ITS Institut Teknologi Sepuluh Nopember

# INFERENSI

## Jurnal Statistika

Volume 2, No.1, Januari 2006

5

## Daftar Isi:



*Diterbitkan oleh:*
Divisi Jurnal
**Jurusan Statistika**
**Fakultas Matematika dan Ilmu Pengetahuan Alam**
**Institut Teknologi Sepuluh Nopember (ITS)**
**Surabaya**

***Sekretariat :***
Divisi Jurnal
Jurusan Statistika FMIPA ITS
Kampus ITS Sukolilo
Jl. Arief Rahman Surabaya
Tlp : (031) 5943352
Fax : (031) 5922940
email : inferensi@statistika.its.ac.id

# Pendugaan Model Aditif untuk Data Deret Waktu dengan Pendekatan Model Linear Campuran

Anik Djuraidah[1] dan Aunuddin[2]

**Abstrak**

Model aditif adalah generalisasi dari regresi berganda, dimana nilai tengah peubah respon dimodelkan sebagai penjumlahan dari bentuk hubungan fungsional setiap peubah penjelas. Metode baku pada model aditif mengasumsikan bahwa galat acak bebas stokastik. Asumsi ini tidak sesuai untuk data deret waktu, dimana galat acak pada umumnya saling berkorelasi (autokorelasi). Model aditif baku perlu dikoreksi untuk data dengan galat acak berkorelasi. Hasil simulasi menunjukkan bahwa model aditif deret waktu memberikan penduga parameter pemulus yang lebih tinggi dan jumlah kuadrat galat yang lebih kecil dibandingkan dengan model aditif baku. Hasil ini menunjukkan bahwa model aditif peka terhadap asumsi kebebasan galat acak.

**Kata-kata kunci** : Model Aditif, Regresi *Spline,* Regresi *Spline* Terpenalisasi, Parameter Pemulus, Model Campuran, REML, AR.

***Abstract***

*Additive model is generalization of multiple regression that expresses the mean of a reponse variable which is modeled as a sum of functional form of predictors. Standard methods of additive models assume that errors are independent stochastic. This assumption is inappropiate for time series data where errors are autocorrellated. Standard methods of*

[1] Jurusan Statistika FMIPA ITS (e-mail : aniksb@plasa.com)
[2] Departemen Statistika FMIPA IPB

*additive model need to be justified when it is used for data with autocorrelated errors. Simulation results show that the time series additive model produce higher smoothing parameter estimator and smaller sum square errors than the standard additive models. This results show that additive model is sensitive to errors asumption.*

***Keywords*** *: Additive Model, Spline Regression, Penalized Spline Regression, Smoothing Parameters, Linear Mixed Models, REML, AR*

## 1. Pendahuluan

Teknik pemulusan nonparametrik digunakan untuk memodelkan hubungan antar peubah tanpa penetapan bentuk khusus tentang fungsi regresinya. Dalam bentuk sederhana, hubungan antara respon $y$ dengan peubah penjelas $x$ dinyatakan dalam model $y_i = s(x_i) + \varepsilon_i$ untuk $i = 1, 2, \ldots, n,$ dimana $s(.)$ adalah bentuk hubungan fungsional nonparametrik dan $\varepsilon_i$ adalah galat acak. Model sederhana ini oleh Stone (1985) dikembangkan untuk peubah penjelas lebih dari satu seperti pada regresi berganda ke dalam bentuk yang fleksibel dan dikenal dengan model aditif. Untuk data $\{(y_i, x_{i1}, \cdots, x_{ip}),\ i = 1, \cdots, n\}$, model aditif didefinisikan sebagai

