

PENGGUNAAN REGRESI KUADRAT TERKECIL PARSIAL (PLS) UNTUK MENGATASI KOLINIERITAS DAN PEMANFAATANNYA UNTUK PREDIKSI CURAH HUJAN TIPE MONSON BERDASARKAN DATA GCM (*Global Circulation Model*)

THE USING OF PARTIAL LEAST SQUARE REGRESSION (PLSR) TO SOLVE THE COLINIERTY PROBLEM AND ITS UTILIZATION IN PREDICTING THE MONSOON RAINFALL , BASED ON GLOBAL CIRCULATION MODEL (GCM) DATA

GUSTI RUSMAYADI¹ dan UMI SALAWATI²

¹Jurusan Budidaya Pertanian Fakultas Pertanian UNLAM

²Jurusan Sosial Ekonomi Pertanian Fakultas Pertanian UNLAM
Jl. Jend. A. Yani Km.36 PO Box 1028 Banjarbaru 70714

ABSTRACT

Prediction of parameter in multiple regression analysis is an interesting topic of some researches. This can be due to some problems that may rise in regression, and one of these is a colinierty problem. The method usually used to solve this problem is Principle Component Regression (PCR), Ridge Regression, and Partial Least Square (PLS). Differ from PCR and Ridge Regression, regression coefficient in the PLS is obtained iteratively and doesn't have a closed formula to get a variant of regression coefficient. GCM data were obtained by using the grade ranged from v2272 – v2727 and the rainfall data from Kotabaru BMKG Station collected from 1996 – 2001. These data were then used as a model, whereas the data collected in 2001 was used to validate the model. All calculation and figures were made by using Minitab software 1412 version and Microsoft Excel 2003. Estimation with PLS algorithm to various iteration resulted in PRESS and R² values. Variant of X variable with 3 components of model could explain a 66.9% of variant of dependent variables. Minimum PRESS with the value of 5.714.118 was reached in the third iteration with the highest prediction of R² value as 7%. Regression analysis using PLS method based on GCM and rainfall data or is known as Statistical Downscaling, resulted in a value prediction of a coefficient regression in the third iteration. This model has been validated and can be used to predict rainfall during the next 12 months. It can conclude that the optimum PLS model using GCM data can be used to predict a local climate, whereas as a suggestion is that we still need to see the relation between rainfall variable and seasonal period, so that it can be used to predict every period of event, as in case of DJF periode (Dec, Jan, and Feb).

Key words : PLS, PRESS, colinierty, iteration

ABSTRAK

Pendugaan parameter dalam analisis regresi berganda merupakan topik menarik dalam beberapa penelitian. Itu disebabkan beberapa kendala yang dapat muncul dalam regresi yang salah satunya adalah masalah kolonieritas. Metode yang biasa digunakan untuk mengatasi kolinearitas adalah regresi komponen utama (PCR), regresi ridge dan regresi kuadrat terkecil parsial (PLS). Berbeda dengan regresi ridge dan regresi komponen utama, pada regresi PLS, koefisien regresi diperoleh secara iteratif dan tidak memiliki formula tertutup untuk mencari ragam koefisien regresi. Data GCM diambil dengan grid antara V2272 – V2727 tahun 1996 – 2001 dan data curah hujan Stasiun BMKG Kotabaru (Sta.310L) tahun 1996 – 2001. Data tahun 1996 – 2000 digunakan untuk membuat model, sedangkan data tahun 2001 dipergunakan untuk memvalidasi model. Semua perhitungan dan gambar menggunakan Software Minitab versi 14.12 dan Microsoft Excel 2003. Pendugaan dengan algoritme PLS untuk berbagai iterasi menghasilkan nilai PRESS dan R². Varian dari peubah X dengan 3 komponen model dapat menjelaskan 66,9% dari varian dalam peubah penjelas. PRESS minimal dengan nilai 5.714.118 tercapai pada iterasi yang ke-3 dengan nilai R² prediksi tertinggi sebesar 7%. Analisis regresi dengan metode PLS terhadap data GCM dan curah hujan atau yang dikenal sebagai *statistical downscaling*, menghasilkan nilai dugaan persamaan regresi dengan iterasi ke-3. Model tersebut dapat dipergunakan untuk memprediksi curah hujan selama 12 bulan ke depan dan telah divalidasi. Kesimpulannya adalah bahwa model PLS yang optimal dengan menggunakan data GCM dapat dimanfaatkan dalam peramalan iklim lokal. Sedangkan saran adalah perlu melihat hubungan antara peubah curah hujan dengan peubah GCM pada periode musiman sehingga diharapkan dapat digunakan untuk memprediksi pada setiap periode kejadiannya, misalnya periode DJF (Desember, Januari dan Februari).

Kata kunci: PLS, PRESS, kolinieritas, iterasi

PENDAHULUAN

1.1. Latar Belakang

Pendugaan parameter dalam analisis regresi berganda merupakan topik menarik dalam beberapa penelitian, karena ada beberapa kendala yang dapat muncul dalam regresi yang salah satunya adalah masalah kolinearitas.

Kolinearitas pada regresi berganda terjadi karena korelasi yang tinggi di antara peubah bebas. Penggunaan metode kuadrat terkecil akan memperlihatkan pengaruh kolinearitas, yaitu koefisien determinasi (R^2) yang tinggi tetapi tidak diikuti oleh hasil uji hipotesis yang nyata dari koefisien penduga. Keadaan tersebut disebabkan oleh koefisien determinasi yang besar tidak didukung oleh ragam koefisien penduga yang kecil (Martens & Naes, 1989).

