

# Pendekatan Multivariate Adaptive Regression Spline (MARS) pada Pengelompokan Zona Musim Suatu Wilayah

BAMBANG WIDJANARKO OTOK

Jurusan Statistika Institut Teknologi Sepuluh Nopember, Surabaya  
Email: bambang\_wo@statistika.its.ac.id; otok\_bw@yahoo.com

## ABSTRAK

Multivariate adaptive regression spline (MARS) adalah salah satu model regresi nonparametrik, yaitu model yang mengasumsikan bentuk hubungan fungsional antara variabel respon dan prediktor tidak diketahui. MARS adalah kombinasi yang kompleks antara metode spline dengan rekursif partisi untuk menghasilkan estimasi fungsi regresi yang kontinu, dan digunakan untuk prediksi dan klasifikasi. Hasil penelitian menunjukkan bahwa estimasi parameter model MARS untuk variabel respon kontinu dilakukan dengan *penalized least square* (PLS). Pemilihan model MARS terbaik dilakukan dengan prosedur *forward* dan *backward stepwise* didasarkan pada nilai GCV. Prosedur *forward* adalah tahapan untuk mendapatkan fungsi basis maksimum yang mencakup pengaruh efek utama, interaksi, dan *knot*. Sedangkan prosedur *backward* adalah tahapan untuk mengeliminasi fungsi basis yang kontribusinya tidak signifikan. Hasil kajian juga menunjukkan bahwa GCV dengan potongan regresi linear dapat terbukti bekerja dengan baik dalam menentukan pemilihan model terbaik pada MARS respon kontinu.

Kata-kata Kunci: MARS, *Penalized Least Square*, GCV.

## 1. PENDAHULUAN

Klasifikasi mempunyai dua arti yang berbeda. Jika diberikan satu set pengamatan dengan tujuan menetapkan keberadaan kelas atau kelompok dalam data tanpa diketahui pengelompokan dengan jelas, maka prosedur klasifikasi yang digunakan dikenal dengan *Unsupervised Learning* (*Clustering*). Sebaliknya jika untuk menetapkan aturan dengan mengolompokkan pengamatan baru ke dalam kelas yang ada dikenal dengan *Supervised Learning*. (Michie, dkk., 1994). Banyak penelitian klimatologi yang bertujuan membuat kelompok stasiun wilayah perkiraan iklim. Penggunaan data kelompok stasiun lebih menguntungkan daripada data masing-masing individu stasiun. Salah satu keuntungan data kelompok stasiun adalah data lebih homogen. Data pengamatan stasiun seringkali tidak homogen, karena adanya penggantian alat pengamatan (*instrumentasi*), perpindahan lokasi stasiun pengamatan, dan pengaruh urbanisasi (Bunkers *et al.* 1996). Metode penggabungan yang seringkali digunakan adalah *complete linkage*, *average linkage*, *Ward's* (Bunkers, dkk., 1996); *complete linkage* (BMG 2003); *Ward's* dan *Centroid* (Wigena 2006). Bunkers, dkk., (1996) menyimpulkan bahwa *average linkage* mempunyai kinerja yang baik. Sementara itu Gong dan Richman (1995) menyimpulkan metode *Ward's* mempunyai kinerja yang baik diantara metode-metode hierarkhi lainnya. Hal ini juga didukung oleh Otok (2007), bahwa metode *Ward's* memberikan hasil ketepatan yang lebih baik dibanding dengan beberapa metode hierarkhi lainnya pada pengelompokan wilayah. Pada beberapa dekade terakhir ini, ada beberapa teknik pemodelan berbasis *machine learning* yang telah dikembangkan untuk permasalahan pendekatan suatu fungsi. Teknik pemodelan tersebut antara lain *Feed-forward Neural Network* (FFNN), *Radial Basis Function Network* (RBFN), *General Regression Neural Network* (GRNN), *Support Vector Machine* (SVM), dan pengembangan basis spline yang lain, *Multivariate Adaptive Regression Spline* (MARS). Otok, dkk (2008, 2005), mengembangkan model berbasis *machine learning*, seperti MARS dan FFNN. Otok (2008), mengkaji secara inferensi fungsi basis pada model MARS.

*Multivariate adaptive regression spline* (MARS) adalah salah satu model regresi nonparametrik, yaitu model yang mengasumsikan bentuk hubungan fungsional antara variabel respon dan prediktor tidak diketahui, dan mempunyai bentuk fungsional yang fleksibel. MARS adalah

metode yang diperkenalkan oleh Friedman (1991). Metode ini adalah implementasi teknik-teknik untuk menyelesaikan masalah regresi dan klasifikasi, dengan tujuan untuk memprediksi variabel respon yang bernilai kontinu berdasar beberapa variabel prediktor. Model MARS disusun pada pengaturan beberapa koefisien fungsi basis dimana secara keseluruhan dikendalikan pada data regresi. Hastie dkk. (2002) menyatakan bahwa model MARS berguna untuk mengatasi permasalahan data dimensi tinggi yang dikenal dengan *curse of dimensionality* dan menghasilkan prediksi variabel respon yang akurat, serta untuk mengatasi kelemahan regresi partisi rekursif (RPR) yaitu menghasilkan model yang kontinu pada knot, yang didasarkan pada nilai *generalized cross validation* (GCV) minimum. Oleh karena itu akan mengkaji Zona Musim (ZOM) BMKG dengan menggunakan pendekatan pengelompokan parametrik dan nonparametrik. Lebih khusus, akan dikaji penggunaan analisis kelompok berhierarkhi (*complete linkage*) pada data curah hujan. Selanjutnya dengan pendekatan MARS dapat diketahui ketepatan pengelompokan tersebut. Dengan demikian, pengembangan model yang dibangun dapat digunakan sebagai *early warning system* yang baik terhadap pihak-pihak terkait dan dapat mengantisipasi tindakan-tindakan yang harus segera dilaksanakan untuk memperbaiki kondisi ekstrim

## 2. TINJAUAN PUSTAKA

### Analisis Kelompok

Analisis kelompok digunakan untuk mengelompokkan obyek-obyek pengamatan menjadi beberapa kelompok berdasarkan pengukuran variabel-variabel yang di amati, sehingga obyek dalam kelompok yang sama mirip dan antar kelompok tidak mirip. Analisis kelompok adalah suatu teknik yang digunakan untuk mengkombinasikan pengamatan ke dalam kelompok (cluster) dengan masing-masing kelompok (cluster) adalah homogen berkenaan dengan karakteristik tertentu. Ini menunjukkan bahwa pengamatan pada setiap kelompok adalah serupa untuk satu sama lain. Masing-Masing kelompok harus berbeda dari kelompok lainnya berkenaan dengan karakteristik yang sama; yang berarti pengamatan satu kelompok harus berbeda dari pengamatan lain kelompok. Tahapan pengelompokkan bila dilakukan dapat disajikan dalam bentuk diagram pohon yang memungkinkan penelusuran pengelompokkan obyek-obyek yang diamati dengan lebih mudah dan informatif. Dalam teknik penggabungan, mula-mula tiap objek merupakan satu kelompok tersendiri. Lalu dua kelompok yang terdekat digabungkan dan seterusnya sehingga diperoleh satu kelompok yang berunsurkan semua objek. Untuk menggabungkan dua kelompok pada tahap awal, dimana tiap kelompok hanya terdiri atas satu objek, diperlukan ukuran ketakmiripan antar objek. Bila suatu kelompok merupakan penggabungan dari beberapa kelompok sebelumnya, maka diperlukan ukuran ketakmiripan antar kelompok, kelompok-kelompok dengan ukuran ketakmiripan terkecil yang nantinya akan digabungkan menjadi kelompok yang baru. Makin besar nilai ukuran ketakmiripan antara dua objek, makin besar pula perbedaan antara kedua objek tersebut sehingga makin cenderung untuk tidak mengangapnya ada dalam kelompok yang sama, atau makin cenderung untuk tidak menggabungkannya dalam satu kelompok yang sama. Andaikan  $d_{ij}$  merupakan ukuran ketakmiripan antara kelompok ke- $i$  dengan kelompok ke- $j$  dan  $d_{k(i,j)}$  merupakan ukuran ketakmiripan antara kelompok ke- $k$  dengan kelompok  $(i,j)$  yang merupakan penggabungan antara kelompok ke- $i$  dengan kelompok ke- $j$ , maka beberapa teknik pengelompokkan antara kelompok menggunakan pautan lengkap (Complete Linkage/Farthest Neighbor) dan ukuran jarak yang digunakan adalah :  $d_{k(i,j)} = \max(d_{ki}, d_{kj})$

