Perbandingan Model Regresi Nonparametrik Spline Multivariabel dengan Menggunakan Metode *Generalized Cross Validation* (GCV) dan Metode *Unbiassed Risk* (UBR) dalam Pemilihan Titik Knot Optimal

Sulistya Umie Rumana Sari
Fakultas Tarbiyah dan Ilmu Keguruan, UIN Maulana Malik Ibrahim Malang

Email : sulistya0706@gmail.com

Info Artikel

Riwayat Artikel:

Diterima: 15 Mei 2017
Direvisi: 1 Juni 2017
Diterbitkan: 31 Juli 2017

Kata Kunci:

Spline
UBR
GCV
Angka Kematian Maternal.

ABSTRAK

Pada regresi nonparametrik spline menentukan titik knot optimal menjadi hal yang sangat penting. Model regresi spline terbaik dihasilkan dari pemilihan titik knot yang paling optimal. Terdapat beberapa metode yang dapat digunakan untuk memilih titik knot optimal, antara lain *Cross Validation* (CV), *Generalized Cross Validation* (GCV), *Unbiassed Risk* (UBR) dan *Generalized Maximum Likelihood* (GML). Penelitian ini akan membahas tentang perbandingan model regresi spline nonparametrik multivariabel dengan menggunakan metode GCV dan UBR sebagai metode pemilihan titik knot optimal. Kriteria pemilihan model terbaik adalah berdasarkan nilai MSE dan R^2_{adj}. Selanjutnya akan dilakukan pemodelan menggunakan data Angka Kematian Maternal di Jawa Timur dengan menggunakan regresi nonparametrik *spline*. Pada hasil penelitian didapatkan menggunakan metode GCV, titik knot optimum adalah menggunakan kombinasi titik knot 2-3-2-2-1 yang menghasilkan nilai MSE 0,002059. Selanjutnya pada uji parameter didapatkan bahwa semua variabel berpengaruh signifikan dan semua asumsi residual terpenuhi dengan nilai R^2_{adj} adalah sebesar 92,4%. Sementara, dengan metode UBR didapatkan titik knot optimum adalah menggunakan satu titik knot yang menghasilkan nilai MSE sebesar 0,01315. Pada uji parameter didapatkan bahwa semua variabel tidak berpengaruh signifikan dan asumsi residual distribusi normal tidak terpenuhi dengan nilai R^2_{adj} adalah sebesar 52,15%. Hal ini membuktikan bahwa pemodelan regresi nonparametrik dengan menggunakan metode GCV lebih baik untuk data angka kematian maternal dibandingkan menggunakan metode UBR dalam pemilihan titik knot.

Korespondensi:

Sulistya Umie Rumana Sari
Fakultas Tarbiyah dan Ilmu Keguruan
UIN Maulana Malik Ibrahim Malang,
Jl. Gajayana No. 50 Malang, Jawa Timur, Indonesia 65144
Email: sulistya0706@gmail.com

