

Keunggulan Pendugaan Model Aditif dengan Pendekatan Model Linear Campuran Dibanding dengan Algoritma *Backfitting*

ANIK DJURAI DAH

Jurusan Statistika FMIPA-ITS
e-mail : aniksb@plasa.com

ABSTRACT

The additive model is the generalized of multiple linear regression that expresses the mean of a response variable as a sum of functional form of predictors. The widely used estimation of additive models described in Hastie and Tibshirani (1990) is backfitting algorithm. However, the algorithm with linear smoothers gave some difficulties when it comes to model selection and its inference. The additive model with P-spline as smooth function admits a mixed model formulation, in which variance components control the non-linearity degree in the smooth function. This research is focused in comparing of estimation additive models using backfitting algorithm and linear mixed model approach. The research results show that estimation of additive models using linear mixed models offer simplicity in the computation, since use low-rank dimension of P-spline, and in the model inference, since based on maximum likelihood method. Estimation additive model using linear mixed model approach can be suggested as an alternative method beside backfitting algorithm

Key words : additive model, backfitting algorithm, P-spline, smoothing parameters, mixed models

1. Pendahuluan

Regresi 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

$$y_i = s(x_i) + \varepsilon_i \text{ untuk } i = 1, 2, \dots, n$$

dengan $s(\cdot)$ menyatakan bentuk hubungan fungsional nonparametrik dan ε_i adalah galat acak. Stone (1985) memperluas model sederhana ini untuk peubah penjelas lebih dari satu dan dikenal dengan model aditif. Untuk data $\{(y_i, x_{i1}, \dots, x_{id}), i = 1, \dots, n\}$. Model aditif didefinisikan sebagai

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

dengan ε bebas stokastik terhadap peubah penjelas \mathbf{X} , 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 s_j 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 GAM yang terkenal adalah algoritma *backfitting* dengan *local scoring*. Meskipun GAM bersifat fleksibel dan efisien, akan tetapi algoritma *backfitting* dengan pemulus linear mempunyai kesulitan dalam seleksi model dan penarikan kesimpulan.

Eilers dan Marx (1996), Ruppert dan Carroll (1997) mengemukakan pemulus dimensi rendah yang disebut dengan P-spline atau regresi spline terpenalti (*penalized spline regression*).

P-spline menggunakan jumlah basis yang sedikit dengan penalti kekasaran untuk mengontrol kemulusan dan mempunyai hubungan matematis yang sederhana dengan model linear campuran seperti yang dibahas oleh Fan dan Zhang (1998), Wang (1998), Brumback *et al.* (1999), Vebyla *et al.* (1999), French *et al.* (2001), Kamman dan Wand (2003), dan Wand (2003). Dalam model linear campuran, penduga P-spline adalah penduga tak bias linear terbaik (*best linear unbiased linear prediction* selanjutnya disingkat BLUP) dan parameter pemulusnya merupakan ratio dari dua komponen ragam.

P-spline dapat dipandang sebagai model aditif untuk satu peubah penjelas. Kajian empiris pada model aditif satu peubah penjelas dengan fungsi mulus P-spline dan spline menunjukkan bahwa perbedaan nilai parameter pemulus dan MSE (mean square error) antara keduanya relatif kecil (Djuraidah dan Aunuddin, 2006a). Model ini dapat dikembangkan untuk banyak peubah penjelas dan diduga dengan pendekatan model linear campuran (Djuraidah dan Aunuddin, 2006b). Perluasan P-spline pada dimensi dua diturunkan dari pemulus spline pada dimensi ganda yang dikenal dengan thin-plate spline dan dapat diduga dengan pendekatan model linear campuran (Djuraidah dan Aunuddin, 2006c).

Pada penelitian ini akan dikaji pendugaan model aditif dengan pendekatan model linear campuran dan algoritma backfitting. Kedua metode pendugaan ini akan dievaluasi tentang kebaikan modelnya, kemudahan komputasi, dan penarikan kesimpulannya.