### Multivariate Adaptive Regression Spline (MARS)

Friedman (1991) menyatakan bahwa model MARS adalah kombinasi yang kompleks antara metode spline dengan rekursif partisi untuk menghasilkan estimasi fungsi regresi yang kontinu. Spline adalah salah satu jenis potongan polinomial, yaitu polinomial yang memiliki sifat tersegmen. Sifat tersegmen ini memberikan fleksibilitas lebih daripada polinomial biasa, sehingga memungkinkan untuk menyesuaikan diri secara lebih efektif terhadap karakteristik lokal pada suatu fungsi atau data. Wahba (1990) menunjukkan bahwa spline memiliki sifat-sifat statistik yang berguna untuk menganalisis hubungan dalam regresi. Spline dalam regresi nonparametrik terus berkembang sampai pada model adaptive (Bilier dan Fahrmeir, 2001) dan multivariat respon (Holmes dan Mallick, 2003). Selain itu, He dan Shi (1998) mengembangkan pendekatan monotonicity untuk mengestimasi fungsi basis spline, sedangkan Hall dan

Opsomer (2005) menggunakan pendekatan penalti kuadrat terkecil. Regresi partisi rekursif (RPR) merupakan pendekatan fungsi yang tidak diketahui dengan menggunakan pengembangan fungsi basis. Selain itu, RPR adalah suatu konsep geometri yang membagi daerah dengan dasar aritmetik, yaitu penjumlahan dan perkalian. Morgan dan Sonquist (1963) memperkenalkan RPR dalam riset *automatic interaction detection* (AID), selanjutnya Morgan dan Messenger (1973) menemukan *Theta AID* (THAID), yang digunakan untuk memproduksi pohon-pohon klasifikasi. Venables dan Ripley (1994) menggunakan aturan pengklasifikasi pada metode THAID untuk memprediksi suatu obyek. Pertimbangkan  $y$  suatu variabel random berdistribusi Bernoulli,  $y \sim Ber(1, \pi(\mathbf{x}))$  dengan  $y \in \{0, 1\}$  dan  $\mathbf{x} \in \mathcal{R}^p$ , maka  $P(Y_i = 1) = \pi(\mathbf{x})$  dan  $P(Y_i = 0) = 1 - \pi(\mathbf{x})$ .  $\mathbf{x} \in \mathcal{R}^p$  adalah vektor dari  $p$  variabel prediktor dan  $P(Y = 1 | \mathbf{x}) = \pi(\mathbf{x})$ . Misalkan model ditulis sebagai,  $y = \pi(\mathbf{x}) + r$  maka  $r = y - \pi(\mathbf{x})$  sehingga,

$$\begin{aligned} E(r) &= E(y - \pi(\mathbf{x})) \\ &= E(y) - E(\pi(\mathbf{x})) = \pi(\mathbf{x}) - \pi(\mathbf{x}) = 0 \end{aligned} \quad (1)$$

$$Var(r) = Var(y - \pi(\mathbf{x})) = \pi(\mathbf{x})(1 - \pi(\mathbf{x})) \quad (2)$$

**Lemma 1: (Otok, 2008)** Jika hubungan dengan model logistik,  $\sigma_L: \mathcal{R} \rightarrow (0, 1)$ ,  $\sigma_L = \pi(x) = [e^z / (1 + e^z)]$  maka inver dari dari  $\sigma_L$  dapat dikatakan sebagai transformasi logit, logit  $\pi(x) = \ln[\pi(x) / (1 - \pi(x))] = z$

Selanjutnya jika Lemma 1 terpenuhi dan  $z = \hat{f}(x)$ , yaitu

$$z = \hat{f}(x) = \beta_0 + \sum_{m=1}^M \beta_m \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})]$$

maka dapat ditulis dalam model,

$$\text{logit } \pi(x) = \ln \left( \frac{\pi(x)}{1 - \pi(x)} \right) = \beta_0 + \sum_{m=1}^M \beta_m \prod_{k=1}^{K_m} [s_{km} \cdot (x_{v(k,m)} - t_{km})]$$

dan dalam bentuk matriks,

$$\text{logit } \pi(x) = \mathbf{B}\beta \quad (3)$$

dengan,

$$\beta = (\beta_0, \beta_1, \dots, \beta_M)^T$$

$$\mathbf{B} = \begin{pmatrix} 1 & \prod_{k=1}^{K_1} s_{1m}(x_{1(1,m)} - t_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(x_{1(M,m)} - t_{Mm}) \\ 1 & \prod_{k=1}^{K_1} s_{1m}(x_{2(1,m)} - t_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(x_{2(M,m)} - t_{Mm}) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \prod_{k=1}^{K_1} s_{1m}(x_{n(1,m)} - t_{1m}) & \cdots & \prod_{k=1}^{K_M} s_{Mm}(x_{n(M,m)} - t_{Mm}) \end{pmatrix}$$

Persamaan (3) dikatakan sebagai model MARS respon biner, dan dalam bentuk fungsi probabilitas dapat dinyatakan sebagai,

$$\pi(x) = \sigma_L(\mathbf{B}^T \beta)$$

Untuk mencari estimasi  $\beta$  digunakan metode maksimum likelihood.