1. PENDAHULUAN

Analisis regresi adalah suatu metode statistika yang dilakukan untuk mengetahui pola hubungan antara satu atau lebih variabel. Bentuk pola hubungan antara variabel prediktor dengan variabel respon dapat diidentifikasi berdasarkan informasi masa lalu atau dengan menggunakan *scatter plot* [1]. Selain itu analisis regresi bertujuan untuk melakukan sebuah prediksi [2]. Terdapat 3 pendekatan dalam melakukan estimasi kurva regesi yaitu pendekatan parametrik, nonparametrik dan pendekatan semiparametrik. Dalam pendekatan nonparametrik, kurva regresi hanya diasumsikan *smooth* (mulus) dalam arti termuat di dalam suatu fungsi tertentu. Kelebihan ketika menggunakan pendekatan nonparametrik yaitu mempunyai fleksibilitas yang tinggi, artinya data diharapkan dapat mencari sendiri bentuk estimasinya, tanpa adanya pengaruh dari subjektivitas si peneliti.Terdapat beberapa metode yang telah banyak digunakan untuk memodelkan regresi dengan menggunakan pendekatan nonparametrik yaitu antara lain: Kernel, Spline, *K-Nearest Neighbor*, Histogram, Estimator Deret Fourier, MARS, Deret Orthogonal, Wavelets, Neural Network. Dari berbagai metode yang banyak digunakan, spline adalah salah satu metode yang mempunyai banyak kelebihan. Spline *Truncated* merupakan fungsi dimana terdapat perubahan pola perilaku kurva yang berbeda pada interval-interval yang berlainan. Berbagai macam kelebihan jika menggunakan spline antara lain: mempunyai intepretasi visual yang baik, bersifat fleksibel, serta mampu menangani karakter fungsi yang bersifat mulus [3,4]. Selain itu, spline dapat mengatasi pola data dan dapat menggambarkan perubahan perilaku data yang berubah-ubah pada sub-sub interval tertentu. Spline juga mempunyai keunggulan dalam mengatasi pola data yang menunjukkan naik/turun yang tajam dengan bantuan titik-titik knot serta kurva yang dihasilkan relatif mulus [3]. Spline diperoleh berdasarkan optimasi yang merupakan perluasan dari optimasi optimasi yang digunakan pada regresi parametrik. Spline merupakan potongan-potongan polinomial yang memiliki sifat tersegmen dan kontinu. Salah satu kelebihan spline seperti yang telah dijelaskan adalah bersifat fleksibel, artinya model ini cenderung mencari sendiri estimasi data kemanapun pola data tersebut bergerak. Hal tersebut terjadi karena dalam spline terdapat adanya titik-titik knot. Titik knot adalah titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data [2,3]. Dengan titik knot ini, spline dapat memberikan fleksibilitas yang lebih baik dari pada polinomial, sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik lokal dari suatu fungsi atau data. Terdapat beberapa metode untuk memilih titik knot yang optimal dalam regresi nonparametrik Spline antara lain metode *Cross Validation* (CV), *Unbiassed Risk* (UBR), *Generalized Cross Validation* (GCV), dan *Generalized Maximum Likelihood* (GML). Dalam penelitian ini akan membandingkan hasil pemilihan titik knot optimal dengan metode GCV dan metode UBR.

GCV dan UBR adalah metode pemilihan titik knot optimal yang mempunyai banyak kelebihan. Adapun kelebihan yang dimiliki metode GCV antara lain adalah sederhana dan efisien dalam perhitungan, optimal secara asimtotik, invarian terhadap transformasi dan tidak memerlukan informasi terhadap σ^2. Pemilihan titik knot dengan metode GCV akan lebih baik jika digunakan pada data yang *Gaussian* (berdistribusi normal), sementara pemilihan titik knot dengan metode UBR akan lebih baik jika digunakan pada data *non-Gaussian* atau tidak berdistribusi normal [5,6]. Berdasarkan pernyataan tersebut, peneliti tertarik untuk membandingkan apakah model regresi nonparametrik spline multivariabel menggunakan metode GCV akan lebih baik jika dibandingkan dengan model regresi nonparametrik spline multivariabel menggunakan metode UBR pada data Angka Kematian Maternal di Provinsi Jawa Timur tahun 2013 yang mempunyai distribusi normal.