Metode-metode yang biasa digunakan untuk mengatasi kolinearitas adalah regresi komponen utama (PCR), regresi ridge dan regresi kuadrat terkecil parsial (PLS). PCR adalah regresi dengan mengambil komponen utama sebagai peubah bebas. Koefisien penduga dari metode ini diperoleh melalui pengurangan dimensi peubah penduga komponen utama, sehingga subset komponen utama yang dipilih harus tetap mempertahankan keragaman yang besar terhadap responnya. Regresi ridge (bertatar) merupakan modifikasi dari metode kuadrat terkecil dengan cara menambah tetapan bias c yang kecil pada nilai diagonal matrik $X'X$. Tetapan bias c yang besar mencerminkan bias dalam koefisien koefisien penduga ridge yang besar, c yang bernilai nol merupakan implimentasi dari metode kuadrat terkecil. Metode PLS melakukan pendugaan internal pada regresi kuadrat terkecil parsial secara iteratif dengan melibatkan struktur keragaman peubah bebas dan responnya.

Berbeda dengan regresi ridge dan regresi komponen utama, pada regresi PLS, koefisien regresi diperoleh secara iteratif dan tidak memiliki formula tertutup untuk mencari ragam koefisien regresi. Pada dasarnya metode PLS dapat menjelaskan hubungan antara dua kelompok peubah sehingga metode ini dapat digunakan untuk pendugaan model kalibrasi peubah ganda dengan banyak peubah y .

Metode PLS adalah salah satu alternatif untuk mengatasi kolinearitas yang sering ditemui dalam pemodelan statistika. Beberapa peneliti yang melakukan penelitian tentang metode ini dalam model kalibrasi ganda diantaranya adalah Almoy (1996), Herwidiati (1997), Indriati (1997), Wigena dan Aunuddin (1998), Antou (2000) dan Wulandari (2000).

Kebanyakan penelitian di atas yang menggunakan metode PLS diterapkan pada bidang Chemometric. Menurut Wigena dan Aunuddin (1998) metode PLS dapat diterapkan dalam bidang ilmu lain seperti biologi, lingkungan hidup, dll.

Sifat terpenting dari model PLS adalah penentuan banyak komponen yang digunakan. Walaupun dimungkinkan untuk menghitung komponen PLS sebanyak pangkat matrik kelompok X , namun umumnya tidak semua komponen digunakan, karena data yang diukur tidak pernah bebas galat dan beberapa komponen yang terakhir hanya menerangkan galat sehingga dapat menyebabkan kolinearitas. Ini artinya harus ada suatu kriteria untuk menghentikan proses. Akurasi pendugaan yang dipergunakan dalam model PLS adalah PRESS (Prediction Residual Sum of Squares).

Model mekanistik seperti GCM dapat memprediksi kondisi suhu muka laut sampai tiga dan enam bulan ke depan dengan tingkat ketepatan yang layak dan juga data iklim lainnya. Akan tetapi resolusi data sangat rendah (350×350) km^2 sehingga informasi ramalan sangat bersifat global dan tidak cocok untuk skala lokal. Oleh karena itu, cara yang digunakan biasanya ialah dengan menggunakan GCM untuk menduga kondisi suhu muka laut atau tekanan dan kemudian hasil ramalan ini dihubungkan dengan keragaman hujan di lokasi yang diinginkan dengan menggunakan persamaan regresi atau yang diistilahkan sebagai statistical downscaling.

1.2. Kolinearitas

Korelasi yang tinggi diantara peubah bebas menyebabkan kolinearitas. Salah satu ukuran mengidentifikasi keberadaan kolinearitas adalah VIF. VIF merupakan faktor yang mengukur seberapa besar kenaikan ragam dari koefisien penduga b_j dibandingkan terhadap peubah bebas lain yang saling orthogonal. Apabila R_j^2 merupakan koefisien determinasi dari peubah bebas X_j yang jika diregresikan terhadap semua peubah bebas X lainnya, maka VIF _{j} adalah:

$$VIF_j = \frac{1}{(1 - R_j^2)} \quad (1.1)$$

(Fox dan Monette, 1992)

Apabila diantara peubah-peubah bebas tersebut terdapat korelasi yang tinggi, maka nilai VIF akan besar. VIF _{j} akan memiliki nilai mendekati 1 (satu) jika peubah bebas X_j tidak ada korelasi dengan peubah-peubah bebas lainnya.

1.3. Koefisien Determinasi

Koefisien determinasi, R^2 , merupakan suatu ukuran dari peubah bebas dalam model regresi yang dirumuskan sebagai:

$$R^2 = \frac{JKR}{JKT(\text{terkoreksi})} = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad (1.2)$$

(Draper dan Smith, 1992)

dengan, JKR adalah Jumlah Kuadrat Regresi dan JKT adalah Jumlah Kuadrat Total.

Dalam hal ini pejumlahan dilakukan untuk $i = 1, 2, \dots, n$. Jadi, R^2 mengukur "proporsi keragaman atau variasi total di sekitar nilai

tengah \bar{Y} yang dapat dijelaskan oleh regresi tersebut". Ukuran ini sering diucapkan sebagai persentase dengan mengalikannya dengan 100. R^2 bernilai antara 0 sampai dengan 1, R^2 yang semakin besar menunjukkan ketepatan model yang semakin baik dalam menerangkan keragaman data. Sesungguhnya, R adalah

korelasi antara Y dengan \bar{Y} , yang biasanya disebut koefisien korelasi ganda.