### Estimator MARS Classification

Misalkan variabel random  $Y$ ,  $y \sim Ber(1, \sigma_L(\mathbf{B}_i \beta))$  dan dinyatakan dalam model,  $y = \sigma_L(\mathbf{B}\beta) + r$  maka  $r = y - \sigma_L(\mathbf{B}\beta)$ . Jika fungsi  $\sigma_L$  diterapkan pada vektor  $z = (z_1, \dots, z_n)^T \in \mathcal{R}^n$  mempunyai interpretasi sebagai  $\sigma_L(z) = (\sigma_L(z_1), \dots, \sigma_L(z_n))^T$ . Analog Persamaan (1) dan (2) diperoleh,

$E(r) = 0$ , dan

$$\begin{aligned} s_i^2 &= \text{Var}(r_i) = \text{Var}(y - \sigma_L(\mathbf{B}_i^T \beta)) = \sigma_L(\mathbf{B}_i^T \beta)(1 - \sigma_L(\mathbf{B}_i^T \beta)) \\ &= \frac{e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}} \left[ 1 - \frac{e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}} \right] = \frac{e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}} \left[ \frac{1 + e^{\mathbf{B}_i^T \beta} - e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}} \right] \\ &= \frac{e^{\mathbf{B}_i^T \beta}}{(1 + e^{\mathbf{B}_i^T \beta})^2} = \sigma'_L(\mathbf{B}_i^T \beta) \end{aligned}$$

Sehingga fungsi likelihood dapat dinyatakan sebagai

$$\begin{aligned} L(\beta) &= \prod_{i=1}^n P(\mathbf{y} = y_i) \\ &= \prod_{i=1}^n \binom{1}{y_i} [\sigma_L(\mathbf{B}_i^T \beta)]^{y_i} [(1 - \sigma_L(\mathbf{B}_i^T \beta))]^{1 - y_i} \\ &= \prod_{i=1}^n \frac{1}{y_i! (1 - y_i)!} [\sigma_L(\mathbf{B}_i^T \beta)]^{y_i} [(1 - \sigma_L(\mathbf{B}_i^T \beta))]^{1 - y_i} \\ &= \prod_{i=1}^n \left( \frac{\sigma_L(\mathbf{B}_i^T \beta)}{(1 - \sigma_L(\mathbf{B}_i^T \beta))} \right)^{y_i} (1 - \sigma_L(\mathbf{B}_i^T \beta)) \\ &= \prod_{i=1}^n \left( \frac{\frac{e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}}}{\frac{1 + e^{\mathbf{B}_i^T \beta} - e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}}} \right)^{y_i} \left( \frac{1 + e^{\mathbf{B}_i^T \beta} - e^{\mathbf{B}_i^T \beta}}{1 + e^{\mathbf{B}_i^T \beta}} \right) \\ &= \prod_{i=1}^n (e^{\mathbf{B}_i^T \beta})^{y_i} \left( \frac{1}{1 + e^{\mathbf{B}_i^T \beta}} \right) \\ &= \prod_{i=1}^n \frac{e^{\mathbf{B}_i^T \beta y_i}}{1 + e^{\mathbf{B}_i^T \beta}} \end{aligned} \tag{4}$$

dan fungsi log-likelihoodnya adalah

$$\begin{aligned} \ln L(\beta) &= \ln \left( \prod_{i=1}^n \frac{e^{\mathbf{B}_i^T \beta y_i}}{1 + e^{\mathbf{B}_i^T \beta}} \right) \\ &= \sum_{i=1}^n \ln(\exp(\mathbf{B}_i^T \beta y_i)) - \sum_{i=1}^n \ln(1 + \exp(\mathbf{B}_i^T \beta)) \\ &= \sum_{i=1}^n \mathbf{B}_i^T \beta y_i - \sum_{i=1}^n \ln(1 + \exp(\mathbf{B}_i^T \beta)) \\ &= \mathbf{y}^T \mathbf{B} \beta - \sum_{i=1}^n \ln(1 + \exp(\mathbf{B}_i^T \beta)) \end{aligned} \tag{5}$$

Untuk menunjukkan fungsi maksimum, diturunkan fungsi log-likelihood terhadap  $\beta$ , yaitu sebagai berikut:

$$\begin{aligned} \frac{\partial}{\partial \beta} \ln L(\beta) &= \frac{\partial}{\partial \beta} \left( \mathbf{y}^T \mathbf{B} \beta - \sum_{i=1}^n \ln(1 + \exp(\mathbf{B}_i^T \beta)) \right) \\ &= \mathbf{B}^T \mathbf{y} - \sum_{i=1}^n \frac{\exp(\mathbf{B}_i^T \beta)}{(1 + \exp(\mathbf{B}_i^T \beta))} \mathbf{B}_i \\ &= \mathbf{B}^T \mathbf{y} - \mathbf{B}^T \sigma_L(\mathbf{B} \beta) \\ &= \mathbf{B}^T [\mathbf{y} - \sigma_L(\mathbf{B} \beta)] \end{aligned}$$

Jika turunan pertama disamadengankan nol, maka  $\hat{\beta}$  adalah estimator maksimum likelihood dari  $\beta$ , yaitu:

$$\mathbf{B}^T [\mathbf{y} - \sigma_L(\mathbf{B}\hat{\beta})] = 0 \quad (6)$$

Untuk membuktikan  $\hat{\beta}$  adalah estimator yang memaksimumkan Persamaan (4) dicari turunan kedua dari fungsi log-likelihood, yaitu:

$$\begin{aligned} \frac{\partial^2}{\partial \beta_s \partial \beta_r} \ln L(\beta) &= \frac{\partial}{\partial \beta_s} (\mathbf{B}^T [\mathbf{y} - \sigma_L(\mathbf{B}\beta)]) \\ &= \frac{\partial}{\partial \beta_s} \left( \sum_{i=1}^n \mathbf{B}_{ir} \mathbf{y}_i - \sum_{i=1}^n \mathbf{B}_{ir} \sigma_L(\mathbf{B}_i^T \beta) \right) \\ &= - \sum_{i=1}^n \mathbf{B}_{ir} \frac{\partial}{\partial \beta_s} \sigma_L(\mathbf{B}_i^T \beta) \\ &= - \sum_{i=1}^n \mathbf{B}_{ir} \mathbf{B}_{is} \sigma_L'(\mathbf{B}_i^T \beta) \\ &= -\mathbf{B}^T D(\beta) \mathbf{B} \end{aligned}$$

dan diperoleh matriks Hessian,  $H_{\ln L}(\beta) \in \mathfrak{R}^p$ , yaitu:

$$H_{\ln L}(\beta) = -\mathbf{B}^T D(\beta) \mathbf{B}$$

dimana  $D(\beta) = d_{ij} \in \mathfrak{R}^n$  matrik diagonal yang didefinisikan sebagai,

$$d_{ij} = \begin{cases} \sigma_L'(\mathbf{B}_i^T \beta) & , \text{jika } i = j \\ 0 & , \text{jika } i \neq j \end{cases} \quad (7)$$

dan  $H_{\ln L}(\beta)$  semi definit negatif untuk setiap  $\beta \in \mathfrak{R}^p$ . Sehingga diperoleh:

$$\mathbf{u}^T H_{\ln L}(\beta) \mathbf{u} = -\mathbf{u}^T \mathbf{B}^T D(\beta) \mathbf{B} \mathbf{u} = - \sum_{i=1}^n (\mathbf{B}_i^T \mathbf{u})^2 \sigma_L'(\mathbf{B}_i^T \beta) \quad (8)$$

Jadi turunan pertama selalu positif, sedangkan dari Persamaan (8) bahwa  $\mathbf{u}^T H_{\ln L}(\beta) \mathbf{u} \leq 0$  untuk semua  $\mathbf{u} \in \mathfrak{R}^p$  dan  $\beta \in \mathfrak{R}^p$ . Jadi dapat dikatakan bahwa  $\hat{\beta}$  adalah estimator yang memaksimumkan.  $\square$

### 3. HASIL EMPIRIS

Analisis pengelompokan yang digunakan adalah analisis kluster hierarki berdasarkan data asli dan data reduksi dengan analisis faktor. Pendekatan *machine learning* (MARS *Clasification*) juga digunakan untuk melihat ketepatan pengelompokan ZOM suatu wilayah dari hasil analisis pengelompokan.