Kematian Maternal menurut batasan dari The Tenth Revision of The International Classification of Diseases (ICD-10) adalah kematian wanita yang terjadi pada saat kehamilan atau dalam waktu 42 hari setelah kehamilan, tidak tergantung dari lama dan lokasi kehamilan, disebabkan oleh apapun yang berhubungan dengan kehamilan, atau yang diperberat oleh kehamilan tersebut, atau penanganannya, akan tetapi bukan kematiaan yang disebabkan oleh kecelakaan atau kebetulan [7]. Saat ini Indonesia masih menghadapi permasalahan tingginya angka kematian maternal. Menurut Survei Demografi Kesehatan Indonesia, Indonesia memiliki angka kematian maternal tertinggi di ASEAN pada tahun 2002-2003 sebesar 307 per 100000 kelahiran. Angka kematian maternal di Indonesia tahun 2007 adalah 228 per 100000 kelahiran pada tahun 2007 dan turun menjadi 220 per 100000 kelahiran pada tahun 2010. Hal tersebut menyebabkan Indonesia menduduki peringkat 51 tertinggi angka kematian maternal di dunia menurut CIA *World Factbook* pada tahun 2010. Akan tetapi, angka kematian maternal meningkat lagi pada tahun 2011 menjadi 228 per 100000 kelahiran, yang membuat angka kematian di Indonesia tertinggi di Asia Tenggara. [12]

2. METODOLOGI PENELITIAN

Variabel-variabel yang digunakan dalam penelitian ini berdasarkan penelitian sebelumnya yang dilakukan oleh Nuraziza Arfan pada tahun 2013 yang meneliti tentang pemodelan regresi nonparametrik spline pada kasus angka kematian maternal di Jawa Timur [13]. Adapun variabel yang digunakan dalam penelitian ini antara lain sebgai berikut

Tabel 1. Variabel Penelitian

Variabel	Nama Variabel
Y	Angka kematian maternal
x_1	Persentase ibu hamil yang mendapatkan tablet Fe1
x_2	Persentase ibu hamil melaksanakan program K1
x_3	Presentase ibu hamil berisiko tinggi/komplikasi yang ditangani
x_4	Persentase penduduk perempuan yang pernah kawin di bawah umur

x_5	Persentase penduduk perempuan dengan pendidikan paling tinggi SD

Metode yang digunakan dalam penelitian ini adalah metode regresi dengan pendekatan nonprametrik spline multivariabel dan menggunakan gcv dan ubr sebagai metode pemilian titik knot optimal. Pendekatan regresi nonparametrik spline digunakan jika kurva regresi antara variabel respon dengan variabel prediktor tidak membentuk suatu pola atau tidak ada informasi masa lalu yang lengkap mengenai pola data. Dalam banyak hal, pengamatan-pengamatan yang akan dikaji tidak selalu memenuhi asumsi-asumsi yang mendasari uji parametrik sehingga kerap kali dibutuhkan teknik-teknik inferensial dengan validitas yang tidak bergantung pada asumsi-asumsi yang kaku. Dalam hal ini, teknik-teknik dalam regresi nonparametrik memenuhi kebutuhan. Regresi nonparametrik memiliki fleksibilitas yang tinggi, karena data diharapkan mencari sendiri bentuk estimasi kurva regresinya tanpa dipengaruhi oleh faktor subyektifitas peneliti [3].

Secara umum, model regresi nonparametrik dapat disajikan sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i \quad , i = 1,2,\cdots,n \tag{1}$$

Spline merupakan model polynomial yang tersegmen. Polinomial tersegmen memegang peranan penting dalam teori dan aplikasi statistika. Regresi spline memiliki titik knot yang merupakan titik perpaduan yang menunjukkan perubahan perilaku kurva pada selang yang berbeda [10]. Secara umum fungsi spline $f(x_i)$ berorde m dengan titik knot $k_1, k_2, \dots, k_J$ dapat dinyatakan sebagai berikut.

$$f(x_i) = \sum_{j=0}^{m} \beta_j x_i^j + \sum_{j=1}^{J} \beta_{m+j} \left(x_i - k_j\right)_+^m \tag{2}$$

dengan β_j merupakan parameter-parameter model dan m merupakan orde spline [8]. Persamaan diatas bila disubstitusikan pada persamaan diperoleh persamaan regresi nonparametrik spline sebagai berikut.