2. Model Linear Campuran

Model linear campuran atau dikenal dengan komponen ragam merupakan perluasan dari model linear yaitu dengan menambahkan efek acak. Metode ini banyak digunakan dalam rancangan percobaan untuk data yang berkorelasi seperti pada percobaan pengukuran berulang (Searle *et al.*, 1992). Bentuk umum model linear campuran adalah

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \text{ dengan } \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} \sim \left(\begin{bmatrix} \mathbf{0} \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{G} & \mathbf{0} \\ \mathbf{0} & \mathbf{R} \end{bmatrix} \right) \quad (2)$$

dimana \mathbf{X} adalah matriks desain dari efek tetap yang teramati, $\boldsymbol{\beta}$ adalah vektor parameter pengaruh efek tetap yang tidak diketahui, \mathbf{Z} adalah matriks desain efek acak yang teramati, \mathbf{u} adalah vektor efek acak yang tidak diketahui, dan $\boldsymbol{\varepsilon}$ adalah vektor galat acak yang tidak diketahui. Sehingga nilai tengah dan matriks ragam-peragam untuk \mathbf{y} adalah $E(\mathbf{y}) = \mathbf{X}\boldsymbol{\beta}$ dan $\text{var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}' + \mathbf{R}$.

Penduga efek tetap dan efek acak ditentukan dari persamaan model campuran yaitu :

$$\begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix} \begin{bmatrix} \boldsymbol{\beta} \\ \mathbf{u} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}$$

Bila matriks \mathbf{G} dan \mathbf{R} diketahui, maka penduga bagi parameter $\boldsymbol{\beta}$ dan \mathbf{u} adalah :

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{u}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}'\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}'\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}'\mathbf{R}^{-1}\mathbf{Z} + \mathbf{G}^{-1} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}' \\ \mathbf{Z}' \end{bmatrix} \mathbf{R}^{-1}\mathbf{y}$$

Dari persamaan matriks di atas, penduga efek tetap $\boldsymbol{\beta}$ dapat dinyatakan sebagai

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1} \mathbf{X}'\mathbf{V}^{-1}\mathbf{y} \quad (3)$$

Penduga efek tetap $\hat{\boldsymbol{\beta}}$ adalah penduga GLS (*generalized least squares*) dan $\mathbf{X}\hat{\boldsymbol{\beta}}$ adalah BLUE (*best linear unbiased estimator*) untuk $\mathbf{X}\boldsymbol{\beta}$. Bila \mathbf{y} mempunyai sebaran normal, maka $\hat{\boldsymbol{\beta}}$ adalah MLE (*maximum likelihood estimator*). Penduga efek acak $\hat{\mathbf{u}}$ pada Persamaan (2) dapat dinyatakan sebagai

$$\hat{\mathbf{u}} = \mathbf{G}\mathbf{Z}'\mathbf{V}^{-1}(\mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}})$$

yang merupakan BLUP dari \mathbf{u} (Christensen, 1987).

Pada pendugaan komponen ragam, metode yang paling banyak digunakan adalah metode kemungkinan maksimum (*maximum likelihood* selanjutnya disingkat ML) dan metode kemungkinan maksimum berkendala (*restricted maximum likelihood* selanjutnya disingkat REML). Metode REML menghasilkan penduga takbias bagi komponen ragam, sedangkan ML menghasilkan penduga yang bias.

3. Pendugaan Model Aditif dengan Pendekatan Model Linear Campuran

Misalkan model aditif dengan 2 peubah penjelas x_1 dan x_2 dimodelkan sebagai

$$y_i = \beta_0 + s_1(x_{1i}) + s_2(x_{2i}) + \varepsilon_i \quad (4)$$

dimana $s_1(x_1)$ dan $s_2(x_2)$ masing-masing adalah fungsi mulus x_1 dan x_2 . Fungsi mulus s_1 dan s_2 dimodelkan sebagai regresi spline linear berderajat-1 yaitu:

$$y_i = \beta_0 + \beta_1 x_{1i} + \sum_{k=1}^{K_1} u_k^1 (x_i - \kappa_{k_1}^1)_+ + \beta_2 x_{2i} + \sum_{k=1}^{K_2} u_k^2 (x_i - \kappa_{k_2}^2)_+ + \varepsilon_i \quad (5)$$

Misalkan vektor koefisien regresi spline adalah $\beta = (\beta_0, \beta_1, \beta_2, u_1^1, \dots, u_{K_1}^1, u_1^2, \dots, u_{K_2}^2)$, sedangkan simpul tetap untuk peubah penjelas penjelas x_1 dan x_2 masing-masing adalah $\kappa_1^1 < \dots < \kappa_{K_1}^1$ dan $\kappa_1^2 < \dots < \kappa_{K_2}^2$, dan $(w)_+ = w \mathbf{I}(w \geq 0)$ adalah TPF (*truncated power function*) berderajat-1 dengan \mathbf{I} fungsi indikator.