#### Deskripsi Variabel Penelitian

Pada bab ini akan dijelaskan proses awal sebelum dilakukan analisis lebih lanjut, yaitu menganalisis karakteristik curah hujan dari masing-masing stasiun curah hujan di Kabupaten Ngawi. Terdapat 18 stasiun yang akan dianalisis mengenai karakteristik curah hujan. Curah hujan yang diamati adalah curah hujan dari bulan Januari hingga Desember, dimulai tahun 1989 hingga tahun 2010. Gambaran umum mengenai curah hujan di Kabupaten Ngawi pada 18 stasiun curah hujan dapat dilihat pada Tabel 1.

**Tabel 1.** Deskriptif Curah Hujan di Kabupaten Ngawi Tahun 1989-2010

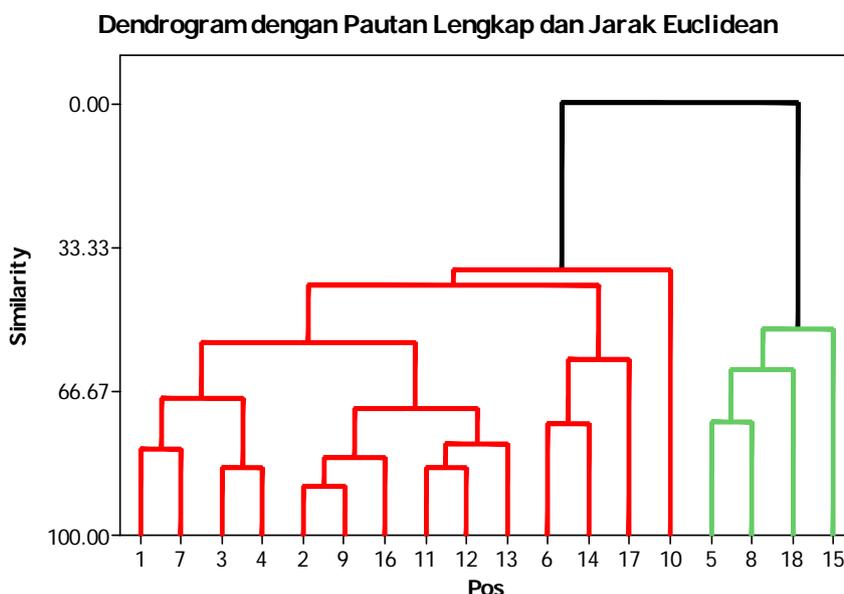
POS	Minimum	Maximum	Mean	Variance
MANTINGAN	32,09	312,86	168,7500	12457,350
NGAWI	20,31	318,66	159,2812	12416,501
NGRAMBE	10,86	330,10	150,6447	13176,275
TRETES	13,73	333,00	147,6390	12390,947
KEDUNG URUNG	20,33	372,27	179,1247	18705,517
KEDUNG GALAR	12,00	255,30	134,3504	<b>8042,355</b>
WALIKUKUN	23,05	349,41	163,8932	13306,074
JOGOROGO	20,36	412,76	197,3838	21422,924
BEKOH	12,70	305,35	159,0253	13027,823
KEDUNGBENDO	<b>5,33</b>	359,44	135,0370	15839,130
NGALE	17,36	296,50	160,3925	11593,535
PARON	13,86	307,14	163,0850	12154,109
MARDIASRI	16,05	317,15	175,0558	13761,222
PADAS	7,29	263,67	148,0588	11072,250
KARANGJATI	22,00	<b>451,64</b>	<b>205,5709</b>	<b>26133,683</b>
SAMBIROTO	9,05	337,32	156,8926	14054,697
BEGAL	18,87	293,47	<b>126,3271</b>	8472,097
GUYUNG	12,59	370,29	175,8186	17386,599

Sumber: Data diolah

Tabel 1 memperlihatkan bahwa rata-rata curah hujan tertinggi adalah pada stasiun Karangjati Karangjati dengan rata-rata curah hujan yang terjadi adalah 205,57 mml. Pada stasiun Karangjati juga memiliki varian terbesar, yaitu 26.133,683 mml, hal ini menunjukkan bahwa keragaman curah hujan pada stasiun tersebut sangat tinggi, dengan curah hujan minimal 22 mml dan curah hujan maksimal sebesar 451,64 mml. Tingginya nilai variasi curah hujan pada stasiun Karangjati mengindikasikan bahwa di daerah sekitar stasiun Karangjati memiliki curah hujan yang ekstrim, dengan kata lain perbedaan tingkat curah hujan pada musim kemarau dan musim penghujan sangat tinggi. Stasiun Begal memiliki curah hujan terendah, dimana rata-rata curah hujan di daerah sekitar stasiun Begal sebesar 126,33 mml, dengan curah hujan paling banyak sebesar 293,47 mml dan curah hujan paling sedikit sebesar 18,87 mml. Stasiun yang memiliki keragaman curah hujan terendah adalah di stasiun Kedung Galar, sehingga pada daerah sekitar stasiun Kedung Galar tidak memiliki perbedaan curah hujan yang mencolok ketika musim penghujan dan musim kemarau. Dari 18 stasiun curah hujan, daerah sekitar Kedung Bendo yang memiliki curah hujan paling rendah dan daerah sekitar Karangjati yang memiliki curah hujan tertinggi.

#### Membangun Zona Musim Pada Data Asli (Tanpa Reduksi)

Data curah hujan yang diperoleh dilakukan penyederhanaan sehingga terbentuk data curah hujan 18 pos dengan masing-masing pos memiliki nilai rata-rata curah hujan tiap bulannya. Data ini dilakukan analisis klaster dengan metode pautan lengkap dan jarak Euclidean. Dengan metode tersebut, dibentuk 2 kelompok zona pos, yaitu zona curah hujan tinggi dan zona curah hujan rendah. Hasil pengelompokan ditunjukkan dengan dendogram pada Gambar 1. Data curah hujan yang telah disederhanakan dan hasil pengelompokannya diringkas dalam Tabel 2.



**Gambar 1.** Dendrogram data asli dengan pautan lengkap dan jarak Euclidean

Gambar 1. menunjukkan bahwa Zona Curah Hujan rendah terdiri dari 14 pos, antara lain Pos Mantingan, Ngawi, Ngrambe, Tretes, Kedunggalar, Walikukun, Bekoh, Kedung Bendo, Ngale, Paron, Mardiasri, Padas, Sambiroto, dan Begal. Zona Curah Hujan tinggi terdiri dari 4 pos lainnya, yaitu Pos Kedung Urung-urung, Jogorogo, Karangjati, dan Guyung. Ketepatan kedua kelompok ini tidak dapat diukur, sehingga akan dilanjutkan dengan metode klasifikasi.

