'Z + kI)^{-1}kI)^2][-kI(Z'Z + kI)^{-1}]\hat{\alpha}_{OLS} \quad (5)$$

$$\hat{\beta}_{MJR} = T\hat{\alpha}_{MJR} \quad (6)$$

dengan :

$$A = (Z'Z + kI)$$

$T$  = Vektor Eigen

$$Z = XT$$

Dengan adanya keheterogenan ragam sisaan model MJR selanjutnya ditambahkan peubah *dummy* dengan cara mengelompokan data curah hujan ditentukan berdasarkan teknik *k-means*, yang diawali dengan membagi objek ke dalam *k cluster*. Kemudian Memodelkan *statistical downscaling* dengan penambahan peubah *dummy*, langkah akhir yaitu melakukan validasi model dengan tujuan untuk mengukur sejauh mana model yang diperoleh mendekati ukuran sebenarnya menggunakan data periode 2018 menggunakan RMSEP dan menghitung korelasi antara  $y$  (data curah hujan aktual) dengan  $\hat{y}$  (prediksi curah hujan) model regresi modifikasi *jackknife ridge* dengan *dummy*.

### 3. Hasil dan Diskusi

Deskripsi data curah hujan di Kabupaten Pangkep ditunjukkan pada Tabel 1. yang menyajikan hasil perhitungan rata-rata curah hujan di bulan Januari hingga Desember pada periode tahun 1997-2018. Berdasarkan tabel tersebut menunjukkan bahwa pada tahun 1997 hingga 2018 di Kabupaten Pangkep memiliki curah hujan rata-rata 292.80 mm/bulan. Adapun curah hujan tertinggi terjadi pada bulan januari sebesar 1540.5 mm/bulan dan curah hujan terendah adalah 0 mm/bulan. Indonesia merupakan negara dengan pola curah hujan monsun atau berpola huruf U yaitu memiliki satu puncak musim hujan. puncak musim hujan terjadi pada bulan Januari dengan rata-rata curah hujan sebesar 744,05 mm/bulan. Bulan januari dan Desember merupakan bulan dengan curah hujan yang cukup tinggi yaitu sebesar 1.540,5 mm/bulan dan 1.252,5 mm/bulan. dilihat dari nilai rata-rata dan maksimum curah hujan, bulan Januari dan Desember termasuk pada bulan curah hujan tinggi dibandingkan dengan bulan hujan lainnya.

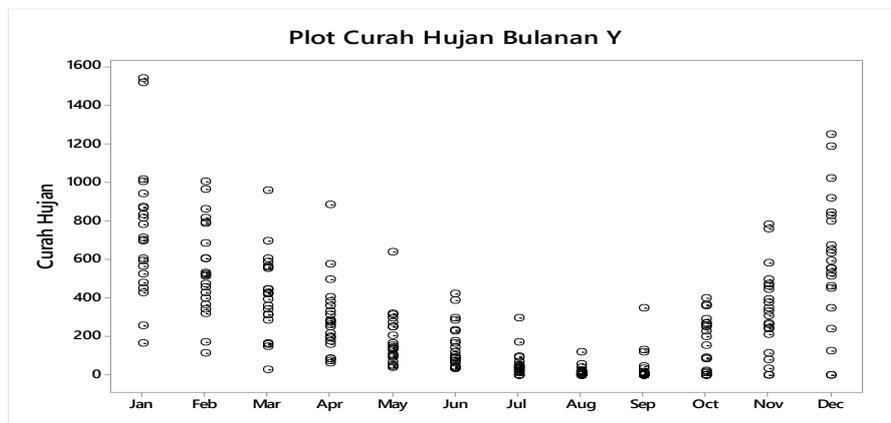
*Time lag* data curah hujan luaran GCM diperoleh dari perhitungan fungsi korelasi silang (CCF) yang di tentukan dari nilai korelasi silang tertinggi antara data curah hujan luaran GCM dengan data curah hujan lokal yang menunjukkan pola yang sama antara data curah hujan tersebut. Berdasarkan perhitungan CCF pada Gambar 1. menunjukan bahwa curah hujan tertinggi di Kabupaten Pangkep rata-rata terjadi pada bulan Januari, Februari dan Desember. Berdasarkan tipe pola curah hujan Indonesia yaitu dengan puncak musim hujan terjadi pada bulan Desember, Januari atau Februari dan puncak musim kemarau terjadi pada bulan Agustus. Hal tersebut menunjukkan bahwa tidak