$$y_i = s_0 + \sum_{j=1}^{p} s_j(x_{ij}) + \varepsilon_i$$

dimana $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_n)'$ bebas stokastik terhadap peubah penjelas $\mathbf{x}_j = (x_{1j}, \cdots, x_{nj})'$, dan memenuhi $E(\varepsilon) = \mathbf{0}$, $\text{cov}(\varepsilon) = \sigma^2\mathbf{I}$, sedangkan $s_j$ adalah bentuk hubungan fungsional antara respon dengan peubah penjelas $\mathbf{x}_j$. Bentuk hubungan fungsional antara peubah penjelas dengan respon dapat parametrik, nonparametrik, atau gabungan keduanya. Setiap bentuk fungsional peubah penjelas dinyatakan secara terpisah, sehingga model tetap mempertahankan interpretasi penting dari model linear.

Hastie dan Tibshirani (1990) memperluas model aditif untuk sebaran keluarga eksponensial menjadi model aditif terampat (*generalized additive model* selanjutnya disingkat GAM). Metode pendugaan pada GAM yang terkenal adalah algoritma *backfitting*. Meskipun GAM bersifat fleksibel dan efisien, akan tetapi algoritma *backfitting* dengan pemulus linear mempunyai kesulitan dalam seleksi model dan penarikan kesimpulan. Kesulitan pada GAM ini dipecahkan oleh Gu dan Wahba (1991) dengan pemulus spline terampat (*generalized smoothing spline* selanjutnya disingkat GSS). Sayangnya perhitungan GSS mendekati $n^3$, dengan n adalah banyaknya data yang akan dimodelkan.

Pada akhirnya muncul pendekatan *spline* terpenalti atau P-*spline* (*penalized spline regression*) dengan menggunakan pemulus dimensi rendah (Eilers dan Mark,1996, Ruppert dan Carroll, 1997). Metode ini menggunakan basis pemulus dengan lokasi simpul yang tetap dan jumlah simpul yang lebih kecil dari data. Faktor pendorong utama pendekatan P-*spline* adalah reduksi komputasi untuk data berukuran besar dan model aditif yang mengandung banyak fungsi pemulus secara simultan. Di samping itu P-*spline* mempunyai hubungan matematis yang sederhana dengan model linear campuran. Metode ini dapat diformulasikan sebagai penduga kemungkinan maksimum dalam kerangka model linear campuran. Efek samping yang menarik dari hubungan ini adalah parameter pemulus berhubungan secara langsung dengan komponen ragam dari model linear campuran.

Metode baku pada model aditif mengasumsikan bahwa galat acak bebas stokastik. Akan tetapi pada beberapa aplikasi sering dijumpai bahwa beberapa jenis data berkorelasi, seperti data deret waktu dan data spatial. Sehingga asumsi galat acak menyebar bebas stokastik dan identik tidak sesuai untuk data deret waktu dimana galat acak saling berkorelasi. Bila model aditif baku digunakan pada data yang berkorelasi, maka akan dihasilkan komponen dugaan yang tidak efisien. Pada tulisan ini akan dikaji tentang pendugaan model aditif pada data deret waktu dengan pendekatan model linear campuran. Data yang digunakan ada dua macam, yaitu data simulasi dan data konsentrasi pencemar udara Ozon di kota Surabaya.

## 2. Model Aditif dengan Satu Peubah Penjelas

Misalkan $(x_i, y_i)$ adalah pengukuran pada peubah penjelas $x$ dan peubah respon $y$ untuk $1 \leq i \leq n$. Misalkan hubungan fungsional antara $x$ dengan $y$ dimodelkan sebagai

$$y_i = s(x_i) + \varepsilon_i \tag{1}$$

dengan $s$ adalah fungsi mulus, $\varepsilon_i$ bebas stokastik dengan ragam $\sigma^2$. Model (1) adalah bentuk regresi nonparametrik yang paling sederhana dan banyak metode pendekatan yang dapat digunakan seperti yang dibahas oleh Eubank (1988), Green dan Silverman (1994), dan Simonoff (1996) Misalkan fungsi mulus $s$ diduga dengan model regresi spline berderajat-p yaitu :

