

MODEL PREDIKSI CURAH HUJAN DENGAN REGRESI *SPLINES* ADAPTIF BERGANDA

(Using *Multivariate Adaptive Regression Spline* to Predict Monthly Rainfall)

Sutikno¹, dan Rizaldi Boer²

¹ Jurusan Statistika FMIPA ITS Surabaya

² Jurusan Geofisika dan Meteorologi FMIPA IPB

ABSTRACT

This paper discusses the use of multivariate adaptive regression splines (MARS), for predicting monthly rainfall. Development of MARS model follow stepwise process based on recursive partitioning with splines. Unlike classical method, MARS does not require assumption. This model used to model relationship ENSO index and monthly rainfall. This result of rainfall prediction from this model was compared with those resulted from model developed using least square method. Based on R^2 and Corrected R^2 , and mean square error (MSE), the MARS gives better estimates than the least square method. The interval R^2 and Corrected R^2 , and MSE MARS method were 25,2 % - 70,5 %; 23,7 % - 68,6 %; 6,859 – 18,781, while least square method were 5,5 % - 52,6 %; 3,9 % - 51,4 %; 8,075 – 27,717. The result of model validation by using MARS method was 30 % - 80 %, while least square method was 30 % - 70 %.

Key word: ENSO, multivariate adaptive regression splines, monthly rainfall prediction regression model

PENDAHULUAN

Latar belakang

Model-model stokastik untuk memodelkan curah hujan yang sudah dikembangkan di Indonesia diantaranya model regresi (regresi berganda, fourier regression), model deret waktu (ARIMA, additive-winter, fungsi transfer), fractal analysis, neural network, transformasi wavelet, geostatistics, dan trend surfaces analysis (e.g. Dupe, 1999; Haryanto, 1999; Boer *et al.*, 2000; Liong *et al*, 2001; Ratag, 2001; Haryoko, 1997; Zifwen, 1999; Andriansyah, 1998; Gunawan *et al.* 2001), sedangkan penggunaan model deterministik (dynamic) baru pada taraf pengujian model.

Model-model statistik yang berkembang saat ini untuk tujuan peramalan iklim belum memberikan hasil yang memuaskan, sehingga upaya untuk meningkatkan keakuratan peramalan masih terus dilakukan baik melalui perbaikan metode peramalan maupun pengembangan model. Metode regresi splines adaptif berganda (RSAB), yang dikenal dengan *Multivariate Adaptive Regression Splines* (MARS), diperkirakan dapat meningkatkan akurasi dalam memperkirakan curah hujan bulanan.

Permasalahan yang muncul dalam pemodelan curah hujan dengan menggunakan metode regresi linier (metode kuadrat terkecil: MKT) adalah membutuhkan asumsi yang sangat ketat. Asumsi-asumsi tersebut diantaranya: kurva regresi harus diketahui, kenormalan sisaan, dan kehomogenan ragam sisaan. Pelanggaran terhadap asumsi berakibat ketidaksahihan kesimpulan. Berdasarkan pengalaman, seringkali data iklim melanggar asumsi tersebut terutama bentuk kurva regresinya tidak diketahui. Sehingga dengan pertimbangan tersebut di atas diperlukan suatu metode yang tidak terlalu ketat terhadap asumsi. Salah satu metode yang tidak ketat dalam asumsi (soft modelling) adalah RSAB yang akan dibahas dalam makalah ini.

Penelitian ini bertujuan untuk : 1) membuat model hubungan indikator ENSO dengan curah hujan bulanan dengan metode RSAB di tiga tipe wilayah hujan (tipe monsun, equatorial dan lokal), 2) membandingkan metode RSAB dan MKT.

TINJAUAN PUSTAKA

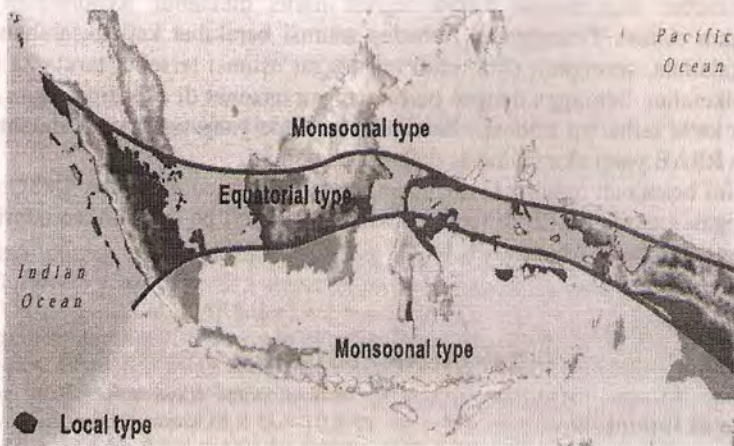
Keragaman iklim di Indonesia

Indonesia merupakan negara yang dilewati oleh garis khatulistiwa dan masuk ke dalam pengaruh kawasan laut pasifik. Posisi ini menjadikan Indonesia sebagai daerah pertemuan sirkulasi meridional (Hadley) dan sirkulasi zonal (Walker), dua sirkulasi yang sangat mempengaruhi keragaman iklim Indonesia. Selain itu karena posisi matahari berpindah dari 23.5 LS ke 23.5 LU sepanjang tahun, aktivitas monsun juga ikut berperan dalam mempengaruhi keragaman iklim. Karena Indonesia merupakan negara kepulauan dengan bentuk topografi yang sangat beragam maka sistem golakan lokal juga cukup dominan dan pengaruhnya terhadap keragaman iklim di Indonesia tidak dapat diabaikan. Faktor lain yang diperkirakan ikut berpengaruh terhadap keragaman iklim Indonesia ialah gangguan siklon tropis. Semua aktivitas dan sistem ini berlangsung secara bersamaan sepanjang tahun.