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

$$J(s) = \sum_{i=1}^n \varepsilon_i^2 + \lambda \beta' \mathbf{D} \beta \quad (6)$$

dengan λ 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 (6) merupakan gabungan antara kriteria pada model regresi spline dengan kriteria dari pemulus spline. Sehingga minimisasi $J(s)$ pada nilai λ tertentu akan memberikan kompromi antara kebaikan pengepasan dengan kehalusan kurva. Model aditif dengan kriteria pendugaan ini disebut juga dengan regresi spline terpenalti. Minimisasi persamaan (6) menghasilkan penduga bagi parameter β yaitu

$$\hat{\beta} = (\mathbf{C}' \mathbf{C} + \lambda \mathbf{D})^{-1} \mathbf{C}' \mathbf{y}$$

sehingga penduga regresi spline terpenalti adalah

$$\hat{y} = \mathbf{C}(\mathbf{C}' \mathbf{C} + \lambda \mathbf{D})^{-1} \mathbf{C}' \mathbf{y} \quad (7)$$

dengan

$$\mathbf{C} = \begin{bmatrix} 1 & x_{11} & x_{21} & (x_{11} - \kappa_1^1)_+ & \cdots & (x_{11} - \kappa_{K_1}^1)_+ & (x_{21} - \kappa_1^2)_+ & \cdots & (x_{21} - \kappa_{K_2}^2)_+ \\ x: & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & (x_{1n} - \kappa_1^1)_+ & \cdots & (x_{1n} - \kappa_{K_1}^1)_+ & (x_{2n} - \kappa_1^2)_+ & \cdots & (x_{2n} - \kappa_{K_2}^2)_+ \end{bmatrix}$$

dan $\mathbf{D} = \text{diag}(0, 0, 0, \lambda_1 \mathbf{1}_{K_1}, \lambda_2 \mathbf{1}_{K_2})$, dimana λ_1 dan λ_2 masing-masing adalah parameter pemulus peubah penjelas penjelas x_1 dan x_2 .

Formulasi regresi spline terpenalti pada Persamaan (5) ke dalam model linear campuran adalah dengan memperlakukan potongan polinomial u_k^1 dan u_k^2 sebagai efek acak dalam model linear campuran. Jika didefinisikan

$$\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)', \quad \mathbf{u} = (u_1^1, \dots, u_{K_1}^1, u_1^2, \dots, u_{K_2}^2)'$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} \\ \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} \end{bmatrix},$$

$$\mathbf{Z} = \begin{bmatrix} (x_{11} - \kappa_1^1)_+ & \cdots & (x_{11} - \kappa_{K_1}^1)_+ & (x_{21} - \kappa_1^2)_+ & \cdots & (x_{21} - \kappa_{K_2}^2)_+ \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ (x_{1n} - \kappa_1^1)_+ & \cdots & (x_{1n} - \kappa_{K_1}^1)_+ & (x_{2n} - \kappa_1^2)_+ & \cdots & (x_{2n} - \kappa_{K_2}^2)_+ \end{bmatrix}$$

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

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \quad \text{dengan} \quad E \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \mathbf{0} \quad \text{dan} \quad \text{cov} \begin{bmatrix} \mathbf{u} \\ \boldsymbol{\varepsilon} \end{bmatrix} = \begin{bmatrix} \sigma_1^2 \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_2^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \sigma_\varepsilon^2 \mathbf{I} \end{bmatrix} \quad (8)$$

Sehingga pendugaan model aditif dapat dilakukan dengan menggunakan model linear campuran. Penduga parameter pemulus x_1 dan x_2 merupakan rasio antara dua komponen ragam yaitu

$$\lambda_1 = \frac{\sigma_\varepsilon^2}{\sigma_1^2} \quad \text{dan} \quad \lambda_2 = \frac{\sigma_\varepsilon^2}{\sigma_2^2}.$$

4. Pendugaan Model Aditif dengan Algoritma *Backfitting*

Pendugaan fungsi mulus $s_0, s_1(\cdot), \dots, s_d(\cdot)$ pada model aditif (1) dengan algoritma *backfitting* menggunakan galat parsial $R_j = Y - s_0 - \sum_{k \neq j} s_k(X_k)$ sehingga $E(R_j | X_j) = s_j(X_j)$ dengan $E(s_j(X_j)) = 0$.