$$s(\mathbf{x};\boldsymbol{\beta}) = \beta_0 + \beta_1 x + \ldots + \beta_p x^p + \sum_{k=1}^{K} u_{pk}(x - \kappa_k)_+^p$$

dengan $\boldsymbol{\beta} = (\beta_0, \ldots \beta_p, u_{p1}, \ldots u_{pK})'$ adalah vektor koefisien regresi spline, $p \geq 1$ adalah bilangan bulat positif, $(w)_+^p = w^p \mathbf{I}(w \geq 0)$ adalah basis fungsi pangkat terpotong berderajat p (*truncated power function* selanjutnya disingkat dengan FPT) dengan $\mathbf{I}$ fungsi indikator, dan $\kappa_1 < \ldots < \kappa_K$ adalah simpul tetap.

Penduga parameter $\hat{\boldsymbol{\beta}}$ ditentukan dengan minimisasi jumlah kuadrat terpenalti, yaitu J(*s*) yang didefinisikan sebagai :

$$J(s) = \sum_{i=1}^{n} (y_i - s(x_i; \boldsymbol{\beta}))^2 + \lambda\ \boldsymbol{\beta}'\ \mathbf{D}\ \boldsymbol{\beta} \qquad (2)$$

dengan $\lambda$ adalah parameter pemulus, dan $\mathbf{D}$ adalah matriks semidefinit positif. Suku pertama pada J(*s*) adalah jumlah kuadrat galat dan suku keduanya adalah penalti kekasaran. Kriteria penentuan model pada persamaan (2) merupakan gabungan antara kriteria pada model regresi *spline* dengan kriteria dari pemulus *spline*. Sehingga minimisasi J(*s*) pada nilai $\lambda$ tertentu akan memberikan kompromi antara kebaikan pengepasan dengan kehalusan kurva. Model aditif degan kriteria pendugaan ini disebut juga dengan regresi *spline* terpenalti.

Misalkan **T** adalah matriks desain untuk regresi *spline* dengan baris ke-*i* dari matriks **T** yaitu :

$$\mathbf{T}_i = \left(1,\ x_i,\ \ldots\ x_i^p,\ (x_i - \kappa_1)_+^p,\ \ldots\ (x_i - \kappa_K)_+^p\right)$$

Misalkan $\mathbf{D} = \mathrm{diag}(\mathbf{0}_{p+1}, \mathbf{1}_K)$ maka dalam notasi matriks J(*s*) dapat dinyatakan sebagai

$$(\mathbf{y} - \mathbf{T}\ \boldsymbol{\beta})'(\mathbf{y} - \mathbf{T}\ \boldsymbol{\beta}) + \lambda\ \boldsymbol{\beta}'\ \mathbf{D}\ \boldsymbol{\beta} \qquad (3)$$

Minimisasi persamaan (3) menghasilkan penduga bagi parameter $\boldsymbol{\beta}$ yaitu

$$\hat{\boldsymbol{\beta}} = (\mathbf{T}'\ \mathbf{T} + \lambda\ \mathbf{D})^{-1}\mathbf{T}'\ \mathbf{y}$$

[1] Jurusan ...
[2] Departemen Statistika ...

sehingga penduga regresi *spline* terpenalti adalah

$$\hat{\mathbf{y}} = \mathbf{T}\hat{\boldsymbol{\beta}} = \mathbf{T}(\mathbf{T}'\mathbf{T} + \lambda\ \mathbf{D})^{-1}\mathbf{T}'\mathbf{y} \qquad (4)$$