BAHAN DAN METODE

2.1. Sumber Data

Data dari Global Circulation Model(GCM) dengan grid antara V2272 – V2727 tahun 1996 – 2001 dan data curah hujan Stasiun dan BMKG Kotabaru (Sta. 310L) tahun 1996 – 2001.

2.2. Metode

Data tahun 1996 – 2000 digunakan untuk membuat model, sedangkan data tahun 2001 dipergunakan untuk memvalidasi model. Semua perhitungan dan gambar menggunakan Software Minitab® versi 14.12 dan Microsoft Excel® 2003.

2.3. Partial Least Square

Partial Least Square merupakan metode "soft modelling" yang dapat mengoptimalkan model kalibrasi peubah Y dan X (Wigena dan Aunuddin, 1998). Metode PLS hampir sama dengan metode Analisis Komponen Utama (AKU). Jika pada AKU komponen utama dihitung masing-masing pada peubah-peubah X dan peubah-peubah Y secara terpisah, maka pada PLS komponen utama disusun berdasarkan konsep NIPALS dengan cara dekomposisi nilai tunggal (Single Value Decomposition).

Pada setiap iterasi dalam PLS keragaman peubah-peubah X dan keragaman peubah-peubah Y saling mempengaruhi. Struktur ragam kelompok peubah Y mempengaruhi kombinasi linear kelompok peubah X dan sebaliknya, struktur ragam kelompok peubah X mempengaruhi kombinasi linear kelompok peubah Y (Young, 1994).

Model ini berdasarkan peubah internal atau peubah latent (latent variable) (Geladi & Kowalski, 1986). Model hubungan tersebut adalah:

1. Hubungan internal

$$\eta = \beta_0 + \beta_1 \xi + \delta \quad (1.3)$$

dengan η dan ξ adalah skor dari Y dan X. β_0 dan β_1 adalah koefisien internal dan δ adalah faktor acak

2. Hubungan eksternal

$$X = \pi_0 + \pi_1 + \varepsilon \quad (1.4)$$

$$Y = (\pi_0 + \pi_1 \beta_0) + \pi_1 \beta_1 \xi + (\varepsilon + \pi_1 \delta) \quad (1.5)$$

dengan π_0 dan π_1 adalah koefisien model eksternal bagi x, dan ε adalah faktor acak.

2.4. Algoritme PLS

Metode PLS digunakan untuk memodelkan hubungan antara peubah tak bebas Y dengan satu kelompok peubah bebas X berdasarkan model (2.3). Algoritmanya disusun berdasarkan algoritme NIPALS (Non Linier Iteratif Partial Least Square) yang menggunakan metode komponen utama dengan cara dekomposisi nilai singular (singular-value decomposition), tetapi ada perbedaan dalam pendugaan modelnya, yaitu dengan cara sebagai berikut (Martens & Naes, 1989):

$$X = t_1 p'_1 + t_2 p'_2 + \dots + t_A p'_A + E_A \quad (1.6)$$

dengan t_a adalah vektor berdimensi n, sebagai vektor skor (*score vector*), sedangkan p_a adalah vektor berdimensi k, sebagai vektor muatan (*loading vector*). Matriks E_A berukuran $n \times k$ adalah matriks sisaan. Dasar metode PLS, hubungan antara X dan Y melalui vektor-vektor tersebut, sehingga

$$Y = t_1 q_1 + t_2 q_2 + \dots + t_A q_A + f_A \quad (1.7)$$

dengan q_a berupa skalar muatan dan f_A adalah vektor sisaan.

Berdasarkan algoritma di atas, maka diperoleh model persamaan regresi pada iterasi ke – A:

$$\hat{y} = Xb_A \quad (1.8)$$

dengan $b_A = W_A(P_A'W_A)^{-1}Q_A$ (Helland, 1988) (1.9)

$$W_A = (w_1, \dots, w_A), P_A = (p_1, \dots, p_A)$$

dan $Q_A = (q_1, \dots, q_A)$

Proses menentukan banyak komponen yang diperlukan (A) digunakan validasi silang. Pada setiap iterasi kumpulan data dihapus dan dilakukan kalibrasi terhadap (n-1) sisanya, kemudian dilakukan pendugaan terhadap titik-titik data dalam data yang dihapus. Jumlah kuadrat selisih y amatan dan y dugaan dari titik-titik data yang dihapus disebut PRESS (Prediction Residual Sum of Square). Jumlah komponen A sesuai dengan iterasi yang memberikan PRESS minimal (Geladi & Kowalski, 1986).

2.5. PRESS

PRESS merupakan salah satu pendekatan yang dipertimbangkan untuk prosedur validasi kestabilan koefisien penduga b (Weiberg, 1985). Nilai PRESS yang lebih kecil atau nilai minimum akan memberikan kestabilan lebih tinggi terhadap model jika ada amatan yang baru.

Lokasi dari nilai minimum tidak selalu terdefinisi dengan baik. Evaluasi terhadap banyak komponen analog dengan konsep pendeteksian, yaitu sinyal terlemah yang bisa dideteksi dalam keberadaan gangguan (Geladi dan Kowalski, 1986).