Secara klimatologis pola iklim di Indonesia dapat dibagi menjadi tiga yaitu pola monsun, pola equatorial dan pola lokal. Pola monsun dicirikan oleh bentuk pola hujan yang bersifat unimodal (satu puncak musim hujan). Selama enam bulan curah hujan relatif tinggi (biasanya disebut musim hujan) dan enam bulan berikutnya rendah (biasanya disebut musim kemarau). Secara umum musim kemarau berlangsung dari April sampai September dan musim hujan dari Oktober sampai Maret. Pola equatorial dicirikan oleh pola hujan dengan bentuk bimodal (dua puncak hujan) yang biasanya terjadi sekitar bulan Maret dan Oktober yaitu pada saat matahari berada dekat equator. Pola lokal dicirikan oleh bentuk pola hujan unimodal (satu puncak hujan) tapi bentuknya berlawanan dengan pola hujan pada tipe monsun. Gambar 1 menunjukkan bahwa wilayah Indonesia di sepanjang garis khatulistiwa sebagian besar mempunyai pola hujan equatorial, sedangkan pola hujan monsun terdapat di pulau Jawa, Bali, NTB, NTT, dan sebagian Sumatera.

Salah satu penyebab terjadinya gangguan pada sirkulasi Walker ialah fenomena ENSO. ENSO merupakan istilah yang mendeskripsikan secara keseluruhan osilasi selatan (fenomena atmosfer) beserta peningkatan suhu muka laut dan juga penurunan suhu muka laut (fenomena lautan). Namun seringkali istilah ini digunakan oleh banyak pakar untuk merujuk kepada kejadian El-Nino (*warm event*) saja, yaitu meningkatnya suhu muka laut di kawasan tengah dan timur ekuator laut pasifik. Osilasi selatan (*southern oscillation*) adalah osilasi tekanan atmosfer kawasan laut pasifik dan atmosfer laut Indonesia-Australia. Untuk memonitor osilasi selatan ini dibuatkan indeks osilasi selatan (SOI) yaitu nilai perbedaan antara tekanan atmosfer di atas permukaan laut di Darwin (Australia) dan Tahiti (Pasifik Selatan), dimana semakin negatif nilai SOI berarti semakin

kuat kejadian panas (*warm event* atau El-Nino) dan sebaliknya semakin positif nilai SOI semakin kuat kejadian dingin (*cold event* atau La-Nina) (Boer 1999).



Sumber : *The effects of the Southern Oscillation and El Nino on Australia*, Information series 2002, Department of Primary Industries, Queensland Government.

Gambar 1. Wilayah Indonesia berdasarkan pola hujan.

Menurut LAPAN (1998) dalam Zifwen (1999) terdapat tiga faktor yang mempengaruhi curah hujan di wilayah Indonesia dalam berbagai skala yaitu : (1) faktor global skala besar interannual, (2) faktor regional skala sedang seasonal dan (3) faktor lokal skala kecil intraseasonal. ENSO merupakan salah satu faktor global sehingga pola iklim (curah hujan) di wilayah Indonesia tidak semua terpengaruh langsung oleh fenomena ini.

Regresi *Splines* Adaptif Berganda

Regresi *splines* adaptif berganda (RSAB), merupakan metode yang dikembangkan oleh Friedman pada tahun 1991. Metode ini terus berkembang seiring dengan perkembangan komputer dan digunakan diberbagai disiplin ilmu. Berikut ini akan dibahas konsep dasar model regresi, recursive partitioning dan modifikasi Friedman (RSAB).

Misalkan Y menunjukkan peubah respon tunggal tergantung pada p peubah prediktor X , dimana $X = (X_1, X_2, \dots, X_p)$, maka Y dapat digambarkan dengan model regresi sebagai berikut :

$$y = f(X_1, X_2, \dots, X_p) + \varepsilon \quad (1)$$

dengan daerah asal (domain) $D \subset R^p$. Peubah acak ε diasumsikan mempunyai rata-rata (*mean*) nol dan ragam σ^2_ε . Jika bentuk kurva $f(X)$ diketahui atau diasumsikan diketahui, maka dapat digunakan regresi parametrik untuk menduga koefisien parameternya. Agar dapat membuat asumsi terhadap bentuk kurva regresi parametrik, diperlukan pengalaman masa lalu atau terdapat sumber-sumber lain yang tersedia dan dapat memberikan informasi detail tentang proses penyelidikan.

Apabila informasi yang tersedia tentang kurva regresi terbatas dan sulit membuat asumsi terhadap bentuk kurva regresi, maka bagian terbesar informasi terletak pada pola data (*data driven*).

Recursive Partitioning

Dari persamaan (1), misalkan terdapat N contoh dari Y dan $X = (X_1, X_2, \dots, X_p)$, dinamakan $\{y_i, x_i\}_{i=1}^N$. Ambil $\{R_j\}_{j=1}^S$, s merupakan himpunan bagian (*subregion*) yang saling lepas dari daerah asal D , sehingga $D = \bigcup_{j=1}^S R_j$. *Recursive Partitioning* (RP) menduga fungsi $f(X)$ yang tidak diketahui dengan :

$$\hat{f}(X) = \sum_{j=1}^S c_j(X) B_j(X) \quad (2)$$

dimana, $B_j(x) = I[x \in R_j]$, $I[\cdot]$ menunjukkan fungsi indikator yang mempunyai nilai 1 (satu) jika pernyataan benar dan 0 (nol) jika salah, $c_j(X)$ merupakan koefisien (konstanta) yang ditentukan dalam subregion. Setiap fungsi indikator merupakan perkalian fungsi tunggal (*univariate step function*) yang menggambarkan setiap subregion R_j .

Secara umum, prosedur RP mempunyai 2 (dua) tahap yang dimulai dari subregion pertama $R_1 = D$. Tahap pertama, atau *forward*, memilah secara iteratif daerah asal D menjadi himpunan bagian (subregion) yang saling lepas $\{R_j\}_{j=2}^M$, untuk $M \geq S$, dimana M ditentukan sembarang. Tahap kedua, atau *backward*, pada tahap ini berlawanan dengan langkah pertama yaitu menghilangkan atau memangkas ($M-S$) subregion dari model dengan dua kriteria yaitu evaluasi dugaan model dan jumlah subregion dalam model. Kedua tahap tersebut mendapatkan sekumpulan subregion yang tidak saling tumpang tindih, sehingga dugaan $\hat{f}(X)$ mendekati $f(X)$ untuk setiap subregion daerah asal.