Regresi *spline* terpenalti mempunyai hubungan matematis yang sederhana dengan model linear campuran seperti yang dibahas oleh Vebyla *et al* (1999), Wang (1998), Fan dan Zhang (1998), Brumback *et al* (1999), French *et al* (2001), Kamman and Wand (2003), dan Wand (2003). Metode ini dapat diformulasikan sebagai model linear campuran. Kunci hubungan antara regresi *spline* terpenalti dengan model linear campuran adalah bahwa penalti dari koefisien $u_{pk}$ pada model (1) ekivalen dengan memperlakukan koefisien ini sebagai efek acak pada model linear campuran. Misalkan didefinisikan parameter

$$\boldsymbol{\beta}^* = (\beta_0, \beta_1, \cdots, \beta_p)', \quad \mathbf{u} = (u_{p1}, \ldots, u_{pK})',$$

dan matriks desain

$$\mathbf{X} = \begin{bmatrix} 1 & x_1 & \cdots & x_1^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & x_1 & \cdots & x_n^p \end{bmatrix}, \text{dan}$$

$$\mathbf{Z} = \begin{bmatrix} (x_1 - \kappa_1)_+^p & \cdots & (x_1 - \kappa_K)_+^p \\ \vdots & \ddots & \vdots \\ (x_n - \kappa_1)_+^p & \cdots & (x_n - \kappa_K)_+^p \end{bmatrix}$$

Kriteria *spline* terpenalti pada persamaan (3) jika dibagi dengan $\sigma_\varepsilon^2$ akan diperoleh

$$\frac{1}{\sigma_\varepsilon^2}(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^* - \mathbf{Z}\ \mathbf{u})'(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}^* - \mathbf{Z}\mathbf{u}) + \frac{\lambda}{\sigma_\varepsilon^2}\ \mathbf{u}'\mathbf{u} \qquad (5)$$

terpenalti. Parameter pemulus $\lambda$ merupakan rasio antara dua komponen ragam dari model linear campuran pada persamaan (6)

Bila galat $\varepsilon_i$ pada model (1) mengikuti proses autoregresif (AR) ordo-1 yaitu

$$\varepsilon_i = \rho\varepsilon_{i-1} + \delta_i \tag{9}$$

dengan galat acak $\boldsymbol{\delta} = (\delta_1, \cdots, \delta_n)'$ bebas stokastik dan identik dimana $E(\boldsymbol{\delta}) = 0$, dan $\text{cov}(\boldsymbol{\delta}) = \sigma_\delta^2 \mathbf{I}$. Sehingga $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ dan $\text{cov}(\boldsymbol{\varepsilon}) = \sigma_\delta^2 \mathbf{W}$ dengan

$$\mathbf{W} = \left(1-\rho^2\right)^{-1} \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{T-1} \\ \rho & 1 & \rho & \cdots & \rho^{T-2} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ \rho^{T-1} & \rho^{T-2} & \rho^{T-3} & \cdots & 1 \end{pmatrix}$$

maka kriteria *spline* terpenalti pada persamaan (3) untuk galat berkorelasi menjadi

$$(\mathbf{y} - \mathbf{T}\,\boldsymbol{\beta})' \mathbf{W}^{-1} (\mathbf{y} - \mathbf{T}\,\boldsymbol{\beta}) + \lambda\, \boldsymbol{\beta}' \mathbf{D}\, \boldsymbol{\beta} \tag{10}$$

Kriteria pada persamaan (10) ekivalen dengan BLUP dari model linear campuran

$$\mathbf{y} = \mathbf{X}\,\boldsymbol{\beta}^* + \mathbf{Z}\,\mathbf{u} + \boldsymbol{\varepsilon}$$

dengan

$$E\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \mathbf{0} \text{ dan } \text{cov}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma_u^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \sigma_\varepsilon^2 \mathbf{W} \end{bmatrix} \tag{11}$$

dengan vektor parameter dan matriks desain sama dengan model linear campuran pada persamaan (6). Pendugaan parameter pada model (11) dilakukan dengan REML (*restricted maximum likelihood*).