PRESS mempunyai kelebihan karena memberi banyak informasi terinci tentang kestabilan berbagai persamaan regresi dalam ruang datanya, dan dapat memusatkan perhatian pada titik data yang berpengaruh (Draper dan Smith, 1992)

Misalkan p adalah parameter termasuk

β_0 dalam satu persamaan regresi dan n adalah jumlah amata, PRESS didefinisikan sebagai:

$$PRESS = \sum_{i=1}^n \left(Y_i - \hat{Y}_{i-i} \right)^2 \quad (1.8)$$

dengan $\hat{Y}_{i-i} = X'_{i-i} b$ dan b_{-i} adalah koefisien regresi dengan cara membuang amatan yang ke-i.

Setelah data amatan ke-1 dihapus, maka langkah berikutnya dilakukan pada data ke-dua, jumlah kuadrat selisih y amatan dan y dugaan dalam langkah kedua ini ditambahkan pada PRESS sebelumnya, demikian seterusnya sampai data ke n dihapus.

Tahapan algoritma PLS adalah sebagai berikut (Martens dan Naes, 1989):

1. Tahapan 1.1 s/d 1.5 diulang untuk setiap komponen:
 - 1.1. Tentukan penduga vektor pembobot w_a
 - 1.2. Tentukan penduga vektor skor t_a
 - 1.3. Tentukan penduga vektor muatan p_a
 - 1.4. Tentukan penduga vektor muatan q_a
 - 1.5. Tentukan penduga E (=residual X) dan f (=residual Y)
2. Tentukan jumlah faktor (A) dalam model
3. Tentukan penduga Y berdasarkan A komponen.

Perhitungan dengan algoritme PLS ini harus dilakukan dengan iterasi sebanyak jumlah peubah X. Dugaan koefisien regresinya dilihat pada saat iterasi dengan nilai PRESS minimal (Geladi dan Kowalski, 1986; Young, 1994).

T u j u a n

Penelitian ini bertujuan untuk menerapkan model PLS untuk mengatasi masalah kolinearitas dalam memanfaatkan data GCM untuk peramalan iklim lokal

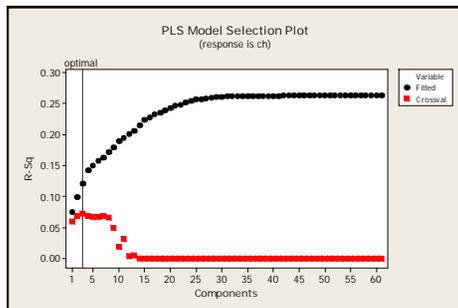
HASIL DAN PEMBAHASAN

3.1. Hasil Algoritma PLS

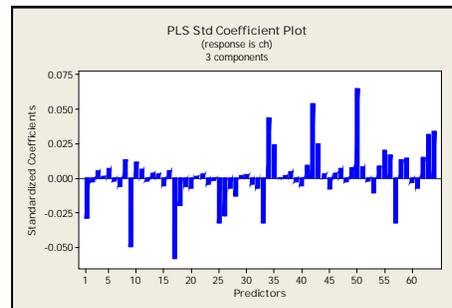
Pendugaan dengan algoritme PLS untuk berbagai iterasi menghasilkan nilai PRESS dan R^2 yang dapat dilihat pada Tabel 4.1. Varian dari peubah X dengan 3 komponen model dapat menjelaskan 66,9% dari varian dalam peubah penjelas. PRESS minimal dengan nilai 5714118 tercapai pada iterasi yang ke-3 dengan nilai R^2 prediksi tertinggi sebesar 7%. Gambaran seluruh nilai PRES dan R^2 ditampilkan pada Gambar 4.1 dan Lampiran 1.2.

Tabel 4. 1. Beberapa komponen hasil metode PLS
 Table 4. 1. Some method yield components PLS

Komponen	Peubah X % varians	PRESS	Error	R (predicted)
1	45,9003	5792611	5702857	0.0596275
2	61,1102	5738135	5547510	0.0684713
3	66,9100	5714118	5419588	0.0723702
4		5738832	5283608	0.0683581
5		5750479	5235905	0.0664672
6		5745194	5191832	0.0673252
7		5740984	5155962	0.0680087
8		5751945	5106592	0.0662292
9		5854075	5057382	0.0496494
10		5792611	5702857	0.0182814



(a)



(b)

Gambar 4.1a. Plot seleksi model PLS dan Plot koefisien regresi PLS distandarkan
 Picture 41a. Selection plot of Model PLS and Plot coefficient regresi PLS standard

Gambar 4.1a menampilkan plot Seleksi Model dan Tabel Validasi. Garis vertikal mengindikasikan bahwa model yang optimal dengan 3 (tiga) komponen. Selain itu, Gambar 4.1a menunjukkan pula kemampuan prediksi seluruh model dengan lebih dari 3 (tiga) komponen yang menurun secara nyata, sampai dengan komponen yang ke-14 dan seterusnya sebanyak jumlah peubah X (64 peubah).

Kemampuan memprediksi nilai R^2 dengan lebih dari 3 (tiga) komponen sama dengan nol (Lampiran 1.2). Menurut Shao (1993) komponen yang dipilih yang memiliki PRESS terkecil karena memberikan kestabilan lebih tinggi terhadap model jika ada pengamatan baru.