Tahapan algoritma ini sebagai berikut:

1. Nilai awal : $s_0 = E(Y)$, $s_1^{(0)} = \dots = s_d^{(0)} = 0$, $m = 0$
2. Iterasi : $m = m + 1$

untuk $j=1$ sampai d , hitung

$$R_j = Y - s_0 - \sum_{k \neq j} s_k^{(m)}(X_k)$$

$$s_j^{(m)} = E(R_j | X_j)$$

3. Ulangi tahap 2 sampai $\text{JKG} = \left[Y - s_0 - \sum_{j=1}^d s_j^{(m)}(X_j) \right]^2$ tidak turun

5. Bahan dan Metode

PM₁₀ adalah partikulat debu dengan ukuran diameter aerodinamik kurang dari 10 mikron. Partikulat ini merupakan campuran yang sangat rumit dari berbagai senyawa organik dan anorganik. Pengaruh partikulat debu yang berada di udara terhadap kualitas lingkungan sangat tergantung kepada ukurannya. Ukuran partikulat debu yang membahayakan kesehatan

umumnya berkisar antara 0.1 mikron sampai dengan 10 mikron. Pada umumnya partikulat debu yang berukuran sekitar 5 mikron dapat langsung masuk kedalam paru-paru dan mengendap di alveoli, sedangkan partikulat yang lebih besar dari 5 mikron dapat mengganggu saluran pernafasan bagian atas dan menyebabkan iritasi. Partikulat debu yang melayang dan berterbangan dibawa angin akan menyebabkan iritasi pada mata dan dapat menghalangi daya tembus pandang mata.

Data konsentrasi PM_{10} ($\mu g/m^3$) yang digunakan pada penelitian ini diukur oleh 5 stasiun pemantau kualitas udara ambien di kota Surabaya. Data yang diteliti mulai dari Januari 2002 sampai Desember 2002 dengan selang waktu pengukuran satu jam. Faktor meteorologis yang diikutsertakan dalam pemodelan adalah kecepatan angin, arah angin, dan curah hujan.

Hasil eksplorasi pada data PM_{10} menunjukkan bahwa dalam data terdapat korelasi temporal dan korelasi spasial. Kedua macam korelasi ini harus diperhitungkan dalam pemodelannya. Pada penelitian ini korelasi temporal dimodelkan dalam model aditif deret waktu dan korelasi spasial dimodelkan dalam model aditif spasial. Selanjutnya kedua model aditif ini digabung secara aditif menjadi model aditif spatio-temporal. Sehingga model aditif spatio-temporal dapat menggabungkan kedua korelasi yang terdapat pada data dalam satu model. Pemodelan data spatio-temporal dengan pendekatan model linear campuran dibahas selengkapnya dalam Djuraidah dan Aunuddin (2006d).

Pada penelitian ini akan dilakukan pendugaan model aditif pada data PM_{10} dengan pendekatan model linear campuran dan algoritma *backfitting*. Kedua metode dievaluasi nilai MSE (*mean square error*), kemudahan dalam penarikan kesimpulan, dan waktu komputasinya. Analisis data menggunakan paket program SAS v9.1.

6. Hasil dan Pembahasan

Hasil Pendugaan Model Aditif dengan Pendekatan Model Linear Campuran

Hasil pendugaan model aditif spatio-temporal pada data berukuran satu tahun disarikan pada Tabel 1. Model spatio-temporal dengan faktor meteorologis (Model-3 sampai Model-8) mempunyai nilai AIC yang lebih kecil dibandingkan dengan model spatio-temporal tanpa faktor meteorologis (Model-1 dan Model-2). Penambahan peubah kecepatan angin pada Model-1 dan Model-2 dapat menurunkan AIC model. Demikian juga penambahan peubah arah angin pada Model-3 dan Model-4 dapat menurunkan AIC model. Penambahan peubah indikator curah hujan ke dalam Model-5 dan Model-6 juga menurunkan AIC model.

Nilai parameter pemulus spasial cenderung lebih rendah pada model dengan faktor meteorologis dibandingkan dengan model tanpa faktor meteorologis. Hal ini disebabkan penambahan faktor meteorologis menurunkan MSE. Sedangkan perbedaan nilai parameter pemulus spasial antara kriging dan spline-2 disebabkan oleh perbedaan fungsi basis radialnya.