RP merupakan metode yang menjanjikan, meskipun demikian secara umum RP terdapat kekurangan dalam pemodelan regresi diantaranya : (1) Model RP menghasilkan subregion yang saling lepas dan diskontinu pada batas subregion, (2) RP tidak cukup mampu dalam menduga fungsi $f(X)$ linier atau aditif, dan (3) Bentuk model RP pada persamaan 2 mengalami kesulitan jika p yang besar.

Modifikasi Friedman

Beberapa inovasi dilakukan oleh Friedman untuk mengatasi kelemahan metode RP. Untuk mengatasi kelemahan RP dalam mengidentifikasi fungsi linier dan aditif, Friedman mengusulkan bahwa tidak menghapus induk (parent) region selama pemilahan subregion berlangsung. Jadi pada iterasi berikutnya parent dan pilahan subregion dapat dipilah lebih lanjut, sehingga diperoleh subregion yang saling tumpang tindih. Dengan modifikasi ini, RP dapat menghasilkan model linier dengan pemilahan berulang pada peubah prediktor yang berbeda. Di samping itu dihasilkan model yang lebih fleksibel.

Modifikasi tersebut di atas masih belum bisa mengatasi adanya diskontinu yang disebabkan perkalian fungsi tunggal. Oleh sebab itu, Friedman mengusulkan untuk mengganti perkalian fungsi tunggal (univariate) dengan regresi linier splines (ordo satu) dengan sisi kiri (-) dan sisi kanan (+) truncated splines.

$$B_j(x) = \prod_{i=1}^{s_j} (s_{ij} (x_{k(i,j)} - x_{k(i,j)}^*))_+ \tag{3}$$

dimana s_j jumlah pilahan subregion ke - j dari domain D, $X_{k(i,j)}^*$ merupakan knot dari peubah prediktor $X_{k(i,j)}$, dan s_{ij} nilainya +1 atau -1 jika knotnya terletak di kanan atau kiri subregion.

Algoritma RSAB, khususnya tahap *for-ward* digunakan untuk mendapatkan subregion - subregion agar dapat menentukan basis fungsi. Penentuan titik pemilah (*knot*) dan koefisien sangatlah penting agar mendapatkan model terbaik. Sedangkan tahap *backward*, mengeluarkan suku model (basis fungsi) yang kontribusinya terhadap nilai dugaan respon kecil. Adapun ukuran kontri-busi yang digunakan tahap ini adalah modifikasi kriteria validasi silang (*Generalized Cross Validation*, GCV) Craven dan Wahba (1979) dalam Lewis (1991) yakni :

$$GCV^*(s) = \frac{(1/N) \sum_{i=1}^N [y_i - \hat{f}_s(x_i)]^2}{[1 - (C(s)^*)/N]^2} \tag{4}$$

Pembilang GCV* adalah rata-rata jumlah kuadrat galat (*average sum square of residual*, ASR), s jumlah subregion yang ditentukan pada tahap *forward*, dan C(s)* merupakan fungsi model kompleks. Model terbaik jika nilai GCV* minimum.

Modifikasi dalam algoritma RP menghasilkan persamaan model RSAB sebagai berikut :

$$\hat{f}(X) = c_0 + \sum_{j=1}^s c_j \prod_{i=1}^{s_j} (s_{ij} (x_{k(i,j)} - x_{k(i,j)}^*))_+ \tag{5}$$

dimana c_0 adalah koefisien konstanta dari basis fungsi B_1 . Koefisien $\{c_j\}_{j=1}^s$ ditentukan dengan menggunakan metode kuadrat terkecil.

DATA DAN METODE

Peubah Respon dan Prediktor

Peubah respon yang digunakan adalah rata-rata curah hujan bulanan dari 3 (tiga) tipe hujan yaitu : (a) tipe monsun meliputi Tuban (1958–1984), Sandakan (1958–1996), dan Rembiga-Ampenan (1958–1990), (b) tipe hujan equatorial meliputi Pekanbaru (1953–1997), Pontianak (1947–1997), dan Medan (1948–1997), dan (c) tipe hujan lokal : Ambon (1961–1997).

Peubah prediktor meliputi : (a) Tekanan permukaan laut di Tahiti (1951–2001), (b) Tekanan permukaan laut di Darwin (1951–2001), (c) Suhu permukaan laut (*sea surface temperature* : SST) terdiri atas Nino 1.2, Nino 3, Nino 4, dan Nino 3.4 dari tahun 1950–1997, (d) Indeks osilasi selatan (*southern oscillation index*: SOI) Tahiti-Darwin, yaitu nilai perbedaan antara tekanan atmosfer di atas permukaan laut di Tahiti (Pasifik Selatan) dan Darwin (Australia) dibagi dengan simpangan baku selisih antara tekanan Tahiti dan tekanan Darwin dari tahun 1950–1997, dan (e) Indeks osilasi Jakarta-Darwin (IO_{JD}) dari tahun 1958–2001, yaitu nilai perbedaan antara tekanan atmosfer di atas permukaan laut di Jakarta dan Darwin (Australia) dibagi dengan simpangan baku selisih antara tekanan Jakarta dan tekanan Darwin.

Analisis Data

Untuk mencapai tujuan, dilakukan analisis data dengan bantuan piranti MARS for windows versi 2.0, Minitab versi 13, dengan langkah sebagai berikut :