Koefisien standar prediktor Gambar 4.1b menunjukkan besaran dan tanda hubungan dari koefisien. Analisis varian per respon berdasarkan model yang optimal atau pada saat nilai PRES minimal, menunjukkan nilai p adalah 0.000, yang kurang dari 0.05, artinya model regresi tersebut nyata (Tabel 4.2).

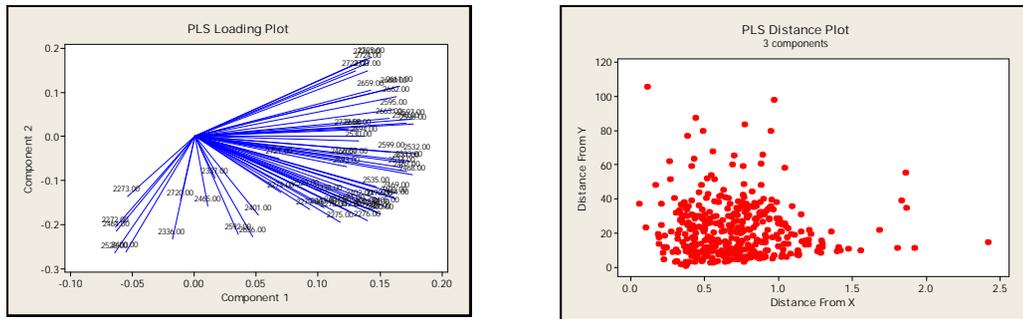
Tabel 4.2. Analysis of Variance for ch
 Table 4.2. Analysis of variance for ch

Source	DF	SS	MS	F	P
Regression	3	740324	246775	17.30	0.000
Residual Error	380	5419588	14262		
Total	383	6159912			

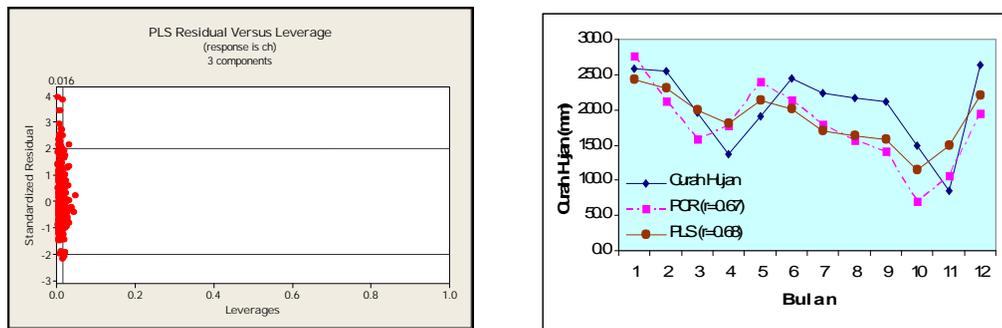
$$R^2 = 46,52\%$$

Ada 35 peubah X yang mempunyai hubungan positif dengan curah hujan dan sisanya 29 yang mempunyai hubungan negatif.

Kemudian dari ke- 35 peubah X tadi, ada 20 peubah di dalamnya yang mempunyai koefisien standar yang besar dan berpengaruh besar pada curah hujan, yaitu grid -2272.00, -2336.00, 2343.00, -2400.00, -2401.00, -2464.00, -2465.00, -2471.00, -2528.00, 2529.00, 2593.00, 2657.00, 2662.00, 2663.00, -2720.00, 2721.00, 2722.00, 2725.00, 2726.00, 2727.00 (Lampiran 1.3). (Tanda – menunjukkan hubungan negatif).



(a) (b)
 Gambar 4.2. Plot loading PLS (a) dan Plot jarak PLS (b)
 Figure 4.2. PLS loading plot (a) and PLS distance plot (b)



(a) (b)
 Gambar 4.3. Sisaan PLS versus Leverage (a) dan Validasi prediksi curah hujan model PLS dan PCR dengan data observasi (b)
 Figure 4.3. Residual PLS vs Leverage PLS (a) and validation of rainfall prediction using PLS and PCR models with observation data.

Gambar 4.2a menunjukkan bahwa grid dengan garis yang panjang mempunyai *loading* yang besar dan berhubungan kuat dengan curah hujan, misalnya grid 2272.00 dan sebaliknya grid dengan garis yang pendek memiliki *loading* yang kecil dan hubungan yang relatif lemah dengan curah hujan, misalnya grid 2273.00.

Gambar 4.2b adalah plot jarak PLS yang menunjukkan seberapa besar jarak pengamatan dari sumbu x dan sumbu y. Pada gambar terlihat sejumlah data dengan jarak cukup besar baik terhadap sumbu X, maupun sumbu Y.

Gambar 4.3a adalah sisaan PLS versus Leverage yang menunjukkan seberapa banyak data pencilan. Dalam gambar terlihat sejumlah data pencilan karena ke luar dari garis referensi baik yang horizontal maupun vertikal.

Secara umum, pencilan merupakan data yang tidak sesuai, juga sangat menyalahi dan tidak wajar dibandingkan dengan mayoritas data (Martens & Naes, 1989). Pencilan sering menyebabkan kesalahan dalam pemilihan model dan biasanya dihilangkan atau dikoreksi.

Padahal, dalam kenyataannya beberapa pencilan merupakan amatan yang sangat informatif dan bernilai. Pencilan dapat disebabkan oleh kesalahan data atau suatu komposisi atau status fisik yang ganjil dari objek yang dianalisis. Kesalahan dalam data dapat berupa gangguan, penyimpangan instrumen, kesalahan operator, maupun kesalahan pencetakan.