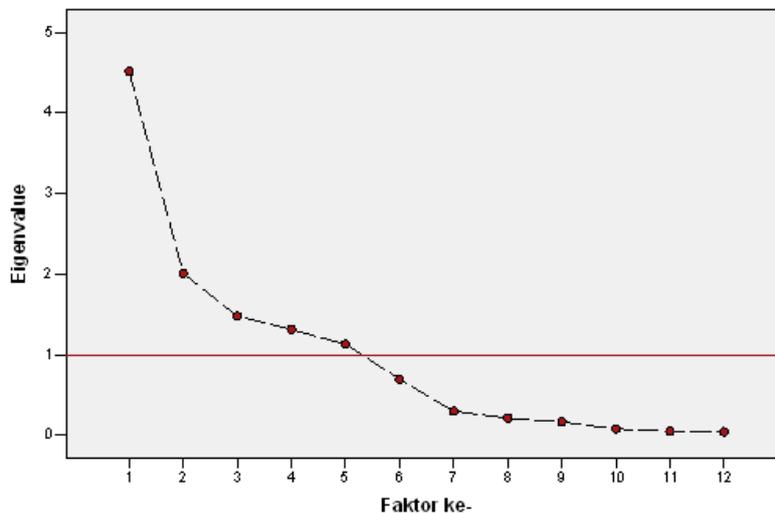
**Tabel 2.** Data Curah Hujan dan Pengelompokan dengan Pautan Lengkap

No	Pos	Jan	Feb	Mar	Apr	Mei	Jun	Jul	Ags	Sep	Okt	Nov	Des	Kelompok Curah Hujan
1	Mantingan	307.45	312.86	298.00	203.91	135.19	49.77	32.09	32.23	41.50	150.05	197.81	264.14	Rendah
2	Ngawi	275.09	267.69	318.66	225.40	100.26	56.09	27.75	20.31	27.03	127.09	205.56	260.44	Rendah
3	Ngrambe	330.10	306.30	246.62	190.98	85.95	52.36	17.23	10.86	32.90	110.00	171.05	253.38	Rendah
4	Tretes	333.00	303.32	241.10	193.11	68.36	75.45	19.50	13.73	19.41	125.14	173.10	206.45	Rendah
5	Kedung Urung	358.73	372.27	346.67	194.80	125.64	65.77	20.33	23.90	27.57	106.86	210.95	296.00	Tinggi
6	Kedunggalar	220.00	255.30	236.25	173.35	89.60	67.20	25.40	12.00	32.05	95.95	203.58	201.53	Rendah
7	Walikukun	349.41	311.86	281.90	198.36	97.00	69.38	25.86	23.05	30.36	157.57	197.95	224.00	Rendah
8	Jogorogo	385.57	412.76	370.41	263.48	118.50	100.91	44.14	20.36	34.14	104.95	197.67	315.71	Tinggi
9	Bekoh	286.16	279.95	305.35	236.05	93.10	53.50	41.35	12.70	18.79	108.50	225.58	247.28	Rendah
10	Kedung Bendo	222.11	359.44	326.44	174.89	113.00	26.78	10.44	5.33	20.89	25.67	111.22	224.22	Rendah
11	Ngale	273.00	296.50	255.57	230.86	120.36	62.82	33.91	17.36	30.86	108.41	252.81	242.24	Rendah
12	Paron	236.62	307.14	294.55	224.77	134.38	55.33	36.05	13.86	37.10	108.43	248.45	260.35	Rendah
13	Mardiasri	301.16	296.10	317.15	232.60	150.45	62.00	33.65	16.05	37.95	120.80	277.11	255.65	Rendah
14	Padas	245.76	263.67	256.59	247.88	83.71	62.35	32.29	14.18	7.29	111.33	248.94	202.71	Rendah
15	Karangjati	401.91	451.64	371.50	325.86	133.50	51.05	25.55	22.00	28.55	105.18	282.71	267.41	Tinggi
16	Sambiroto	245.11	313.56	337.32	229.83	99.53	50.21	25.79	9.05	23.00	100.85	207.67	240.81	Rendah
17	Begal	213.33	234.53	293.47	166.93	73.80	35.60	34.80	44.00	18.87	71.47	145.81	183.31	Rendah
18	Guyung	319.35	323.94	370.29	303.00	104.53	72.88	34.06	12.59	21.71	111.18	198.35	237.94	Tinggi

### Membangun Zona Musim Pada Data Reduksi

Pendekatan lainnya adalah dengan melakukan reduksi variabel terlebih dahulu, yaitu dengan melakukan analisis faktor. Pengujian asumsi kecukupan data dan adanya multikolinieritas dilakukan sebelum melakukan analisis reduksi variabel. Berdasarkan nilai Kaiser-Meyer-Olkin Measure of Sampling Adequacy (KMO MSA) sebesar 0,544, yaitu lebih dari 0,5, dapat disimpulkan bahwa data sudah cukup untuk dilakukan analisis faktor. Berdasarkan hasil uji *Bartlett's test of Sphericity*, dimana *p-value* yang dihasilkan sebesar 0, dapat disimpulkan bahwa terdapat multikolinieritas dalam data sehingga dapat dilanjutkan dengan analisis faktor.

Banyaknya faktor dapat ditentukan dengan melihat *Scree Plot*, yaitu memilih faktor yang *eigenvalue*-nya bernilai lebih besar daripada 1. Berdasarkan *Scree Plot* pada Gambar 4.2, dapat diketahui bahwa ada 5 faktor yang memiliki *eigenvalue* yang bernilai lebih besar daripada 1 dan dapat dibentuk. Namun, dalam penelitian ini hanya dibentuk 4 faktor, yang menunjukkan faktor musim, yaitu musim kemarau dan musim hujan, dan bulan pergantian musim. Rotasi VARIMAX diterapkan dalam analisis faktor dalam perhitungan skor faktornya.



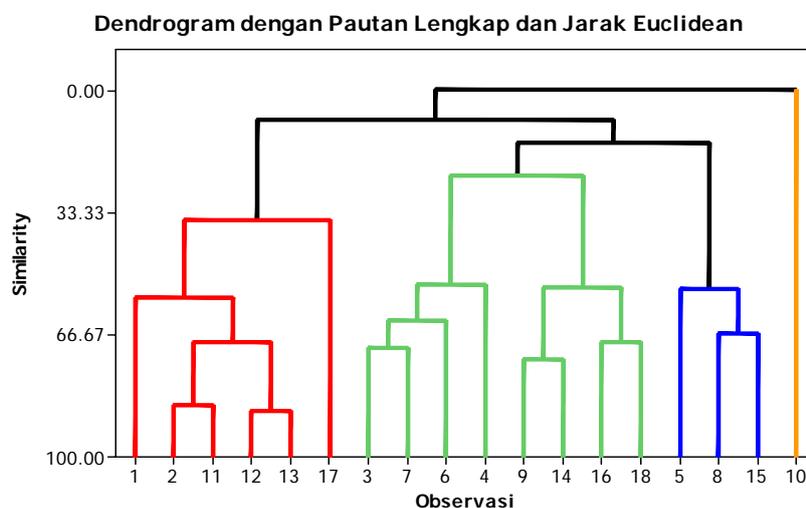
**Gambar 3..** *Scree plot* analisis faktor

Asumsi dependensi dan syarat kecukupan data telah terpenuhi, maka dari itu analisis faktor dapat dilakukan. Hasil analisis faktor adalah seperti terlihat pada Tabel 3.