**3. Model Aditif**

Misalkan model aditif deret waktu untuk Ozon dimodelkan sebagai

$$y_i = \beta_0 + f(t_i) + g(m_i) + \varepsilon_i \tag{12}$$

dimana f($t$) dan g($m$) adalah fungsi mulus untuk peubah penjelas waktu dan faktor meteorologis. Fungsi mulus $f$ dan $g$ dimodelkan sebagai regresi *spline* linear berderajat-1 yaitu:

$$y_i = \beta_0 + \beta_t t_i + \sum_{k=1}^{K_t} u_k^t (t_i - \kappa_{k_t}^t)_+ + \beta_m m_i + \sum_{k=1}^{K_m} u_k^m (m_i - \kappa_{k_m}^m)_+ + \varepsilon_i \tag{13}$$

dan diduga dengan minimisasi jumlah kuadrat terpenalti. Misalkan vektor koefisien regresi *spline* adalah

$$\boldsymbol{\beta} = \left(\beta_0, \beta_t, \beta_m, u_1^t, \cdots, u_{K_t}^t, u_1^m, \cdots, u_{Km}^m\right)',$$

sedangkan simpul tetap untuk peubah penjelas waktu dan faktor meteorologis masing-masing adalah $\kappa_1^t < \ldots < \kappa_{K_t}^t$ dan $\kappa_1^m < \ldots < \kappa_{K_m}^m$. Pendugaan model aditif dengan peubah penjelas lebih dari satu dapat diturunkan dengan cara yang sama seperti pada model aditif dengan satu peubah penjelas seperti yang telah diuraikan di atas. Bila galat pada model (12) bebas stokastik, maka penduga regresi *spline* terpenalti dari model (13) diturunkan melalui generalisasi persamaan (4), dan diperoleh

$$\hat{\mathbf{y}} = \mathbf{C}(\mathbf{C}'\mathbf{C} + \mathbf{\Lambda})^{-1}\mathbf{C}'\mathbf{y} \tag{14}$$

dengan

$$\mathbf{C} = \begin{bmatrix} 1 & t_1 & m_1 & (t_1-\kappa_1^t)_+ & \cdots & (t_1-\kappa_{K_t}^t)_+ & (m_1-\kappa_1^m)_+ & \cdots & (m_1-\kappa_{K_m}^m)_+ \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & t_n & m_n & (t_n-\kappa_1^t)_+ & \cdots & (t_n-\kappa_{K_t}^t)_+ & (m_n-\kappa_1^m)_+ & \cdots & (m_n-\kappa_{K_m}^m)_+ \end{bmatrix}$$

dan

$$\mathbf{\Lambda} = \mathrm{diag}\left(0,\ 0,\ 0,\ \lambda_t \mathbf{1}_{K\times 1},\ \lambda_m \mathbf{1}_{K\times 1}\right),$$

dimana $\lambda_t$ dan $\lambda_m$ masing-masing adalah parameter pemulus peubah penjelas waktu dan faktor meteorologis.

Formulasi regresi *spline* terpenalti pada persamaan (13) ke dalam model linear campuran adalah dengan memperlakukan potongan polinomial $u_k^t$ dan $u_k^s$ sebagai efek acak dalam model linear campuran. Jika didefinisikan

$$\boldsymbol{\beta} = (\beta_0, \beta_t, \beta_m)', \quad \mathbf{u} = \left(u_1^t, \cdots, u_{K_t}^t, u_1^m, \cdots, u_{K_m}^m\right)'$$

$$\mathbf{X} = \begin{bmatrix} 1 & t_1 & m_1 \\ \vdots & \vdots & \vdots \\ 1 & t_n & m_n \end{bmatrix},$$

$$\mathbf{Z} = \begin{bmatrix} (t_1-\kappa_1^t)_+ & \cdots & (t_1-\kappa_{K_t}^t)_+ & (m_1-\kappa_1^m)_+ & \cdots & (m_1-\kappa_{K_m}^m)_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (t_n-\kappa_1^t)_+ & \cdots & (t_n-\kappa_{K_t}^t)_+ & (m_n-\kappa_1^m)_+ & \cdots & (s_n-\kappa_{K_m}^m)_+ \end{bmatrix}$$

maka penduga kuadrat terkecil terpenalti adalah ekivalen dengan BLUP model linear campuran