Berdasarkan analisis pencilan di atas (Lampiran 2.1 s/d 2.4) terlihat bahwa amatan ke-74, 157, dan 184 memisah dari amatan lainnya. Kemudian dalam artikel ini dicoba untuk menghilangkan pencilan, yaitu pada amatan ke-184 dan hasilnya dapat dilihat pada Lampiran 2.8. Pada gambar tersebut terjadi peningkatan koefisien korelasi antara data dugaan dengan data amatan dari 0,68 menjadi 0,69. Dugaan persamaan regresi dalam sub bab 4.2 hanya menyajikan analisis korelasi antara data dugaan dengan data amatan sebelum pencilan dibuang.

3.2. Dugaan Persamaan Regresi

Analisis regresi dengan metode PLS terhadap data GCM dan curah hujan, menghasilkan nilai dugaan persamaan regresi dengan iterasi ke-3 seperti yang terdapat pada Lampiran 1.3. Model tersebut dapat dipergunakan untuk memprediksi curah hujan selama 12 bulan ke depan dan divalidasi seperti pada Gambar 4.3b. Jika pencilan data no. 184 pada sta. Kotabaru dibuang, maka koefisien korelasi menjadi 0,69 (Lampiran 2.4). Koefisien korelasi antara curah hujan dengan prediksinya sebesar 0,68. Dalam gambar tersebut juga diperlihatkan kemampuan model Partial Component Regression (PCR) untuk memprediksi. Koefisien korelasi antara model PCR dengan data curah hujan adalah 0,67.

Menurut Tobias (1995) aplikasi metode PLS lebih ditekankan untuk pendugaan peubah respon daripada bentuk modelnya. Ini dapat dipahami, karena model yang eksplisit tidak terlalu penting, namun ketepatan kedua kriteria berikut terpenuhi, yaitu kesesuaian model dugaan dengan seluruh data yang ada dan kesesuaian model dengan data baru (validasi).

Berdasarkan hasil kesesuaian antara model dugaan dengan seluruh data yang ada yang dicerminkan oleh nilai R^2 dan kesesuaian model dengan data baru yang dicerminkan dengan koefisien korelasi, maka model tersebut dianggap layak untuk menduga peubah respon.

SIMPULAN

Model PLS yang optimal dicapai dengan menggunakan tiga komponen pada iterasi ke-3 dengan nilai PRESS terkecil. Model PLS dapat menghilangkan masalah kolineritas dan data GCM dapat dimanfaatkan dalam peramalan iklim lokal.

SARAN

Yang perlu dilakukan berikutnya adalah melihat hubungan antara peubah curah hujan dengan peubah GCM pada periode musiman sehingga diharapkan dapat digunakan untuk memprediksi pada setiap periode kejadiannya, misalnya periode DJF (Desember, Januari dan Februari).

DAFTAR PUSTAKA

Almoy, T. 1996. A Simulation study on comparison of prediction methods when only a few componen are relevant. *Computational Statistical & data Analysis*. 21; p87-107.

- Antou, N. Konda. 2000. Kajian pencilan pada model regresi kuadrat terkecil parsial. Tesis S2, PPs-IPB. Bogor: 57p.
- Draper, NR. & Smith H.. 1992. Analisis Regresi Terapan. Ed. 2. Pen. Bambang Sumatri. Gramedia Pustaka Utama. Jakarta; 688 hal.
- Fox, John and G. Monette. 1992. Generalized Collinearity Diagnostics. *Journal of the Ameracan Statistical Association*. Vol 87, No. 417: p 178 – 183.
- Geladi, P & Kowalski, BR. 1986. Partial Least Square Regression. Tutorial. *Analyca Chimica Acta*. 185; p 1 – 17.
- Herwindiati, DE. 1997. Pengkajian Regresi Komponen Utama, Regresi Ridge dan Regresi Kuadrat Terkecil Parsial. Tesis S2, PPs-IPB. Bogor: 47p.
- Indriati, K. 1997. Pengkajian Regresi Komponen Utama, Regresi Ridge dan Regresi Kuadrat Terkecil Parsial. Tesis S2, PPs-IPB. Bogor: 43p.
- Manly, BFJ. 1986. *Multivariate Statistical Methods A PRIMER*. Chapman and Hall. NY: 159p.
- Marten, H. & Naes, T. 1989. *Multivariate Calibration*. John Wiley & Sons. NY.
- Minitab versi 14.12. 2004. *Minitab Help*. Minitab Inc.
- Nater, J, Wasserman, W & Kutner, MH. 1990. *Applied Linier Statistical Models Regression, Analysis of Variance and Experimental Design*. Richard D. Irwin Inc. Illinois. Toppan Company, LTD. Tokyo.
- Shao, J. 1993. Linear Model Selection by Cross-Validation. *JASA* 88: p486-498.
- Tobias, RD. 1995. An Introduction to Partial Least Square Regression in Proceedings of the Twentieth Annual SAS User Group International Conference, Cary, NC: SAS Institute Inc.; 1250-1257.
- Wigena & Aunuddin. 1998. Metode PLS Untuk Mengatasi Kolinearitas dalam Kalibrasi Ganda. *Forum Statistika dan Komputasi*. Vol 3, No.1; p 1 – 4. ISSN0853-8115.
- Weisberg, S. 1985. *Applied Linear Regression*. 2nd Edition. John Wiley & Sons. New York.
- Wulandari, Sri P. 2000. Analisis hubungan antara peubah ekonomi dengan kesejahteraan menggunakan metode PLS (Lartial Least Square). Tesis S2, PPs-IPB. Bogor: 43p.
- Young, PJ. 1994. A Reformation of The Partial Least Square Regression Algorithm. *SIAM J. SCL STAT Comput*. 5, 1: p 225-230.