1. Langkah awal untuk membangun model regresi membuat plot antara peubah respon (curah hujan) dan peubah prediktor (indikator ENSO). Dari plot tersebut akan terdeteksi pola hubungan, pengamatan pencilan atau penyimpangan kelinieran. Karena periode data yang tersedia tidak sama maka perlu disamakan periodenya.
2. Pembentukan model regresi dengan menggunakan metode kuadrat terkecil, akan diperoleh model regresi, sidik ragam, R^2 , $R^2_{\text{terkoreksi}}$, dan peubah prediktor yang nyata.
3. Pembentukan model regresi dengan menggunakan metode RSAB. Model regresi yang diperoleh :

$$Y = B_0 + B_1 * BF_1 + B_2 * BF_2 + \dots + B_k * BF_k$$

dimana Y merupakan peubah respon, B_0 adalah konstanta, B_1, B_2, \dots, B_k merupakan koefisien basis fungsi spline ke-1, 2, ..., k dan BF_1, BF_2, \dots, BF_k merupakan basis fungsi ke-1, 2, ... ,k. Pendugaan parameter terlebih dahulu menentukan maksimum basis fungsi, maksimum jumlah interaksi, minimum jumlah pengamatan diantara knot. Di samping model regresi, diperoleh juga R^2 , $R^2_{\text{terkoreksi}}$, dan peubah prediktor yang terpenting. Penentuan peubah terpenting berdasarkan seberapa besar peubah tersebut memberikan kontribusi terhadap model. Kriteria penentuan peubah prediktor yang relatif penting adalah GCV. Semakin kecil nilai GCV suatu peubah, semakin penting peubah tersebut terhadap model yang dibangun.

Verifikasi dan Validasi Model

Verifikasi model digunakan untuk mengetahui kemampuan model dalam menjelaskan keragaman data. Validasi model digunakan untuk melihat keterandalan model tersebut dalam peramalan, yang biasanya digunakan data bebas. Verifikasi model menggunakan kriteria R^2 , $R^2_{\text{terkoreksi}}$, dan MSE, sedangkan validasi model menggunakan analisis korelasi.

HASIL DAN PEMBAHASAN

Pemodelan Hubungan antara Indikator ENSO dan Curah Hujan

Berdasarkan eksplorasi data melalui diagram pencar antara peubah respon dan peubah prediktor menunjukkan bahwa masing-masing daerah tipe hujan mempunyai pola hubungan yang berbeda (Lampiran 1). Pada daerah tipe hujan monsun, tekanan Darwin, dan tekanan Tahiti mempunyai pola yang lebih jelas terhadap curah hujan daripada peubah prediktor lainnya. Untuk daerah tipe hujan equatorial hampir semua prediktor tidak menampakkan pola hubungan yang jelas (acak) dengan peubah respon. Demikian juga pada daerah tipe hujan lokal menunjukkan pola hubungan yang acak antara peubah respon dan peubah prediktor.

Secara umum wilayah dengan tipe hujan monsun dipengaruhi oleh pola dan kejadian ENSO. Sedangkan untuk wilayah dengan tipe hujan equatorial dan lokal masih belum nampak jelas pengaruhnya. Hal ini disebabkan karena pengaruh sirkulasi meridional sangat kuat di daerah sekitar equator, meskipun terjadi El-Nino dampaknya tidak terasa. Demikian juga daerah tipe hujan lokal, pengaruh lokal (kondisi geografi) cukup dominan bila dibandingkan pengaruh global. Hasil ini mendukung penelitian yang dilakukan oleh Tjasyono (1997) dalam Boer (2000).

Berdasarkan statistik *uji t* secara parsial, hasil pendugaan dengan metode kuadrat terkecil untuk daerah tipe hujan monsun menunjukkan bahwa peubah tekanan Darwin (Stasiun Tuban), tekanan Darwin, dan SOI (Stasiun Sandakan), tekanan Darwin dan IO_{JD} (Stasiun Ampenan) nyata pada $\alpha = 5\%$. Untuk tipe hujan equatorial peubah prediktor tekanan Darwin, dan Nino 1.2 (Stasiun Pontianak), tekanan Darwin, Nino 3, dan Nino 1.2 (Stasiun Pekanbaru), Nino 1.2 (Stasiun Medan) nyata pada $\alpha = 5\%$. Sedangkan untuk daerah tipe hujan lokal peubah tekanan Darwin, tekanan Tahiti, Nino 3, Nino 1.2, dan SOI nyata pada $\alpha = 5\%$.

Hasil pemeriksaan sisaan terhadap model curah hujan tipe hujan monsun menunjukkan adanya ragam yang tidak homogen karena plot antara sisaan dan dugaan respon yang membentuk pola corong (Lampiran 2). Demikian juga dengan model tipe hujan lokal menunjukkan adanya ragam yang tidak homogen. Berbeda dengan model tipe hujan equatorial menunjukkan adanya ragam yang homogen.

Adanya keterbatasan informasi bentuk fungsi regresi dan hasil eksplorasi data yang menampakkan pola hubungan yang tidak jelas kelinierannya antara peubah respon dan peubah prediktor, serta terdapat kasus heteroskedastisitas, multikolinieritas dan kemungkinan autokorelasi, maka dengan pertimbangan ini dilakukan metode regresi splines adaptif berganda.

Pendugaan Model Regresi dengan RSAB

Sebagai ilustrasi, dipilih Stasiun Tuban untuk dibahas secara rinci. Model curah hujan Stasiun Tuban dibentuk dengan kriteria input : Minspan (minimal banyaknya pengamatan setiap knot) = 10, maksimal interaksi (MI) = 3, dan maksimal basis fungsi sebesar 40. Sedangkan model untuk stasiun lainnya disajikan dalam Lampiran 3 dan 4.

Pendugaan model regresi *splines* adaptif berganda di Stasiun Tuban menghasilkan model dengan satu intersep dan 20 basis fungsi, meliputi : 3 interaksi level pertama, 8 interaksi level dua, dan 9 interaksi level tiga. Jumlah nilai *knot* atau *threshold* sebanyak 16 buah di antaranya : 1 nilai untuk peubah Nino-3, 2 nilai untuk peubah tekanan Darwin, tekanan Tahiti, Nino 3.4, dan IO_{JD}, 3 nilai untuk peubah Nino 1.2 dan 4 nilai untuk peubah SOI. Nilai ragam sisaan (*mean square error*: MSE) sebesar 7133, R² sebesar 70,5% dan R²_{terkoreksi} sebesar 68.6%. Peubah SOI mempunyai nilai *knot* terbanyak karena mempunyai keragaman data terbesar di antara peubah prediktor lainnya.

Interpretasi model RSAB terletak pada komponen sidik ragam. Tabel 1 menunjukkan komponen basis fungsi yang membentuk model regresi curah hujan Tuban baik interaksi level pertama maupun interaksi antar peubah.