Model-7 dan Model-8 mempunyai nilai AIC terkecil dan hampir sama, juga mempunyai nilai parameter pemulus Lag1-PM, Lag2-PM, Jam yang sama, serta MSE yang sama, sedangkan nilai parameter pemulus spasial berbeda. Fungsi korelasi spasial, ACF, dan PACF dari galat Model-7 dan Model-8 hampir sama, sehingga untuk menentukan model terbaik ditentukan dari pola ragam spasialnya. Pola ragam spasial berhubungan dengan nilai parameter pemulus spasial.

Pola ragam spasial dari Model-7 dan Model-8 masing-masing disajikan pada Gambar 1(a) dan 1(b). Pada Gambar 1(b) tampak pola ragam spasial dari kriging pada lokasi yang tidak diamati ditentukan oleh jarak terhadap lokasi stasiun pemantau, sedangkan pola ragam spasial pada spline-2 pada gambar 1(a) tampak homogen. Dengan demikian kriging dapat menggambarkan keragaman spasial dari pada spline-2

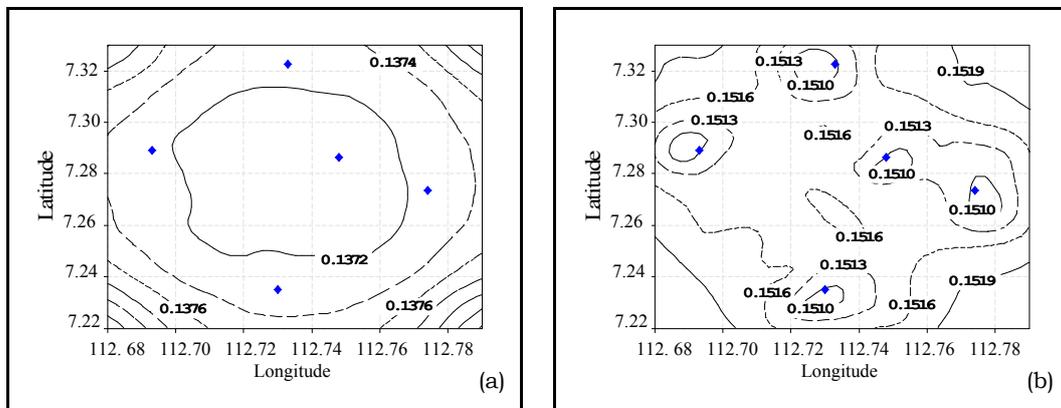
Tabel 1. Nilai AIC, MSE dan parameter pemulus dari model aditif spatio-temporal PM₁₀

Model	Komponen Model	AIC	MSE	Nilai Parameter Pemulus (λ)			
				Lag1-PM	Lag2-PM	Jam	Spatial
1	AAR(2) + Spline-2(K=5)	32291.0	0.145	1.68	5.39	3.98	20.03
2	AAR(2) + Kriging(K=5)	32289.8	0.145	1.68	5.39	3.98	3.54
3	AAR(2) + Spline-2(K=5) + KA	31302.4	0.141	1.76	5.22	3.98	15.99
4	AAR(2) + Kriging(K=5) + KA	31302.3	0.141	1.76	5.22	3.98	2.62
5	AAR(2) + Spline-2(K=5) + KA + AA(I)	30574.8	0.138	1.85	5.25	4.22	17.98
6	AAR(2) + Kriging(K=5) + KA + AA(I)	30574.6	0.138	1.85	5.25	4.22	2.94
7	AAR(2) + Spline-2(K=5) + KA + AA(I) + HJ(I)	30245.1	0.137	1.91	5.50	4.26	17.81
8	AAR(2) + Kriging(K=5) + KA + AA(I) + HJ(I)	30244.7	0.137	1.91	5.50	4.26	2.93

Keterangan :