Lampiran 1. Analisis P L S

Lampiran 1. 1. Analisis ragam untuk ch pada Sta. Kotabaru
Appendix 1.1. Analysis of variance for ch at Kotabaru BMKG Station

Source	DF	SS	MS	F	P
Regression	3	740324	246775	17.30	0.000
Residual Error	380	5419588	14262		
Total	383	6159912			

$$R^2 = 46,52\%$$

Lampiran 1. 2. Model Selection and Validation for ch pada Sta. Kotabaru
Appendix 1.2. Model selection and validation for ch at Kotabaru BMKG Station

Components	X Variance	Error SS	R-Sq	PRESS	R-Sq (pred)
1	0.459003	5702857	0.074198	5792611	0.0596275
2	0.611102	5547510	0.099417	5738135	0.0684713
3	0.669100	5419588	0.120184	5714118	0.0723702
4		5283608	0.142259	5738832	0.0683581
5		5235905	0.150003	5750479	0.0664672
6		5191832	0.157158	5745194	0.0673252
7		5155962	0.162981	5740984	0.0680087
8		5106592	0.170996	5751945	0.0662292
9		5057382	0.178985	5854075	0.0496494
10		4992445	0.189526	6047300	0.0182814
11		4961351	0.194574	5970513	0.0307469
12		4920086	0.201273	6141678	0.0029600
13		4889647	0.206215	6134112	0.0041884
14		4838691	0.214487	6267545	0.0000000
...
61		4542390	0.262588	6403388	0.0000000

Lampiran 1. 3. Koefisien regresi pada Sta. Kotabaru
 Appendix 1.3. Regression coefficient at Kotabaru BMKG Station

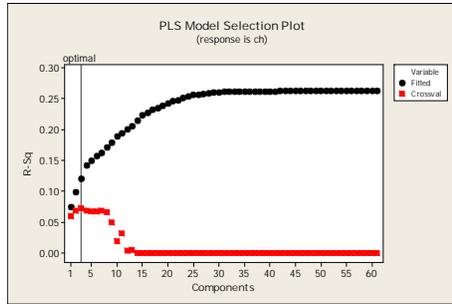
Constant	ch	ch standardized	ch	ch standardized
	199.458	0.0000000		
272.00	-0.048	-0.0286240	2528.00	-0.076
2273.00	-0.003	-0.0026225	2529.00	0.089
2274.00	0.008	0.0056683	2530.00	0.041
2275.00	0.001	0.0012746	2531.00	-0.001
2276.00	0.010	0.0072512	2532.00	0.003
2277.00	-0.004	-0.0021543	2533.00	0.011
2278.00	-0.020	-0.0059084	2534.00	-0.009
2279.00	0.081	0.0132703	2535.00	-0.030
2336.00	-0.098	-0.0496637	2592.00	0.017
2337.00	0.020	0.0119546	2593.00	0.114
2338.00	0.010	0.0068310	2594.00	0.041
2339.00	-0.002	-0.0018512	2595.00	0.005
2340.00	0.005	0.0037155	2596.00	-0.015
2341.00	0.006	0.0029656	2597.00	0.009
2342.00	-0.017	-0.0055191	2598.00	0.023
2343.00	0.034	0.0052462	2599.00	-0.011
2400.00	-0.107	-0.0581367	2656.00	0.018
2401.00	-0.048	-0.0199540	2657.00	0.157
2402.00	-0.013	-0.0061669	2658.00	0.014
2403.00	-0.009	-0.0071536	2659.00	-0.003
2404.00	0.002	0.0012105	2660.00	-0.019
2405.00	0.006	0.0031085	2661.00	0.019
2406.00	-0.012	-0.0044390	2662.00	0.057
2407.00	-0.007	-0.0012489	2663.00	0.057
2464.00	-0.073	-0.0323314	2720.00	-0.078
2465.00	-0.055	-0.0270356	2721.00	0.029
2466.00	-0.011	-0.0071160	2722.00	0.028
2467.00	-0.020	-0.0126587	2723.00	-0.005
2468.00	0.004	0.0021871	2724.00	-0.012
2469.00	0.005	0.0023952	2725.00	0.028
2470.00	-0.013	-0.0041786	2726.00	0.070
2471.00	-0.038	-0.0071413	2727.00	0.103

Lampiran 1.4. Predicted Response for New Observations Using Model for ch pada Sta. Kotabaru
 Appendix 1.4. Predicted Response for New Observations Using Model for ch at Sta. Kotabaru