Deskripsi peubah prediktor untuk stasiun Tuban selengkapnya disajikan pada Tabel 2.

Tabel 1. Komponen sidik ragam model curah hujan Tuban

Basis Fungsi	Parameter	Koefisien	S.E	T-rasio	P-value
0	Konstanta	62.805	10.64	5.903	0.000
1	max(0, DARWIN - 10.500)	-13.992	5.57	-2.512	0.013
2	max(0, 10.500 - DARWIN)	50.773	12.084	4.202	0.000
3	max(0, NINO3 - 24.490) * BF2	54.169	17.935	3.02	0.003
4	max(0, 24.490 - NINO3) * BF2	-42.967	12.125	-3.544	0.000
5	max(0, SOI - 6.900) * BF2	-5.262	0.984	-5.35	0.000
6	max(0, 6.900 - SOI) * BF2	3.794	0.804	4.717	0.000
7	max(0, TAHITI - 13.300) * BF2	314.355	103.304	3.043	0.003
8	max(0, 13.300 - TAHITI) * BF2	-13.027	3.631	-3.587	0.000
9	max(0, IO _{JD} + 3.600) * BF7	-394.274	58.257	-6.768	0.000
10	max(0, NINO12 - 26.810)	63.468	23.957	2.649	0.008
11	max(0, 26.810 - NINO12)				
12	max(0, IO _{JD} + 2.500) * BF11				
13	max(0, - 2.500 - IO _{JD}) * BF11	-32.367	9.822	-3.296	0.001
14	max(0, NINO34 - 24.580) * BF7	361.02	79.424	4.545	0.000
15	max(0, TAHITI - 13.200) * BF5	17.563	6	2.927	0.004
18	max(0, 26.030 - NINO34) * BF12	17.631	4.706	3.746	0.000
20	max(0, 8.900 - DARWIN) * BF11	19.279	5.229	3.687	0.000
22	max(0, - 8.300 - SOI) * BF20	-5.093	1.114	-4.571	0.000
23	max(0, SOI - 3.100) * BF8	3.118	0.814	3.831	0.000
27	max(0, NINO12 - 22.690) * BF3	-113.484	31.765	-3.573	0.000
28	max(0, 22.690 - NINO12) * BF3	-159.053	36.029	-4.415	0.000
29	max(0, NINO12 - 23.250) * BF3	115.125	32.13	3.583	0.000
F-statistic = 36.286		S.E. of Regression = 84.458			
P-value = 0.000		Residual Sum of Squares = 2161344.229			
[MDF,NDF] = [20, 303]		Regression Sum of Squares = 5176698.833			
R-squared = 0.705; Adj R-squared = 0.686					

Model regresi yang diperoleh :

$$Y = 62.806 - 13.993 * BF1 + 50.772 * BF2 + 54.127 * BF3 - 42.970 * BF4 - 5.263 * BF5 + 3.795 * BF6 + 314.356 * BF7 - 13.028 * BF8 - 394.276 * BF9 + 63.470 * BF10 - 32.368 * BF13 + 361.017 * BF14 + 17.563 * BF15 + 17.631 * BF18 + 19.280 * BF20 - 5.093 * BF22 + 3.118 * BF23 - 113.484 * BF27 - 159.053 * BF28 + 115.125 * BF29$$

Tabel 2. Nilai rata-rata, simpangan baku, minimal, dan maksimal peubah prediktor curah hujan stasiun Tuban

Peubah	Rataan	Simpangan baku	Minimal	Maksimal
Tek. Darwin	9.842	2.594	4.000	15.500
Tek. Tahiti	12.565	1.690	7.400	16.600
Nino 1.2	23.032	2.246	18.800	28.770
Nino 3	25.748	1.263	23.380	29.160
Nino 3.4	26.945	0.930	24.580	29.350
Nino 4	28.336	0.608	26.610	29.740
SOI	-0.157	9.856	-35.700	31.500
IO _{JD}	-0.852	1.119	-3.600	2.800

Peubah Prediktor yang Relatif Penting

Peubah prediktor yang relatif penting untuk stasiun Tuban adalah tekanan Darwin. Hal ini ditunjukkan pada nilai GCV terkecil (terbesar untuk GCV^{-1}) diantara peubah lainnya. Pendeteksian awal peubah prediktor yang relatif penting ditentukan dari besarnya nilai simpangan bakunya (Tabel 2). Seperti halnya pada prosedur regresi bertatar setiap peubah dikeluarkan satu per satu dari model kemudian dihitung kebaikan model (*goodness of fit*). Peubah yang relatif penting jika mempunyai pengaruh yang terbesar terhadap kebaikan model dan sebaliknya untuk peubah yang tidak penting.

Secara umum, peubah prediktor ENSO yang relatif penting terhadap model curah hujan adalah tekanan Darwin. Tercatat 6 stasiun (yaitu Tuban, Ampenan, Sandakan, Pontianak, Pekanbaru, dan Ambon) peubah ini memberikan pengaruh yang nyata terhadap model curah hujan. Hal ini ditunjukkan pula dengan nilai korelasi yang tinggi antara curah hujan dan tekanan Darwin (Tabel 3) bila dibandingkan dengan indikator ENSO lainnya.

Bila ditinjau dari nilai korelasinya, tekanan Darwin sebagian besar bernilai negatif terutama di daerah tipe hujan monsun. Fenomena ini terjadi karena pada saat tekanan permukaan laut di Darwin meningkat dengan membawa uap air yang sedikit (karena melalui daratan benua) akan bergerak ke wilayah Indonesia, sehingga mengakibatkan curah hujan berkurang.