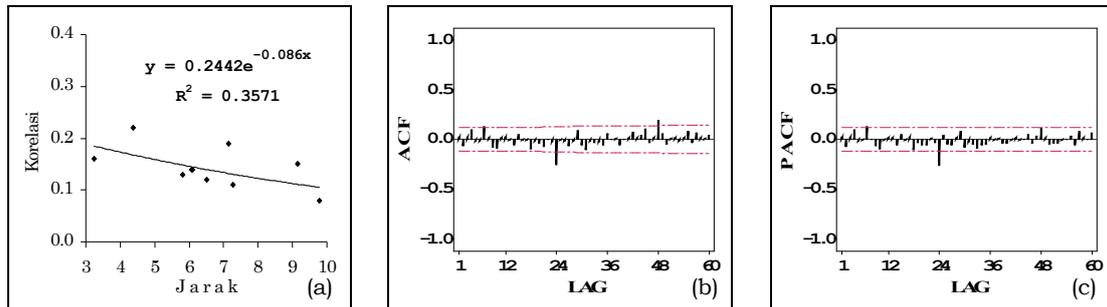
AAR(2) = Lag1-PM(K=5) + Lag2-PM(K=5) + Jam(K=23)

K= jumlah simpul, KA = (Lag-1 Kecepatan Angin)², AA = Arah Angin, HJ = Hujan, I = Peubah Indikator



Gambar 1. Pola ragam spasial untuk PM₁₀ (a) Model-7 dan (b) Model-8

Plot korelasi spasial, ACF, dan PACF dari galat Model-8 masing-masing disajikan pada Gambar 2(a), 2(b), dan 2(c). Pada Gambar 2(a) tampak nilai korelasi spasial dari galat Model-8 sudah kecil, meskipun masih mempunyai pola eksponensial akan tetapi nilai R² sudah kecil. Pada Gambar 2(b) dan 2(c) tampak plot ACF dan PACF dari galat Model-8 mempunyai nilai autokorelasi yang kecil dan sudah tidak berpola. Dengan demikian Model-8 adalah model aditif spatio-temporal terbaik untuk PM₁₀.



Gambar 2. (a) Plot korelasi spasial, (b) Plot ACF, dan (c) Plot PACF dari galat Model-8

Hasil Pendugaan Model Aditif dengan Algoritma *Backfitting*

Algoritma *backfitting* hanya mampu menganalisa data berukuran dua bulan dengan fungsi mulus spasial spline-2. Hasil pemodelan untuk data dua bulan dengan algoritma *backfitting* dan model linear campuran masing-masing disarikan pada Tabel 2 dan Tabel 3. Nilai MSE pada algoritma *backfitting* lebih tinggi dari pada model linear campuran. Nilai parameter pemulus pada algoritma *backfitting* tampak sama untuk semua model, sedangkan pada model linear campuran tampak berbeda antar model.

Tabel 2. Hasil pendugaan model aditif dengan algoritma *backfitting* untuk data dua bulan

Model	Komponen Model	MSE	Nilai Parameter Pemulus (λ)			
			Lag1-PM	Lag2-PM	Jam	Spatial
1	AAR(2) + Spline-2(K=5)	0.151	1	1	1	1.70E-11
2	AAR(2) + Spline-2(K=5) + KA	0.141	1	1	1	1.70E-11
3	AAR(2) + Spline-2(K=5) + KA + AA(I)	0.134	1	1	1	1.70E-11
4	AAR(2) + Spline-2(K=5) + KA + AA(I) + HJ(I)	0.134	1	1	1	1.70E-11

Keterangan :

AAR(2) = Lag1-PM(K=5) + Lag2-PM(K=5) + Jam(K=23)

K= jumlah simpul, KA = (Lag-1 Kecepatan Angin)², AA = Arah Angin, HJ = Hujan, I = Peubah Indikator

Tabel 3. Hasil pendugaan model aditif dengan model linear campuran untuk data dua bulan

Model	Komponen Model	AIC	MSE	Nilai Parameter Pemulus (λ)			
				Lag1-PM	Lag2-PM	Jam	Spatial
1	AAR(2) + Spline-2(K=5)	6144.9	0.142	1.72	3.65	3.35	26.96
2	AAR(2) + Spline-2(K=5) + KA	5854.3	0.135	1.70	3.45	3.37	9.69
3	AAR(2) + Spline-2(K=5) + KA + AA(I)	5536.8	0.128	1.97	3.53	3.61	12.12
4	AAR(2) + Spline-2(K=5) + KA + AA(I) + HJ(I)	5487.3	0.127	1.84	3.46	3.61	11.60

Keterangan :

AAR(2) = Lag1-PM(K=5) + Lag2-PM(K=5) + Jam(K=23)

K= jumlah simpul, KA = (Lag-1 Kecepatan Angin)², AA = Arah Angin, HJ = Hujan, I = Peubah Indikator

Waktu Komputasi

Perbandingan waktu komputasi dari kedua metode pendugaan model aditif untuk data dua bulan disajikan pada Tabel 4. Algoritma *backfitting* tidak mampu mengolah data berukuran besar seperti yang dilakukan oleh model linear campuran. Waktu komputasi model aditif dengan pendekatan model linear campuran lebih cepat dibandingkan dengan algoritma *backfitting*.