terjadi pergeseran waktu pada data presipitasi  $X_5$ . Sebaliknya, terjadi pergeseran waktu pada presipitasi  $X_1, X_2, X_7$  dan  $X_8$ . Hasil CCF ysng diperoleh menunjukkan bahwa presipitasi  $X_1, X_2, X_5, X_7$  dan  $X_8$  memiliki korelasi silang tertinggi dengan curah hujan  $Y$  berturut-turut pada *time lag* ke-2 (penundaan 2 bulan), ke-1 (penundaan 1 bulan), ke-0 (tidak ada penundaan), ke-(-1) (penundaan 1 bulan) dan ke-10 (penundaan 11 bulan). Hal ini berarti bahwa curah hujan bulan Januari terjadi di bulan Maret pada data presipitasi  $X_1$ , terjadi di bulan Februari pada data presipitasi  $X_2$ , terjadi di bulan Desember pada data presipitasi  $X_7$ , dan terjadi di bulan November pada data presipitasi  $X_8$ .

Tabel 1. Deskripsi curah hujan bulanan Tahun 1997-2018

Bulan	Rata-Rata	Simpangan Baku	Minimum	Maximum
Jan	744,05	339,16	166,5	1.540,5
Feb	558,20	239,47	116	1.006,5
Mar	417,14	206,52	27,5	958
Apr	288,18	187,89	62	884
May	182,09	135,58	38	641,5
Jun	145,52	114,98	31,5	419
Jul	51,52	67,42	0	296,5
Aug	14,86	26,73	0	117
Sept	33,77	79,46	0	350
Oct	151,64	139,05	0	398
Nov	327,45	216,85	0	779,5
Dec	599,11	338,33	0	1.252,5

*Sumber : Data diolah 2020*



Gambar 1. Pola curah hujan bulanan Kabupaten Pangkep 1997-2018

Hasil perhitungan menunjukkan bahwa data curah hujan luaran GCM-*lag* memiliki nilai VIF berkisar 4,9–2.828,2. Menurut Wirnancy (2015) [5], nilai VIF >10

mengindikasikan adanya multikolinieritas yang tinggi antara grid data curah hujan GCM-lag yang saling berdekatan sehingga data presipitasi GCM tidak bisa langsung digunakan untuk pemodelan. Pemodelan data yang mengandung multikolinieritas menyebabkan dugaan yang tidak tepat sehingga perlu dilakukan penanganan masalah. Untuk mengatasi masalah multikolinieritas dengan metode MJR, terlebih dahulu dilakukan pembakuan variabel dengan menggunakan metode pemusatan dan penskalaan. Dalam pembakuan variabel  $X$  dan  $Y$  menggunakan transformasi korelasi dengan cara menghitung nilai rata-rata dan simpangan baku. Selanjutnya menggunakan Persamaan (6), diperoleh model koefisien regresi modifikasi *jackknife ridge*  $\beta_{MJR}$  pada Tabel 2.