Tabel 3. Korelasi antara indikator ENSO dan curah hujan menurut stasiun

Stasiun	Indikator ENSO							
	IO _{JD}	SOI	Nino1.2	Nino 3	Nino 4	Nino 3.4	Tahiti	Darwin
Tuban	-0.120	0.154	0.279	-0.085	-0.354	-0.307	-0.457	-0.710
Sandakan	-0.152	0.126	-0.120	-0.343	-0.217	-0.362	-0.189	-0.379
Ampenan	-0.102	0.185	0.243	-0.099	-0.312	-0.297	-0.393	-0.653
Medan	-0.088	0.102	-0.353	-0.285	-0.052	-0.174	0.176	0.102
Pekanbaru	-0.078	0.047	0.038	-0.031	-0.089	-0.077	-0.189	-0.251
Pontianak	-0.058	0.019	-0.020	-0.042	0.004	-0.029	-0.127	-0.150
Ambon	-0.086	0.142	-0.211	-0.041	-0.004	0.037	0.416	0.423

Perbandingan Metode

Hasil pendugaan model regresi dan validasi model dengan menggunakan metode regresi splines adaptif berganda lebih baik dari metode kuadrat terkecil. Hal ini nampak pada nilai R^2 , dan $R^2_{\text{terkoreksi}}$ dari metode metode regresi splines adaptif berganda lebih besar dari metode kuadrat terkecil. Di samping itu nilai ragam sisaan dari metode kuadrat terkecil lebih besar (Tabel 4). Hasil validasi model dengan metode RSAB dan MKT dapat dilihat pada Lampiran 5. Bila ditinjau dari nilai korelasi antara hasil prediksi/ramalan dan data aktual, nilainya berkisar antara 30% hingga 80%. Sedangkan nilai korelasi MKT berkisar antara 30% hingga 70% (Tabel 5).

Tabel 4. Nilai R^2 , $R^2_{\text{terkoreksi}}$, dan MSE model curah hujan menurut tipe hujan berdasarkan MKT dan RSAB

MKT				
Tipe Hujan	Stasiun	Kriteria		
		R^2	$R^2_{\text{terkoreksi}}$	MSE
Monsun	Tuban	52,6	51,4	11.047
	Ampenan	45,3	44,1	8.075
	Sandakan	30,1	28,4	21.557
Equatorial	Pontianak	5,5	3,9	17.671
	Pekanbaru	9,7	8,2	9.959
	Medan	13,6	12,2	10.910
Lokal	Ambon	31,9	30,8	27.717
Metode RSAB				
Tipe Hujan	Stasiun	Kriteria		
		R^2	$R^2_{\text{terkoreksi}}$	MSE
Monsun	Tuban	70,5	68,6	7.133
	Ampenan	53,8	52,5	6.859
	Sandakan	59,2	57,0	12.942
Equatorial	Pontianak	27,9	26,1	13.595
	Pekanbaru	25,2	23,7	8.273
	Medan	28,7	26,7	9.101
Lokal	Ambon	54,7	53,1	18.781

Tabel 5. Nilai korelasi validasi model metode RSAB dan MKT

Stasiun	Korelasi	
	RSAB	MKT
Tuban	0.701	0.728
Sandakan	0.659	0.600
Ampenan	0.574	0.603
Pekanbaru	0.642	0.399
Medan	0.778	0.635
Pontianak	0.298	0.316
Ambon	0.768	0.621

Model curah hujan dengan metode RSAB lebih kompleks bila dibandingkan dengan metode kuadrat terkecil, meskipun tersedia selang kepercayaan koefisien regresi, namun cenderung bersifat spesifik lokal. Parameter model akan berubah menurut pola data. Berbeda dengan metode kuadrat terkecil yang mempunyai sifat global. Sehingga untuk kepentingan peramalan metode RSAB membutuhkan *update data*.

KESIMPULAN

Dari hasil analisis data dan pembahasan dapat disimpulkan sebagai berikut :

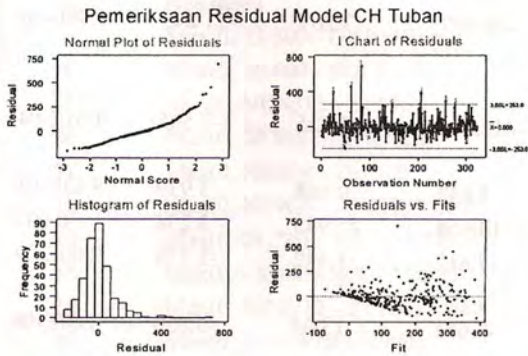
1. Wilayah dengan tipe hujan monsun dipengaruhi oleh pola dan kejadian ENSO. Sedangkan untuk wilayah dengan tipe hujan equatorial dan lokal masih belum jelas pengaruhnya.
2. Jika asumsi bentuk fungsi hubungan antara peubah respon dan peubah prediktor tidak diketahui, seperti data curah hujan maka metode RSAB mempunyai potensi prediksi curah hujan yang lebih baik dari MKT.
3. Berdasarkan nilai korelasi Tekanan Darwin merupakan prediktor yang nyata dalam memodelkan curah hujan di Indonesia.

DAFTAR PUSTAKA

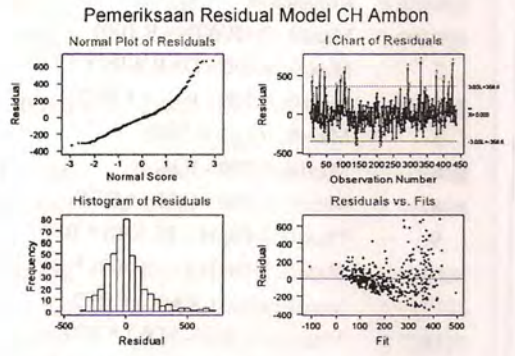
- Andriansyah. 1998. Perbandingan Model Geostatistika dan Box-Jenkins dalam Peramalan Curah Hujan Bulanan. Skripsi. Jurusan Statistika FMIPA IPB, Bogor.
- Boer, R., Notodiputro, K.A., and Las, I. 2000. Prediction of daily rainfall characteristics from monthly climate indices. Proceeding of "The Second International Conference on Science and Technology for the Assessment of Global Climate Change and Its Impacts on Indonesian Maritime Continent, 29 November-01 December 1999.
- Dupe, Z.L. 1999. Prediction Nino3.4 SST anomaly using simple harmonic model. Paper Presented at The Second International Conference on Science and Technology for the Assessment of Global Climate Change and Its Impacts on Indonesian Maritime Continent, 29 November-01 December 1999.
- Friedman, J.H. 1991. Multivariate Adaptive Regression Splines (with discussion). <http://www.salford-systems.com/MARS.pdf>.
- Gunawan, D., Soetamto, Nuryadi, dan Heru, R. 2001. Prakiraan jangka panjang di badan meteorologi dan geofisika. hlm 51-59. Di dalam M.A Ratag *et al* (Penyunting). Prediksi Cuaca dan Iklim Nasional. Prosiding Temu Ilmiah LAPAN. Bandung.
- Haryoko,U. 1997. Peramalan Hujan Bulanan di Ambon Berdasarkan ENSO dengan Fungsi Tranfer. Skripsi. Jurusan Statistika FMIPA IPB. Bogor.
- Liong *et al*. 2001. Prediksi Hujan Bulanan Dengan Logika Samar dan Jaringan Sel saraf Tiruan. hlm 1-8. Di dalam M.A Ratag *et al* (Penyunting). Prediksi Cuaca dan Iklim Nasional. Prosiding Temu Ilmiah LAPAN. Bandung.
- Lewis, P.A.W and J.G.Stevens. 1991. Nonlinear Modelling of Time Series Using Multivariate Adaptive Regression Splines. *J. Amer. Statist. Assoc.* Vol. 86. No. 416. pp. 864 – 877.