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}$$

dengan

$$\mathrm{E}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \mathbf{0} \text{ dan } \mathrm{cov}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma_t^2\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_m^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\varepsilon^2\mathbf{I} \end{bmatrix} \quad (15)$$

Sehingga pendugaan model aditif dapat dilakukan dengan menggunakan model linear campuran. Penduga parameter pemulus pada regresi *spline* terpenalti untuk waktu dan faktor meteorologis merupakan rasio antara dua komponen ragam yaitu

$$\lambda_t = \frac{\sigma_\varepsilon^2}{\sigma_t^2} \text{ dan } \lambda_m = \frac{\sigma_\varepsilon^2}{\sigma_m^2}.$$

Bila galat $\varepsilon_i$ pada model (13) mengikuti proses autoregresif (AR) ordo-1 seperti pada persamaan (9), maka formulasi model linear campuran merupakan perluasan persamaan (11) dan persamaan (15), yaitu

$$\mathbf{y} = \mathbf{X\beta} + \mathbf{Zu} + \boldsymbol{\varepsilon}$$

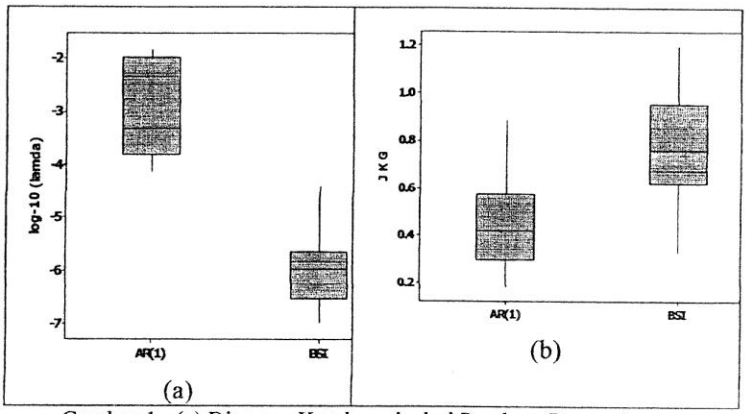
dengan

$$\mathrm{E}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \mathbf{0} \text{ dan } \mathrm{cov}\begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma_t^2\mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_m^2\mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\varepsilon^2\mathbf{W} \end{bmatrix} \quad (16)$$