Tabel 2. Koefisien MJR  $\beta_{MJR,J}$

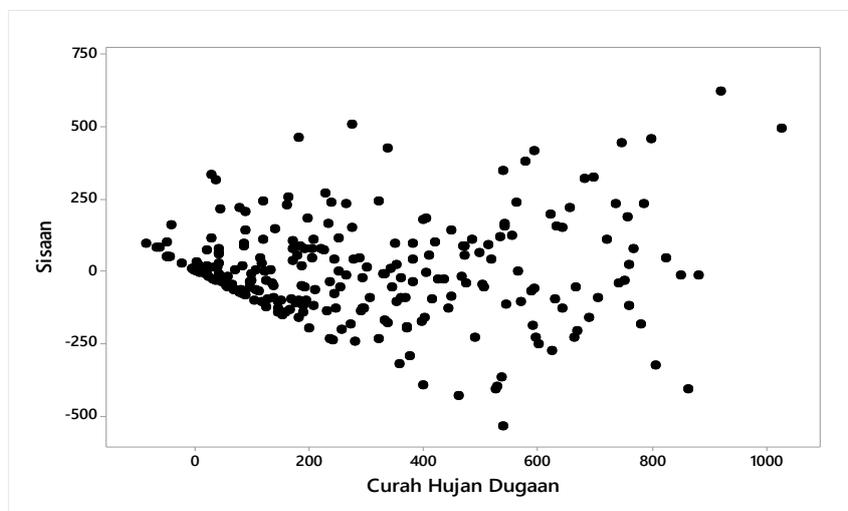
$\beta_{MJR_1}$	0,075	$\beta_{MJR_{17}}$	-0,012	$\beta_{MJR_{33}}$	-2,898	$\beta_{MJR_{49}}$	0,885
$\beta_{MJR_2}$	-0,350	$\beta_{MJR_{18}}$	0,928	$\beta_{MJR_{34}}$	-0,565	$\beta_{MJR_{50}}$	0,461
$\beta_{MJR_3}$	1,237	$\beta_{MJR_{19}}$	1,492	$\beta_{MJR_{35}}$	-0,917	$\beta_{MJR_{51}}$	1,553
$\beta_{MJR_4}$	-0,396	$\beta_{MJR_{20}}$	1,203	$\beta_{MJR_{36}}$	0,262	$\beta_{MJR_{52}}$	0,377
$\beta_{MJR_5}$	0,170	$\beta_{MJR_{21}}$	-0,290	$\beta_{MJR_{537}}$	-0,125	$\beta_{MJR_{53}}$	0,229
	⋮		⋮			⋮	
$\beta_{MJR_{14}}$	-0,272	$\beta_{MJR_{30}}$	0,503	$\beta_{MJR_{46}}$	-0,195	$\beta_{MJR_{62}}$	-0,076
$\beta_{MJR_{15}}$	0,336	$\beta_{MJR_{31}}$	0,177	$\beta_{MJR_{47}}$	0,210	$\beta_{MJR_{63}}$	-0,246
$\beta_{MJR_{16}}$	0,198	$\beta_{MJR_{32}}$	-0,161	$\beta_{MJR_{48}}$	-0,134	$\beta_{MJR_{64}}$	-0,358

Koefisien regresi MJR  $\beta_{MJR}$  yang telah diperoleh pada Tabel 2, selanjutnya dikembalikan dari bentuk  $\beta_j^*$  ke bentuk awal  $\beta_j$ , serta untuk memperoleh nilai  $\beta_0$  dengan melakukan transformasi persamaan regresi sehingga diperoleh model MJR pada persamaan 7.

$$\begin{aligned} \hat{Y} = & -515,32 + 8,71977 X_1 + -36,2567 X_2 + 84,41 X_3 & (7) \\ & - 39,0256 X_4 + 16,43743 X_5 - 33,9284 X_6 \\ & - 49,5679 X_7 + 2,06411 X_8 + 100,2457 X_9 + \dots \\ & + (-10,68 X_{62}) - 38,547 X_{63} - 60,95 X_{64} \end{aligned}$$

Berdasarkan model yang diperoleh, selanjutnya diagnostik model dilihat dengan mencari  $R^2$  dari model tersebut. Hasil yang diperoleh menunjukkan bahwa nilai  $R^2$  yang dihasilkan sebesar 68% dengan *root mean square error* (RMSE) yaitu 165,569. Pemodelan SD dengan regresi modifikasi *jackknife ridge* hanya mampu menjelaskan keragaman data curah hujan terbesar 68%.. Salah satu asumsi yang harus dipenuhi dalam model regresi adalah sisaan memiliki ragam yang homogen (homoskedastisitas). Sisaan memiliki ragam yang heterogen apabila plot antara nilai dugaan dengan sisaannya