- Ratag, M.A. 2001. Pengembangan dan Uji Validasi Model Iklim Berbasis Transformasi Wavelet. hlm 9-14. Di dalam M.A Ratag *et al* (Penyunting). Prediksi Cuaca dan Iklim Nasional. Prosiding Temu Ilmiah LAPAN. Bandung.
- Steinberg, D., Colla, P.L, and Kerry Martin. 2001. MARS User Guide, San Diego, CA: Salford Systems. <http://www.salford-systems.com>
- Tjasyono, B., dan R.K Lestari. 2001. Peramalan Cuaca Bulanan dengan Metoda Statistik. hlm.19-26. Di dalam M.A Ratag *et al* (Penyunting). Prediksi Cuaca dan Iklim Nasional. Prosiding Temu Ilmiah LAPAN. Bandung.
- Zifwen.1999. Peramalan ENSO dan Pemodelan Hubungan ENSO dengan Curah Hujan Monson. Skripsi. Jurusan Statistika FMIPA IPB, Bogor.

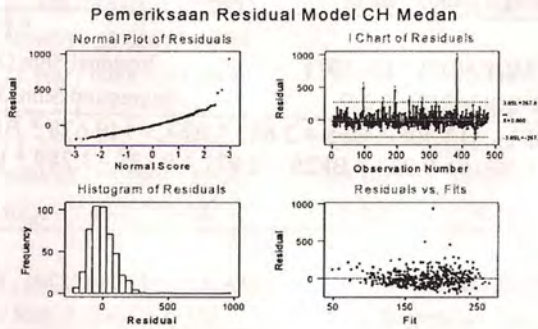
Lampiran 2. Pemeriksaan asumsi sisaan : tipe hujan monsun (a), tipe hujan lokal (b), dan tipe hujan equatorial (c)



(a)



(b)



(c)

Lampiran 3. Komponen sidik ragam model stasiun pekanbaru (tipe hujan equatorial)

Basis Fungsi	Parameter	Koefisien	S.E	T-rasio	P-value
0	Konsatanta	307.36	11.093	27.707	9.99E-16
1	Max(0, DARWIN - 9.100)	-38.697	3.267	-11.843	9.99E-16
2	Max(0, 9.100 - DARWIN)				
4	Max(0, 0.700 - IO _{JD}) * BF2				
5	Max(0, IO _{JD} - 0.700)	-413.314	115.925	-3.565	4.01E-04
6	Max(0, 0.700 - IO _{JD})				
8	Max(0, 3.400 - SOI) * BF6	2.632	0.668	3.939	9.43E-05
9	Max(0, NINO4 - 26.530) * BF5	149.68	45.945	3.258	0.001
12	Max(0, NINO12 - 18.800) * BF6	-7.612	1.442	-5.278	2.00E-07
13	Max(0, IO _{JD} + 2.600) * BF2				
16	Max(0, - 0.300 - SOI) * BF2	-3.654	1.005	-3.636	3.07E-04
24	Max(0, 28.620 - NINO34) * BF1				
25	Max(0, NINO12 - 23.470) * BF24	9.923	3.151	3.149	0.002
31	Max(0, NINO3 - 23.050) * BF4	2.971	1.044	2.844	0.005
40	Max(0, TAHITI - 7.400) * BF13	-3.289	0.705	-4.662	4.08E-06

F-STATISTIC = 17.555 S.E. Of Regression = 90.955
 P-VALUE = .999201E-15 ; MDF,NDF] = [9, 470] Residual Sum Of Squares = 3888235.657
 R-SQUARED: 0.252; ADJ R-SQUARED: 0.237 Regression Sum Of Squares = 1307056.174

$$Y = 307.360 - 38.697 * BF1 - 413.308 * BF5 + 2.632 * BF8 + 149.678 * BF9 - 7.612 * BF12 - 3.654 * BF16 + 9.923 * BF25 + 2.971 * BF31 - 3.289 * BF40$$

Lampiran 4. Komponen sidik ragam model curah hujan ambon (tipe hujan lokal)