**Tabel 3.** *Loading* Faktor dengan Rotasi Varimax

Bulan	<i>Loading</i>			
	Faktor 1	Faktor 2	Faktor 3	Faktor 4
JANUARI	<b>0,635</b>	0,697	0,149	-0,055
FEBRUARI	<b>0,904</b>	0,141	0,214	-0,097
MARET	<b>0,889</b>	-0,158	0,053	0,217
APRIL	0,607	0,192	-0,102	<b>0,673</b>
MEI	0,424	-0,136	<b>0,797</b>	0,329
JULI	-0,005	0,168	-0,004	<b>0,791</b>
AGUSTUS	-0,007	0,037	<b>0,072</b>	0,006
SEPTEMBER	0,038	0,206	<b>0,918</b>	-0,035
OKTOBER	-0,209	<b>0,782</b>	0,270	0,260
NOVEMBER	0,046	0,233	0,323	<b>0,798</b>
DESEMBER	<b>0,655</b>	0,280	0,540	0,125

Analisis faktor digunakan untuk mereduksi dimensi sedemikian hingga sejumlah faktor tersebut mampu menjelaskan sebesar mungkin keragaman data yang dijelaskan oleh variabel asal. Analisis faktor dengan rotasi Varimax untuk data curah hujan menghasilkan 5 Faktor dari 12 variabel, namun dari 5 Faktor yang diperoleh hanya 4 Faktor yang akan dianalisis. Pertimbangan pengambilan 4 Faktor untuk di analisis terkait dengan iklim di Indonesia, khususnya di Kabupaten Ngawi yang hanya terdiri dari 2 musim, yaitu musim hujan dan musim kemarau. Berdasarkan acuan kondisi iklim tersebut, maka Faktor 1 merupakan musim penghujan dan faktor 4 adalah musim kemarau, sedangkan Faktor 2 dan 3 adalah musim transisi. Sesuai dengan iklim di Indonesia, dimana ketika bulan April hingga September merupakan musim kemarau, bulan yang termasuk dalam Faktor 1 (musim penghujan) adalah bulan Januari, Februari, Maret dan Desember. Variabel yang termasuk dalam faktor 4 (musim kemarau) adalah bulan April, Juli dan November dengan puncak musim kemarau adalah bulan November yang merupakan puncak musim hujan. Faktor 2 dan 3 merupakan musim transisi yang terjadi pada bulan Mei, Juni, Agustus, September dan Oktober. Berdasarkan tabel 4.2 dapat dijelaskan bahwa musim penghujan di Kabupaten Ngawi terjadi selama 4 bulan, musim kemarau terjadi selama 3 bulan dan musim transisi terjadi selama 5 bulan. Hal tersebut mengindikasikan bahwa musim transisi dan musim hujan terjadi lebih lama daripada musim kemarau. Total keragaman yang dapat dijelaskan oleh Faktor 1 adalah sebesar 25,5%, Faktor 2 dapat menjelaskan total keragaman sebesar 42,9%, sedangkan total keragaman untuk faktor 3 dan 4 adalah 59,8% dan 76,6%, sehingga empat faktor yang terbentuk dapat menjelaskan total keragaman sebesar 76,6%. Hasil pengelompokan observasi dengan analisis kluster dengan pautan lengkap dan jarak Euclidean pada data reduksi ditampilkan dengan dendrogram pada Gambar 4.3. Berdasarkan Gambar 4.3 ini, dapat diketahui bahwa kelompok pertama terdiri dari Pos Mantingan, Ngawi, Ngale, Paron, Mardiasri, dan Begal. Kelompok kedua terdiri dari Pos Ngrambe, Tretes, Kedunggal, Walikukun, Bekoh, Padas, Sambiroto, dan Guyung. Kelompok ketiga terdiri dari Kedung Urung-urung, Jogorogo, dan KarangJati. Sedangkan kelompok keempat hanya Pos Kedung Bendo. Tabel 4 merupakan hasil ringkasan pengelompokan data reduksi dengan skor faktor masing-masing pos.



**Gambar 4.** Dendrogram data reduksi dengan pautan lengkap dan jarak Euclidean

Tabel 4. Skor Faktor Data Curah Hujan Hasil Reduksi

No	Pos	Faktor 1	Faktor 2	Faktor 3	Faktor 4	Kelompok
1	Mantingan	-0.07216	0.11044	-0.09869	2.27458	1
2	Ngawi	-0.27284	0.02410	0.15942	0.23481	1
3	Ngrambe	-0.26021	0.71870	-1.69302	0.15388	2
4	Tretes	-1.00630	1.59954	-1.24897	-1.11847	2
5	Kedung Urung	1.13990	0.40041	-0.95486	0.49539	3
6	Kedunggalar	-1.29817	0.00018	-0.47121	0.19639	2
7	Walikukun	-0.63740	1.47346	-0.60620	0.44869	2
8	Jogorogo	1.51899	1.38791	0.40470	-0.08828	3
9	Bekoh	-0.38299	-0.23255	1.11582	-0.74475	2
10	Kedung Bendo	1.00719	-2.35983	-2.15573	-0.85155	4
11	Ngale	-0.40664	-0.04148	0.73410	0.52489	1
12	Paron	0.07462	-0.70951	0.82517	1.10742	1
13	Mardiasri	0.31985	-0.22048	0.97524	1.28964	1
14	Padas	-1.17995	0.00540	1.23721	-1.68636	2
15	KarangJati	2.13017	-0.02286	0.67299	-0.44750	3
16	Sambiroto	0.22193	-0.67960	0.06719	-0.69988	2
17	Begal	-1.52663	-1.71653	0.14106	0.30164	1
18	Guyung	0.63064	0.26268	0.89576	-1.39055	2

### Multivariate Adaptive Regression Splines (MARS)

Metode regresi nonparametrik seperti metode Multivariate Adaptive Regression Splines (MARS) dapat digunakan dalam permasalahan klasifikasi. Metode MARS diterapkan untuk klasifikasi pada data curah hujan dengan 2 kelompok zona, yaitu zona curah hujan rendah dan tinggi, sehingga pos yang termasuk di dalam zona curah hujan rendah diberi kode 0, sedangkan pos yang termasuk di dalam zona curah hujan tinggi diberi kode 1. Pada penelitian ini, metode MARS dilakukan dengan dua macam cara, yaitu tidak melibatkan interaksi antar variabel bebasnya dan menggunakan interaksi. Hasil klasifikasi dengan metode MARS cara pertama menghasilkan model sebagai berikut.

$$\hat{Z} = -0.4082 + 0.0033(X_{\text{Januari}} - 213.33)_+ + 0.0053(X_{\text{Maret}} - 236.25)_+$$

dengan  $X_{\text{Januari}}$  adalah nilai curah hujan pada bulan Januari dan  $X_{\text{Maret}}$  adalah nilai curah hujan pada bulan Maret. Karena terdapat 2 kelompok, apabila diperoleh nilai  $\hat{Z} < 0.5$ , maka pos yang bersangkutan akan diklasifikasikan ke dalam kelompok zona curah hujan rendah. Sedangkan apabila nilai  $\hat{Z} > 0.5$ , maka pos yang bersangkutan akan diklasifikasikan ke dalam kelompok zona curah hujan tinggi. Model MARS pertama ini menghasilkan GCV sebesar 0.09127748 dan  $R^2$  sebesar 72.45%.