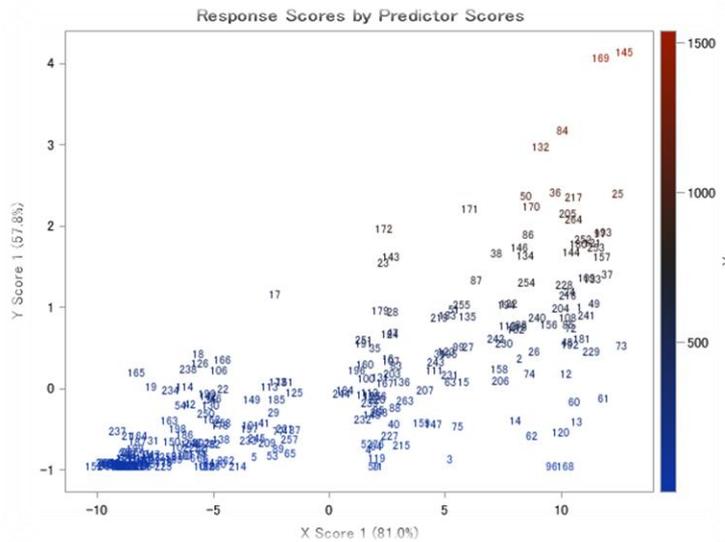
membentuk pola tertentu. Tebaran sisaan dari model regresi modifikasi *jackknife ridge* membentuk pola divergen yang semakin melebar seiring meningkatnya nilai dugaan curah hujan (Gambar 2). Oleh karena itu, peubah *dummy* digunakan untuk mengatasi ragam sisaan yang heterogen.



Gambar 2. Plot sisaan model regresi modifikasi *Jackknife Ridge*

Berdasarkan plot sisaan model regresi MJR yang ditunjukkan pada Gambar 2, menunjukkan bahwa pola keragaman sisaan yang heterogen. Peubah *dummy* ditambahkan ke dalam model regresi MJR bertujuan untuk menangani masalah keragaman sisaan. Pengelompokan data curah hujan ditentukan berdasarkan teknik *k-means*, yang diawali dengan membagi objek ke dalam *k cluster* dimana pada kasus ini peneliti membagi intensitas curah menjadi 4 kelompok seperti pada Gambar 3. Dengan demikian jumlah peubah *dummy* yang digunakan dalam pemodelan SD adalah 3 peubah yakni *D1*, *D2*, dan *D3*.

Berdasarkan Gambar 3. Pengelompokan 3 peubah *dummy* yaitu *D1* digunakan untuk setiap data curah hujan dalam kelompok 1 diberikan nilai 1 pada peubah *dummy* dan 0 untuk kelompok yang lain, jumlah observasi pada *D1* sebanyak 32 yang pada umumnya terjadi pada bulan januari, november dan desember. *D2* setiap data curah hujan dalam kelompok 2 diberikan nilai 1 pada peubah *dummy* dan 0 untuk kelompok yang lain, jumlah observasi pada *D2* sebanyak 92 yang pada umumnya terjadi pada bulan februari dan maret. *D3* setiap data curah hujan dalam kelompok 3 diberikan nilai 1 pada peubah *dummy* dan 0 untuk kelompok yang lain, jumlah observasi pada *D3* sebanyak 134 yang pada umumnya terjadi pada bulan april hingga oktober.