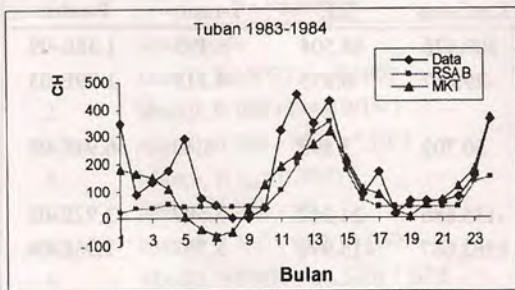
Basis Fungsi	Parameter	Koefisien	S.E	T-rasio	P-value
0	Konstanta	300.476	48.504	6.195	1.38E-09
2	Max(0, 13.300 - DARWIN)	-29.667	6.875	-4.315	1.99E-05
4	Max(0, 10.600 - SOI)				
7	Max(0, NINO12 - 20.550)	30.702	5.592	5.49	6.94E-08
8	Max(0, 20.550 - NINO12)				
11	Max(0, NINO4 - 26.530)	-135.686	24.242	-5.597	3.92E-08
12	Max(0, NINO3 - 24.780) * BF8	1562.667	415.976	3.757	1.96E-04
13	Max(0, 24.780 - NINO3) * BF8				
15	Max(0, - 0.900 - SOI) * BF13	102.428	30.378	3.372	8.15E-04
16	Max(0, NINO12 - 22.700) * BF4	-1.729	0.37	-4.674	3.98E-06
20	Max(0, DARWIN - 9.900)	-117.237	18.627	-6.294	7.75E-10
22	Max(0, SOI + 35.700) * BF20	6.024	0.614	9.815	9.99E-16
23	Max(0, TAHITI - 14.100) * BF20				
25	Max(0, NINO34 - 26.710) * BF23	220.577	37.657	5.858	9.45E-09
28	Max(0, 15.700 - TAHITI) * BF11	19.62	4.148	4.73	3.07E-06
29	Max(0, NINO3 - 23.960) * BF20				
30	Max(0, 23.960 - NINO3) * BF20	191.266	61.753	3.097	0.002
31	Max(0, IO _{JD} + 1.200) * BF29	8.595	2.111	4.072	5.58E-05
32	Max(0, - 1.200 - IO _{JD}) * BF29	56.508	16.346	3.457	6.02E-04
35	Max(0, NINO3 - 24.590) * BF23	-49.381	14.827	-3.331	9.43E-04
37	Max(0, IO _{JD} + 0.400) * BF22	-1.455	0.35	-4.153	3.97E-05

F-STATISTIC = 34.066
P-VALUE = .999201E-15 ; [MDF,NDF] = [15, 423]
R-SQUARED: 0.547; ADJ R-SQUARED: 0.531

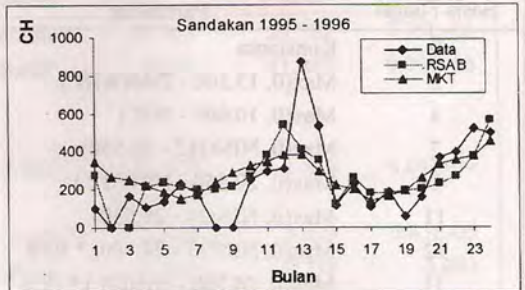
S.E. Of Regression = 137.044
Residual Sum Of Squares = 7944426.677
Regression Sum Of Squares = 9597000.757

$$\begin{aligned}
Y = & 300.476 - 29.667 * BF2 + 30.702 * BF7 - 135.686 * BF11 + 1562.667 * BF12 \\
& + 102.428 * BF15 - 1.729 * BF16 - 117.236 * BF20 + 6.024 * BF22 \\
& + 220.577 * BF25 + 19.620 * BF28 + 191.267 * BF30 + 8.595 * BF31 \\
& + 56.508 * BF32 - 49.381 * BF35 - 1.455 * BF37
\end{aligned}$$

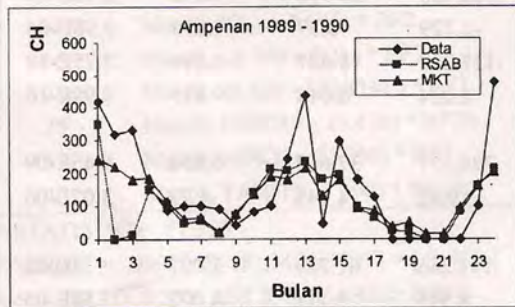
Lampiran 5. Validasi model curah hujan dengan metode rsab dan mkt meliputi : tuban (a), sandakan (b), ampenan (c), pekanbaru (d), medan (e), pontianak (f), dan ambon (g)



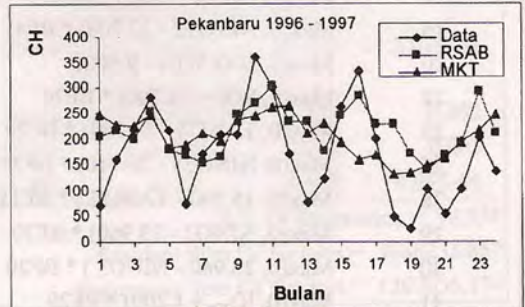
(a)



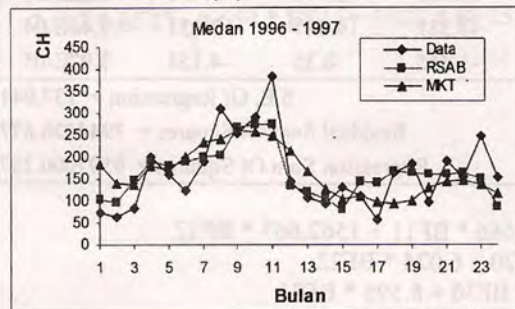
(b)



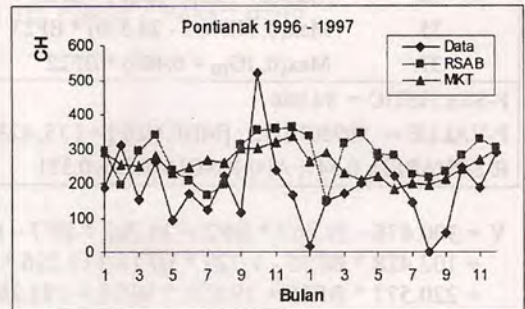
(c)



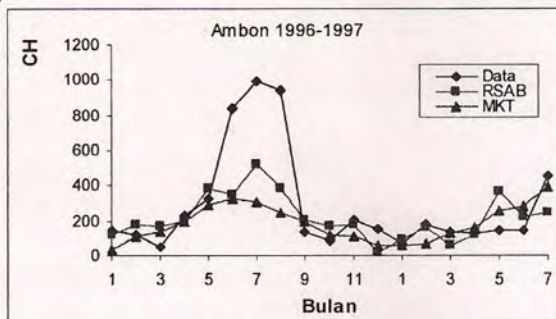
(d)



(e)



(f)



(g)