Model MARS dengan cara kedua, yaitu dimungkinkan ada interaksi antar variabel, adalah sebagai berikut.

$$\hat{Z} = -0.0800077 + 0.00004316(X_{Februari} - 234.53)_+ \cdot (X_{Maret} - 236.25)_+$$

dengan  $X_{Februari}$  adalah nilai curah hujan pada bulan Februari dan  $X_{Maret}$  adalah nilai curah hujan pada bulan Maret. Karena terdapat 2 kelompok, apabila diperoleh nilai  $\hat{Z} < 0.5$ , maka pos yang bersangkutan akan diklasifikasikan ke dalam kelompok zona curah hujan rendah. Sedangkan apabila nilai  $\hat{Z} > 0.5$ , maka pos yang bersangkutan akan diklasifikasikan ke dalam kelompok zona curah hujan tinggi. Model MARS pertama ini menghasilkan GCV yang lebih kecil sebesar 0.07250143 dan  $R^2$  yang lebih besar yaitu 72.78%.

**Tabel 5.** Tabel Ketepatan Klasifikasi Model MARS Data Tanpa Reduksi

Ketepatan Klasifikasi			Prediksi		Total	% Ketepatan
			Zona 1	Zona 2		
MARS tanpa interaksi	Asli	Zona 1	14	0	14	100%
		Zona 2	0	4	4	
MARS dengan interaksi	Asli	Zona 1	14	0	14	94%
		Zona 2	1	3	4	

Penerapan kedua model MARS yang diperoleh terhadap data curah hujan menghasilkan ketepatan klasifikasi yang berbeda. Hal ini ditunjukkan dengan tabel ketepatan klasifikasi pada Tabel 5, karena persentase ketepatan klasifikasi model MARS tanpa interaksi lebih besar daripada model MARS dengan interaksi, model MARS tanpa interaksi lebih disarankan untuk digunakan dalam klafikasi pos pengukuran curah hujan. Dengan cara yang sama seperti pada data asli tanpa reduksi, dilakukan pembentukan model MARS tanpa interaksi saja karena model MARS dengan interaksi yang optimum (memiliki GCV paling kecil) adalah dengan interaksi sebanyak 1 atau tanpa interaksi. Tabel klasifikasi pada data data curah hujan yang telah direduksi dengan model MARS ditampilkan pada Tabel 6.

**Tabel 4.6.** Tabel Ketepatan Klasifikasi Model MARS Data Reduksi

Ketepatan Klasifikasi			Prediksi		Total	% Ketepatan
			Zona 1	Zona 2		
MARS tanpa interaksi	Asli	Zona 1	14	0	14	100%
		Zona 2	0	4	4	

### 3 KESIMPULAN

Berdasarkan hasil dan pembahasan, maka hasil analisis dengan diskriminan menunjukkan bahwa ketepatan klasifikasi Zona Musim (ZOM) BMKG sebesar 94,4%. Pengelompokan Zona Musim (ZOM) Kabupaten Ngawi menggunakan metode Hirarki *clustering* (pautan lengkap) didapatkan kelompok optimun sebanyak 2 kelompok, baik pada data tanpa reduksi dan reduksi. Berdasarkan nilai ketepatan klasifikasi dari metode MARS pada data reduksi tanpa melibatkan variabel interaksi memberikan kinerja yang lebih baik. Sehingga sebagai early warning ketahanan pangan, penggunaan data kelompok stasiun (ZOM) di kabupaten ngawi terdapat dua zona musim. Zona\_1 (Curah Hujan rendah) terdiri dari 14 pos, antara lain Pos Mantingan, Ngawi, Ngrambe, Tretes, Kedunggalar, Walikukun, Bekoh, Kedung Bendo, Ngale,

Paron, Mardiasri, Padas, Sambiroto, dan Begal. Zona\_2 (Curah Hujan Tinggi) terdiri dari 4 pos lainnya, yaitu Pos Kedung Urung-urung, Jogorogo, Karangjati, dan Guyung.

### Daftar Pustaka

- Aldrian E, Susanto RD. 2003. Identification Of Three Dominant Rainfall Regions Within Indonesia And Their Relationship To Sea Surface Temperature. *Int. J. Climatology* 23: 1435–1452
- Billier, C. and Fahrmeir, L. (2001). Bayesian varying-coefficient model using adaptive regression spline. *Statistical modeling*.
- Busuioc A, von Storch H, Schnur R.1999.Verification of GCM-Generated Regional Seasonal Precipitation for Current Climate and of Statistical Downscaling Estimates under Changing Climate Conditions. *J. Climate* 12:258-272.
- Bunkers WJ, Miller JR, DeGaetano AT.1996.Definition of Climate Regions in the Northern Plains Using an Objective Cluster Modification Technique. *J.Climate* 9:130-146.
- [BMG] Badan Meteorologi dan Geofisika. 2003. Pemilayahan Daerah Tipe Hujan dan Evaluasi Musim Kemarau 2003 serta Prakiraan Sementara Musim Hujan 2003/2004 Kabupaten Indramayu. Jakarta. BMG.
- [BMG] Badan Meteorologi dan Geofisika. 2006. Prakiraan Musim Hujan Tahun 2006/2007 di Indonesia. BMG. Jakarta.
- Chan JCL, Shi J. 1997. Application of Projection-Pursuit Principal Component Analysis Method to Climate Studies. *Int. J. Climatology* 17:103-113.
- Chen C.2002. Robust Regression and Outlier Detection with the Robustreg Procedure. SAS Institute.Inc.Cary, NC.
- Corte-Real J, Zhang X, Wang X. 1995. Downscaling GCM information to regional scales: a non-parametric multivariate regression approach. *Climate Dynamics* 11 (7): 413-424.
- Cavazos T, Hewitson B.2005. Performance of NCEP-NCAR reanalysis variables in statistical downscaling of daily precipitation. *Clim Res*, 28:95-107.
- Cavazos T, Hewitson B.2002. Relative Performance of Empirical Predictors of Daily Precipitation. Proc.of the 1<sup>st</sup> Biennial Meeting of the IEMSS, Lugano, Switzerland, 2:349-354.
- De Veaux RD,Gordon, Arnold L. Comiso, Joey CB, Nadine E. 1993. Modeling of topo-graphic effects on Antarctic sea ice using multivariate adaptive regression splines. *Journal of Geophysical Research*, 98 (C11): 20307-20320.
- Friedman, J. H. (1991). Multivariate Adaptive Regression Splines (with Discussion). *The Annals of Statistics*, 19:1-141.
- Finizio M, Palmieri S. 1998. Non-linear modelling of monthly mean vorticity time changes: an application to the western Mediterranean.*Ann.Geophysicae* 16:116-124.
- Falcon WP, Naylor RL, Wada N, Smith WL, Burke MB, McCullough EB. 2004. Using Climate Models To Improve Indonesian Food Security. *Bulletin of Indonesian Economic Studies* 40: 355-377.
- Giorgi F, Hewitson B, Christensen J, Hulme M, Von Storch H, Whetton P, Jones R, Mearns L, Fu C. 2001. the scientific basis. Contribution of Working Group I to the Third Assesment Report of the Intergovernmental Panel on Climate Change IPCC.University Press. Cambrige.UK.
- Gong X, Richman MB. 1995. On the Application of Cluster Analysis to Growing Season Precipitation Data in North America East of the Rockies. *J.Climate* 8:897-931.
- Giannini A. 2005. ENSO and the seasonal cycle of precipitation in Southeast Asia. <http://iri.columbia.edu>. [September 2007]
- Hall, P. and Opsomer, J. D. (2005). Theory for penalized spline regression. *Biometrika*, 92:1-105.
- Haryoko U.2004. Pendekatan Reduksi Dimensi Luaran GCM untuk Penyusunan Model *Statistical Downscaling* [Tesis]. Bogor: Sekolah Pascasarjana, Institut Pertanian Bogor.
- Hastie, T., Tibshirani, R. and Friedman, J. H. (2001). *The Element of Statistical Learning: Data Mining, Inference, and Prediction*. Springer series in Statistics, New York.
- He, X. and Shi, P. (1988). Monote B-Spline Smoothing, *Journal of American Statistical Association*, 93.
- Holmes, C. C. and Mallick, B. K. (2003). Generalized Nonlinear Modeling with Multivariate Free-Knot Regression Spline, *Journal of American Statistical Association*, 98.
- Huth R. 1999. Statistical downscaling in central Europe: evaluation of methods and potential predictors. *Clim Res*,13: 91–101.
- Jesús M, Angel F. 2004. Comparison of statistical methods commonly used in predictive modelling. *Journal of Vegetation Science* 15: 285-292.
- Katz RW, Parlange MB, Naveau P. 2002. Statistics of extremes in hydrology. *Advances in Water Resources* 25:1287-1304.
- Katz RW. 2002. Technique for estimating uncertainty in climate change scenarios and impact studies. *Clim Res* 20:167-185.
- Keppenne C, Lall U.1996. Complex Singular Spectrum Analysis and Multivariate Adaptive Regression Splines Applied to Forecasting the Southern Oscillation. <http://www.cpc.noaa.gov/products/predictions/> [11 Oktober 2006]