Gambar 3. Plot curah hujan Kabupaten Pangkep

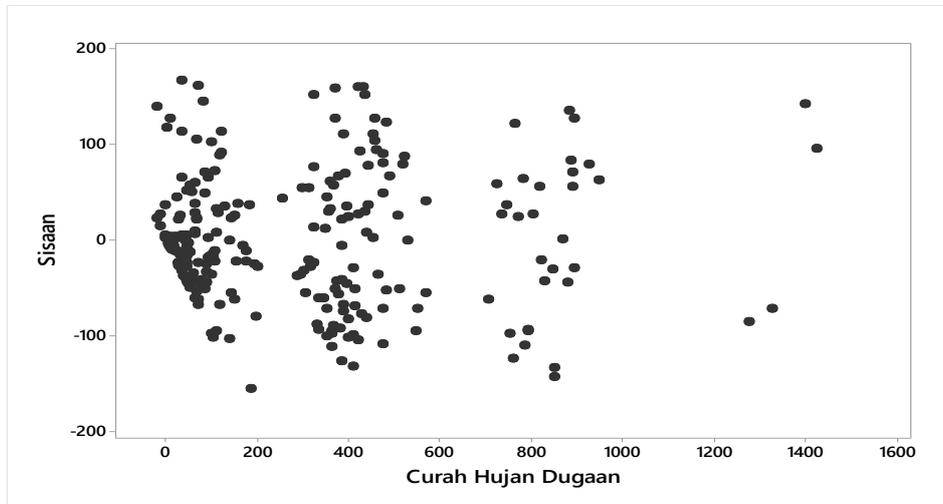
Dengan adanya penambahan peubah *dummy* dalam penelitian dapat mengurangi keheterogenan ragam sisaan dari model Gambar 2. Adapun model MJR *dummy* yang diperoleh dari pemodelan *statistical downscaling* dapat dilihat pada persamaan 8:

$$\begin{aligned} \hat{Y} = & 1.374,474 - 36,972X_1 - 1,216 X_2 + 120,671 X_3 - 34,925 X_4 \quad (8) \\ & + 13,4X_5 - 38,751X_6 + 7,075X_7 + 4,31X_8 - 2,767X_9 + \dots \\ & + (-36,36X_{62}) + 17,774X_{63} + 0,561 X_{64} - 552,444 D_1 \\ & - 962,147 D_2 - 1.195,97 D_3 \end{aligned}$$

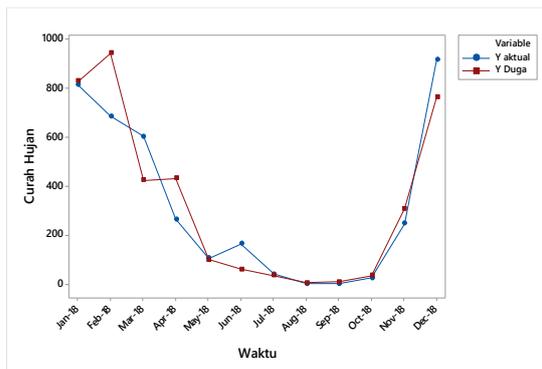
Berdasarkan model MJR *dummy*, dapat dikatakan bahwa model regresi modifikasi *jackknife ridge dummy* merupakan model yang lebih baik dibandingkan model regresi modifikasi *jackknife ridge* tanpa *dummy* berdasarkan nilai *mean square error* (MSE) yang pun relatif kecil, yaitu 66,603 . Selain itu, model regresi modifikasi *jackknife ridge dummy* mampu menjelaskan keragaman data lebih besar, yaitu dengan  $R^2$  sebesar 95%. Berdasarkan nilai RMSE dan  $R^2$  yang diperoleh berarti dengan penambahan peubah *dummy* pada model SD dengan regresi modifikasi *jackknife ridge* nilai berkurang sebesar 98,96, serta nilai  $R^2$  meningkat sebesar 27%. Pada Gambar 4 tebaran sisaan model regresi modifikasi *jackknife ridge dummy* lebih homogen dibandingkan sisaan tanpa peubah *dummy*.

Data curah hujan periode Januari hingga Desember Tahun 2018 digunakan sebagai data validasi. Hasilnya menunjukkan bahwa model regresi modifikasi *jackknife ridge dummy* mampu menjelaskan keragaman curah hujan yang akurat. Model regresi modifikasi *jackknife ridge dummy* menghasilkan nilai korelasi sebesar 0,936 dengan nilai RMSEP 120,545 dibandingkan dengan model regresi modifikasi *jackknife ridge* tanpa *dummy* menghasilkan nilai korelasi sebesar 0,85 dengan nilai RMSEP 176,768. Secara grafik, model regresi modifikasi *jackknife ridge dummy* memperlihatkan performa model