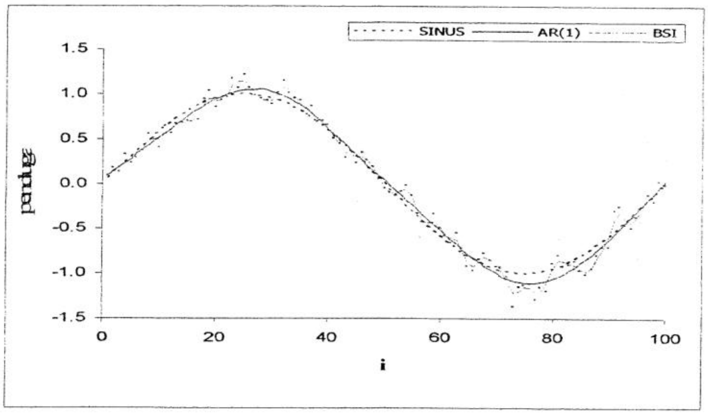
**4. Simulasi**

Pada data simulasi, dibangkitkan 100 pengamatan dari fungsi $y_i = \sin\left(\frac{2\pi i}{100}\right) + \varepsilon_i$ dengan $i$ = 1, 2, ..., 100. Galat $\varepsilon_i$ dimodelkan sebagai $\varepsilon_i = 0.5\varepsilon_{i-1} + \delta_i$ dengan $\delta_i$ menyebar normal nilai tengah 0 dan ragam 0.1. Galat acak $\delta_i$ dibangkitkan sebanyak 100 ulangan. Pada setiap ulangan dilakukan pendugaan model aditif dengan asumsi galat acak bebas stokastik (selanjutnya disingkat BSI) dan galat acak mengikuti proses AR(1). Pendugaan model aditif menggunakan pendekatan model linear campuran. Parameter pemulus $\lambda$ merupakan rasio antara dua komponen ragam dari model linear campuran pada persamaan (6). Kebaikan model dievaluasi dari jumlah kuadrat galat (JKG) yaitu $\mathrm{JKG} = \sum_{i=1}^{100} (y_i - \hat{y}_i)^2$ .

Perbandingan hasil pendugaan antara model aditif deret waktu dengan model aditif baku disajikan pada Gambar 1(a) dan Gambar 1(b). Pada Gambar 1(a) tampak penduga parameter pemulus pada model aditif AR(1) lebih tinggi dibandingkan dengan model BSI, sedangkan JKG model aditif AR(1) lebih rendah dibandingkan dengan model BSI. Nilai parameter pemulus yang rendah akan menghasilkan kurva yang kasar. Hal ini terlihat pada hasil prediksi pada Gambar 2. Prediksi model aditif AR(1) lebih mulus dibandingkan model aditif BSI.




Gambar 1. (a) Diagram Kotak-garis dari Penduga Parameter Pemulus
(b) Diagram Kotak-garis dari JKG




Gambar 2. Hasil Prediksi dengan model aditif AR(1) dan BSI

## 5. Penerapan pada Data Pencemar Udara Ozon

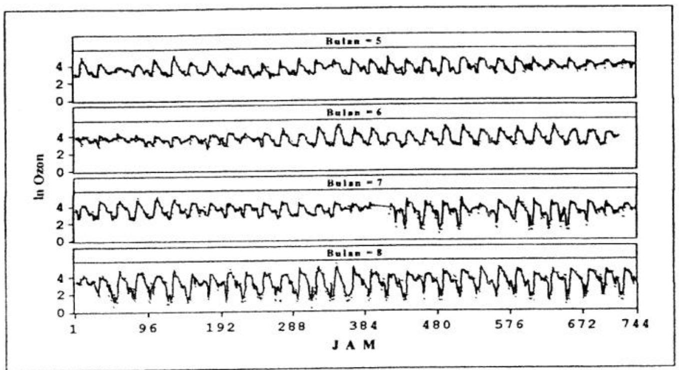
Ozon permukaan adalah pencemar sekunder yang terbentuk akibat reaksi gas nitrogen dengan hidrokarbon akibat pemanasan sinar matahari. Data konsentrasi ozon ($\mu g/m^3$) yang digunakan pada penelitian ini diukur oleh stasiun pemantau kualitas udara kota Surabaya. Lokasi pengamatan di Taman Prestasi Surabaya pada bulan Mei 2002 sampai Agustus 2002. Data ditransformasi logaritma agar mempunyai bentuk sebaran yang simetrik. Peubah penjelas yang disertakan dalam pemodelan adalah waktu, kecepatan angin, dan suhu udara.

Fungsi mulus pada model aditif dimodelkan dengan regresi *spline* terpenalti (P-*spline)* berderajat-1 seperti pada persamaan (13) Pendugaan model aditif menggunakan pendekatan model linear campuran. Dari ketiga peubah penjelas di atas, dan jumlah simpul yang digunakan, serta asumsi galat acak akan diperoleh beberapa alternatif model.