- Lewis PAW, Stevens JG. 1991. Nonlinear Modelling of Time Series Using Multivariate Adaptive Regression Splines. *J. Amer. Statist. Assoc* 86: 864 – 877.
- Leathwick JR, Rowe D, Richardson J, Elith J, Hastie T. 2005. Using multivariate adaptive regression splines to predict the distributions of New Zealand's freshwater diadromous fish. *Freshwater Biology* 50: 2034–2052.
- Mearns LO, Hulme M, Carter TR, Leemans R, Lal M, Whetton P. 2001. The scientific basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change IPCC. University Press. Cambridge. UK.
- Mimmack GM, Mason SJ, Galpin JS. 2001. Choice of Distance Matrices in Cluster Analysis: Defining Region. *J. Climate* 14:2790-2797.
- Morgan, J. N. and Sonquist, J. S. (1963). Problem in the analysis of survey data and a proposal. *Journal of American Statistical Association*, 58:45-434.
- Morgan, J. and Messinger, R. (1973). THAID: A sequential search program for the analysis of nominal scale dependent variables. (Technical Report). Institute for social Research, University of Michigan, Ann Arbor, Michigan.
- Naylor RL, Falcon WP, Daniel Rochberg D, Nikolaswada. 2001. Using El Niño/Southern Oscillation Climate Data To Predict Rice Production In Indonesia. *Climatic Change* 50: 255–265.
- Naylor RL, Falcon WP, Wada N, Rochberg D. 2002. Using El Niño-Southern Oscillation Climate Data To Improve Food Policy Planning In Indonesia. *Bulletin of Indonesian Economic Studies* 38: 75–91.
- Naylor RL, Falcon WP, Wada N, Battisti D, Vimont DJ, Burke MB. 2007. Assessing risk of climate variability and climate change for Indonesian rice agriculture. *PNAS* 104: 7752–7757.
- Otok, B.W. 2005. *Teknik Pengelompokkan dan Ketepatan Klasifikasi Wilayah Indonesia berdasarkan PDRB*, UNS, Solo.
- Otok, B.W., Subanar, Guritno, S.. 2006. *Optimize Knot and Basis Function at Truncated Spline and Multivariate Adaptive Regression Splines*, Proceedings of The First International Conference on Mathematics and Statistics (ICoMS-1), Bandung, West Java, Indonesia.
- Otok, B.W., Subanar, Guritno, S. 2006. *Basis Function of Spline and MARS With Bootstrap*, Second Joint Conference Indonesia Malaysia On Mathematics and Statistics, ITS, Surabaya, Indonesia.
- Otok, B.W., Subanar, Guritno, S. 2006. Faktor-faktor yang mempengaruhi volume perdagangan saham Menggunakan Multivariate adaptive regression splines, *Jurnal Widya Manajemen & Akuntansi*, Vol 6, Nomer 3, UWM, Surabaya
- Otok, B.W., Sjahid, M. A., Subanar, Guritno, S. 2007. Pendekatan Bootstrap pada Klasifikasi Pemodelan Respon Ordinal. *Jurnal Ilmu Dasar*, Volume 8, Nomor 1, FMIPA, Uiniversitas Jember.
- Otok, B.W. 2008. Asimtotik Estimator Multivariate adaptive regression splines. *Jurnal Nature Indonesia (JNI)*, Riau.
- Sutikno, Boer R. 2005a. Model Prediksi Curah Hujan Dengan Regresi Splines Adaptif Berganda. *Jurnal Agromet Indonesia* 18:36-52.
- Sutikno, Boer R. 2005b. Penggunaan Regresi Splines Adaptif Berganda, Regresi Komponen Utama, dan Jaringan Syaraf Tiruan untuk Statistical Downscaling GCM. Di Dalam Prosiding seminar Nasional Statistika VII ITS Surabaya. 26 Nopember 2005.
- Stephen DJ, Walker GK, Lyons TJ. 1994. Forecasting Australian wheat yield with a weighted rainfall index. *Agriculture and Forest Meteorology* 71: 247-263.
- Timbal B, Dufour A, McAvaney B. 2003. An estimate of future climate change for western France using statistical downscaling technique. *Climate Dynamics* 20:807- 823.
- Trigo RM, Palutikof JP. 2001. Precipitation Scenario over Iberia. A Comparison between Direct GCM Output and Different Downscaling Techniques. *J. Climate* 14:4422-4446.
- Uvo CB, Olsson J, Morita O, Jinno K, Kawamura A, Nishiyama K, Koreeda N, Nakashima T. 2001. Statistical atmospheric downscaling for rainfall estimation in Kyushu Island Japan. *Hydrology and Earth System Sciences* 5: 259-271.
- Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns LO. 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. <http://www.ipcc-ddc.cru.uea.ac.uk/guidelines/> [8 Desember 2004].
- Wetterhall F. 2002. Statistical downscaling in climate studies. An overview of used approaches. Report Series A No.59. Uppsala Universitet.
- Wetterhall F. 2005. Statistical downscaling of Precipitation from Large-scale Atmospheric Circulation: Comparison of Methods and Climate Regions. [Dissertation] Uppsala Universitet. <http://urn.kb.se/resolve> [September 2007]
- Wigena AH, Aunuddin. 2004. Aplikasi *Projection Pursuit* dan Jaringan Syaraf Tiruan dalam Pemodelan Statistical Downscaling. *Jurnal Statistika UNISBA*.4(2) Edisi Khusus:7-10.

- Wigena AH. 2006. Pemodelan Statistical Downscaling dengan Regresi Projection Pursuit untuk Peramalan Curah Hujan Bulanan. Kasus Curah Hujan Bulanan di Indramayu. [Disertasi]. Bogor: Sekolah Pascasarjana, Institut Pertanian Bogor.
- von Storch H, Hewitson B, Mearns L. 1999. Review of Empirical Downscaling Techniques. <http://w3g.gkss.de/staff/storch/> [8 Desember 2004].
- von Storch H, Zorita E, Cubash U. 1993. Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. *J. Climate* 6:1161-1171.