Tabel 4. Waktu komputasi pendugaan model aditif dengan model linear campuran dan algoritma *backfitting*

Model	Model linear campuran ^{*)}		Algoritma <i>backfitting</i> ^{*)}	
	real time	cpu time	real time	cpu time
1	2 menit 18.3 detik	1 menit 7 detik	2.30 detik	1.30 detik
2	3 menit 10.4 detik	2 menit 10.9 detik	2.64 detik	1.37 detik
3	4 menit 8.7 detik	2 menit 52. detik	3.02 detik	1.53 detik
4	4 menit 5.9 detik	2 menit 58 detik	2.92 detik	1.49 detik

^{*)} menggunakan software SAS v9.1, jumlah pengamatan 7320

7. Kesimpulan

Pendugaan model aditif dengan pendekatan model linear campuran mempunyai keunggulan, yaitu model bersifat fleksibel, berbasis metode kemungkinan maksimum, waktu komputasi yang cepat untuk data berukuran besar, dan memperluas metode pemulusan pada data spatial dengan kriging. Pendugaan model aditif dengan pendekatan model linear campuran dapat dipertimbangkan sebagai metode alternatif untuk pendugaan model aditif yang selama ini menggunakan algoritma *backfitting*.

Daftar Pustaka

- [1]. 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. *J Amer Stat Ass* 94: 794-797.
- [2]. Christensen, R. 1984. *Plane Answers to Complex Questions. The Theory of Linear Models*. New York : Springer-Verlag.
- [3]. Djuraidah, A., dan Aunuddin. 2006a. Pendugaan Regresi Spline Terpenalti dengan Pendekatan Model Linear Campuran. *Statistika FMIPA-UNISBA* 6(1): 39-46.
- [4]. Djuraidah, A., dan Aunuddin. 2006b. Pendugaan Model Aditif untuk Data Deret Waktu dengan Pendekatan Model Linear Campuran. *Inferensi Jurnal Statistika FMIPA-ITS* 2(1):76-92.
- [5]. Djuraidah, A., dan Aunuddin. 2006c. Kriging dan Thin-Plate Spline dengan Pendekatan Model Campuran. *Matematika Integratif Jurnal Matematika FMIPA-UNPAD* 5(2):1-12.
- [6]. Djuraidah, A., dan Aunuddin. 2006d. *Estimation of Spatio-temporal Additive Model Using Linear Mixed Model Approach with Application to Ozone Data in Surabaya*. Proceedings of The first International Conference on Mathematics and Statistics : 151-157
- [7]. Eilers, P. H. C, dan Marx, B. D. 1996. Flexible smoothing with B-splines and penalties (with discussion). *Stat Sci* 11:89-121.
- [8]. 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. *J Amer Stat Ass* 93: 961-994.
- [9]. 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. *J Amer Stat Ass* 96:1285-1288.
- [10]. Hastie, T. J., dan Tibshirani, R. J. 1990. *Generalized Additive Models*. London: Chapman & Hall.
- [11]. Kammann, E. E., Wand MP. 2003. Geoadditive models. *Appl Stat* 52:1-18.
- [12]. Ruppert, D., Carroll, R. J. 1997. Penalized regression splines. Unpublished manuscript. [terhubung berkala]. <http://www.orie.cornell.edu/~davidr/papers/index/index/index.html>
- [13]. Searle, S. R., Casella, G., dan McCulloch, C. E. 1992. *Variance Component*. New York : John Wiley & Sons.

- [14]. Stone, C. J. 1985. Additive Regression and Other Nonparametric Models. *Ann Stat* 13: 689–705.
- [15]. 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). *J R Stat Soc, Series C* 48: 269-312.
- [16]. Wand M. 2003. Smoothing and mixed models. *Comp Stat* 18:223–249.
- [17]. Wang Y. 1998. Mixed effects smoothing spline analysis of variance. *J R Stat Soc, Series B* 60:159-174.