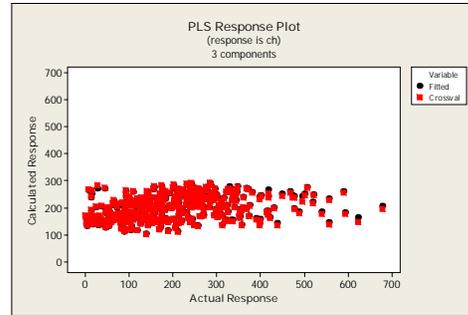
Row	Fit	SE Fit	95% CI	95% PI
1	242.591	11.8262	(219.338, 265.844)	(6.628, 478.554)
2	230.300	11.3993	(207.887, 252.714)	(-5.581, 466.182)
3	198.873	8.8023	(181.566, 216.180)	(-36.578, 434.324)
4	179.625	9.6441	(160.662, 198.587)	(-55.954, 415.204)
5	213.940	12.4933	(189.375, 238.504)	(-22.156, 450.035)
6	201.857	7.0817	(187.933, 215.781)	(-33.370, 437.084)
7	170.553	9.9273	(151.034, 190.072)	(-65.071, 406.177)
8	163.670	12.1106	(139.857, 187.482)	(-72.349, 399.688)
9	158.225	12.8297	(132.999, 183.451)	(-77.940, 394.391)
10	114.919	15.0948	(85.239, 144.599)	(-121.764, 351.602)
11	149.133	9.6088	(130.240, 168.026)	(-86.440, 384.706)
12	220.073	11.9264	(196.623, 243.523)	(-15.910, 456.055)

Lampiran 2. Analisis P L S dalam Grafik

Lampiran 2.1. PLS Model Selection Plot (a) dan PLS Response Plot (b)
 Appendix 2.1. PLS Model selection plot (a) and PLS Response Plot (b)

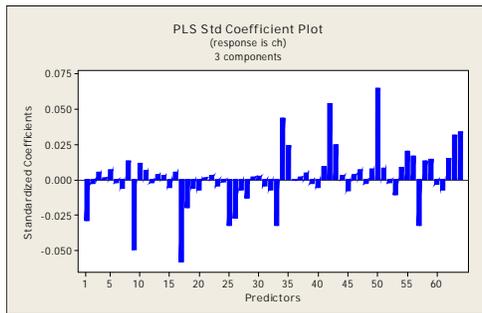


(a)

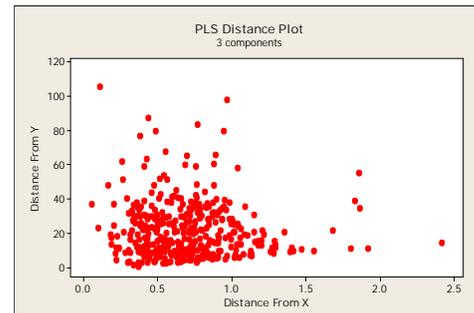


(b)

Lampiran 2.2. PLS Std Coefficient Plot (a) PLS Distance Plot (b)
 Appendix 2.2. PLS STD coefficient plot (a) and PLS Distance Plot (b)

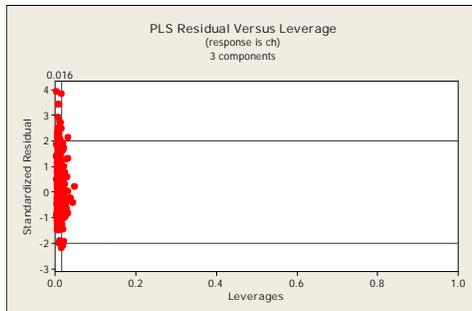


(a)

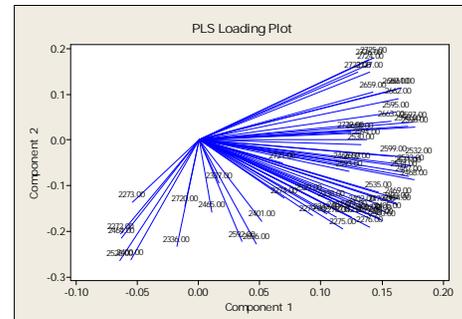


(b)

Lampiran 2.3. PLS Residual Versus Leverage PLS Loading Plot
 Appendix 2.3. Residual PLS of Versus Leverage PLS Loading Plot

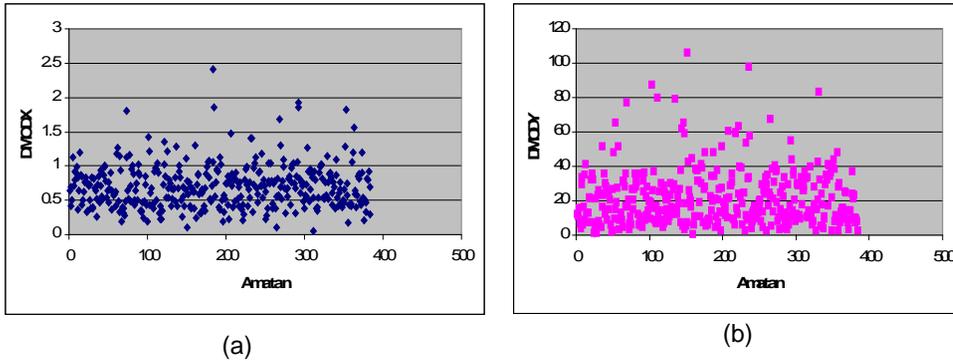


(a)



(b)

Lampiran 2.4. Plot dmodx dari 384 (Kotabaru) dan Plot dmody dari 384
 Appendix 2.4. Plot dmodx out of 384 (Kotabaru) and Plot dmody out of 384



Lampiran 2.5. Validasi prediksi curah hujan model PLS, PLS amatan pencila dibuang dan PCR dengan data observasi pada Sta. Kotabaru
 Appendix 2.5. Validation of rainfall prediction using PLS and PCR models with observation data at Kotabaru BMKG Station.