Hasil seleksi model menghasilkan model terbaik memberikan ragam galat sebesar 0.1084, $R^2$ = 98.39%, dan AIC (*Akaike Information Criteria*) sebesar 3814.4 Model terbaik ini menggunakan asumsi galat acak mengikuti proses AR(1). Penduga parameter pemulus dari setiap peubah penjelas disajikan pada Tabel 1. Hasil prediksi model pada bulan Mei sampai Agustus 2002 ditampilkan pada Gambar 3. Pada Gambar 3 tampak nilai prediksi model aditif deret waktu sangat dekat dengan data.

Tabel 1. Penduga Parameter Pemulus

| Peubah | Parameter Pemulus |
|---|---|
| Jam | 3.317647 |
| Kecepatan Angin | 1.015955 |
| Temperatur | 43.33699 |




Gambar 3. Plot tebaran data dan hasil pendugaan model aditif deret waktu

## 6. Kesimpulan

Pendugaan model aditif dengan pendekatan model linear campuran memberikan kemudahan dalam pemodelan data deret waktu, dimana asumsi galatnya dimodelkan sebagai komponen ragam pada model linear campuran. Pada data deret waktu sebaiknya menggunakan metode aditif deret waktu, karena model aditif peka terhadap asumsi kebebasan dari galat acak.

## 7. Daftar Pustaka

Brumback, B.A., Ruppert, D., dan Wand, M.P, 1999, Comment on Variable selection and function estimation in additive nonparametric regression using a data-based prior by Shively, Kohn and Wood, *Journal of the American Statistical Association*, **94**: 794-797.

Eilers, P.H.C. dan Marx, B.D., 1996, Flexible smoothing with B-splines and penalties (with discussion), *Statistical Science*, **11**: 89-121.

Eubank, R.L., 1988, *Spline Smoothing and Nonparametric Regression*, Marcel Deker, New York.

Fan, J. dan Zhang, J.T., 1998, Comment on Smoothing spline models for the analysis of nested and crossed samples of curves by Brumback and Rice, *Journal of the American Statistical Association*, **93**: 961-994.

French, J.L., Kammann, E.E., dan Wand, M.P., 2001, Comment on Semiparametric nonlinear mixed-effects models and their applications by Ke and Wang, *Journal of the American Statistical Association*, **96**:1285-1288.

Green, P.J. dan Silverman, B.W., 1994, *Nonparametric Regression and Generalized Linear Models : a Roughness Penalty Approach*, Chapman & Hall, London.

Gu, C., Wahba, G., 1991, Minimizing GCV/GML scores with multiple smoothing parameters via the Newton method, *SIAM J. Sci. Statist. Comput.*, **12**: 383–398.

Hastie, T.J. dan Tibshirani, R.J., 1990, *Generalized Additive Models*, Chapman and Hall, London.

Kammann, E.E. dan Wand, M.P., 2003, Geoadditive models, *Applied Statistics*, **52**:1-18.

Ruppert, D. dan Carroll, R.J., 1997, *Penalized regression splines*, Unpublished manuscript, [terhubung berkala], http://www.orie.cornell.edu/~davidr/papers/index/index.html.

Simonoff, J. S., 1996, *Smoothing Methods in Statistics*, Springer-Verlag, New York.

Stone, C.J., 1985, Additive Regression and Other Nonparametric Models, *The Annals of Statistics*, **13**, 689–705.

Verbyla, A.P., Cullis, B.R., Kenward, M.G., dan Welham, S.J., 1999, The analysis of designed experiments and longitudinal data by using smoothing splines (with discussion), *Journal of the Royal Statistics Society*, Series C **48**: 269-312.

Wand, M., 2003, Smoothing and mixed models, *Computational Statistics*, **18**:223–249.

Wang, Y., 1998, Smoothing spline models with correlated random errors, *Journal of the American Statistical Association*, **93**:341-348.

Diterima : Agustus 2005
Diterima untuk dipublikasikan : Desember 2005