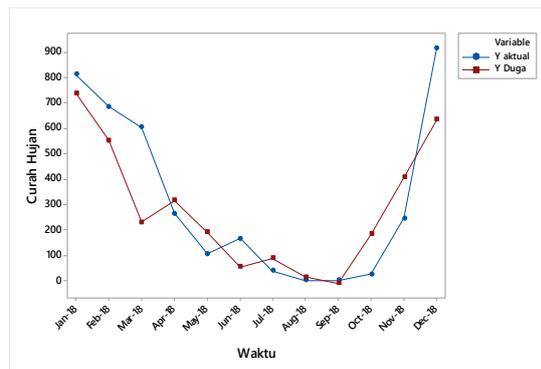
yang baik pada Gambar 5. Pemodelan *statistical downscaling* dengan regresi modifikasi *jackknife ridge dummy* mampu melakukan pendugaan curah hujan yang pola curah hujan aktual dan mampu menangkap curah hujan. Selain itu, dilihat dari jarak yang dekat antara curah hujan aktual dan hasil dugaannya.



Gambar 4. Plot sisaan model regresi modifikasi *Jackknife Ridge dummy*



Gambar 5. Plot curah hujan aktual dan curah hujan dugaan regresi MJR



Gambar 6. Plot curah hujan aktual dan curah hujan dugaan MJR tanpa *dummy*

#### 4. Kesimpulan

Penambahan variabel *dummy* pada regresi modifikasi *jackknife ridge* untuk model *statistical downscaling* digunakan untuk menangani masalah ragam sisaan yang heterogen. Adanya 3 peubah *dummy* yang disertakan ke dalam model merupakan hasil pengelompokan teknik *cluster* berbasis *k-means* yang mengelompokkan data curah hujan berdasarkan jarak kedekatannya. Model *statistical downscaling* dengan regresi

modifikasi *jackknife ridge dummy* memiliki nilai  $R^2$  lebih besar dan mean square of error (MSE) yang lebih rendah dibandingkan regresi modifikasi *jackknife ridge* tanpa *dummy*. Dan berdasarkan nilai *root mean square error prediction* yang dihasilkan juga lebih rendah. Selain itu, hasil prediksi curah hujan yang dihasilkan mampu menangkap pola curah hujan aktual dengan baik yaitu sebesar 0,937.

### **Daftar Pustaka**

- [1] Sipayung, S. B. Dampak variabilitas iklim terhadap produksi pangan di Sumatera. *Jurnal Sains Dirgantara*, 2(2): 111–126, 2005.
- [2] Sahrinan, S., Anisa, & Koerniawan V. Pemodelan *Statistical Downscaling* dengan Peubah Dummy berdasarkan Teknik Cluster Hierarki dan Non Hierarki untuk penduga curah hujan. *Indonesian Journal of Statistics and Its Applications*, Vol 3 No 3: 295 – 309, 2019.
- [3] Notodiputro, A. K., Wigena, A. H., & Fitriadi. Pendekatan Regresi Komponen Utama dan ARIMA untuk *Statistical Downscaling*. Disertasi. Bogor (ID): Institut Pertanian Bogor. *Journal Sains Teknologi*, Vol. 11 (3), 2005.
- [4] Zorita, E., & Storch, V. H. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. *J Clim*, 12 : 2474-2489, 1999.
- [5] Wirnancy, J. S. Model *Statistical Downscaling* dengan Regresi Kuantil Komponen Utama Fungsional untuk Prediksi Curah Hujan Ekstrem [tesis]. Bogor (ID): Institut Pertanian Bogor, 2015.
- [6] Wigena, A. H. Pemodelan *Statistical Downscaling* dengan *Regresi Projection Pursuit* untuk Peramalan Curah Hujan Bulanan: Kasus Curah Hujan Bulanan diIndramayu. Disertasi. Bogor (ID): Institut Pertanian Bogor, 2006.
- [7] Batah, F. S., Ramanathan, T. V., & Gore, S. D. The efficiency of modified jackknife and ridge type regression estimators: A comparison. *Survey In Mathematics and its Applications*, 3: 111-122, 2008.