$$y_i = \sum_{j=0}^{m} \beta_j x_i^j + \sum_{j=1}^{J} \beta_{m+j} \left(x_i - k_j\right)_+^m + \varepsilon_i \tag{3}$$

dimana : $i = 1,2,...,n$

dengan fungsi truncated diberikan oleh :

$$\left(x_i - k_j\right)_+^m = \begin{cases} \left(x_i - k_j\right)^m & , x_i \geq k_j \\ 0 & , x_i < k_j \end{cases} \tag{4}$$

Metode GCV menjadi metode yang sangat sering digunakan untuk memilih titik knot yang paling optimal. Spline terbaik ditandai dengan titik knot optimal yang diperoleh. Salah satu metode pemilihan titik knot optimal adalah GCV. Model spline yang terbaik dengan titik knot optimal didapat dari nilai GCV yang terkecil [9]. Seperti yang telah dipaparkan sebelumnya, metode GCV adalah bentuk modifikasi dari metode CV. Dimana persamaan CV adalah sebagai berikut.

$$\mathrm{CV}(\tilde{k}) = n^{-1} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(\frac{\mathrm{y_i} - \hat{\mu}_{\mathrm{k(J)}}}{1 - \mathbf{A}(k_1, k_2, ..., k_{\mathrm{J}})} \right)^2 \tag{5}$$

Sementara, persamaan GCV diberikan oleh persamaan 6.

$$\mathrm{GCV}(\tilde{k}) = n^{-1} \sum_{\mathrm{i}=1}^{\mathrm{n}} \left(y_i - \hat{\mu}_{\mathrm{k(J)}} \right)^2 \mathbf{Q} \tag{6}$$

dimana :

$$\mathbf{Q} = \left\{ \frac{1 - \mathbf{A}(k_1, k_2, ..., k_J)}{n^{-1}\,\mathrm{trace}\left[\mathbf{I} - \mathbf{A}(k_1, k_2, ..., k_J)\right]} \right\}^2$$

Persamaan GCV juga dapat ditulis dengan persamaan sebagai berikut.

$$GCV(\tilde{k}) = n^{-1} \frac{\sum_{i=1}^{n} \left(y_i - \hat{f}(x) \right)^2}{\left[1 - n^{-1}\,\mathrm{trace}\left[\mathbf{A}(\tilde{k})\right]\right]^2} \tag{7}$$

dimana :

$$\hat{f}(x) = \mathbf{A}(\tilde{k})\tilde{Y} = \left[\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\right]\tilde{Y}$$

$$\mathbf{A}(\tilde{k}) = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$$

$$MSE = n^{-1}\sum_{i=1}^{n}(y_i - \hat{f}(x))^2$$

I : matriks identitas

n : banyak pengamatan

Sementara, metode UBR adalah salah satu metode yang juga digunakan untuk memilih titik knot optimal dalam pemilihan model spline terbaik. Adapun persamaan yang digunakan untuk mencari nilai UBR adalah sebagai berikut. [11]

$$U(\tilde{k}) = \frac{1}{n}\left[\left\|\left(\mathbf{I} - \mathbf{A}(\tilde{k})\right)\tilde{Y}\right\|^2 + \frac{\sigma^2}{n} trace\left[\mathbf{I} - \mathbf{A}(\tilde{k})\right]^2 + \frac{\sigma^2}{n} trace\left[\mathbf{A}^2(\tilde{k})\right]\right] \quad (8)$$

dimana :

$$\hat{\sigma}^2 = \frac{\left\|\left[\mathbf{I} - \mathbf{A}(\tilde{k})\right]\tilde{Y}\right\|^2}{\text{trace}\left[\mathbf{I} - \mathbf{A}(\tilde{k})\right]}$$

$$\mathbf{A}(\tilde{k}) = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$$

I : matriks identitas

n : banyak pengamatan

Adapun langkah- langkah yang dilakukan dalam penelitian ini adalah sebagai berikut :

a. Membuat *scatter plot* antara variabel respon (y) dengan variabel prediktor (x).
b. Memodelkan data dengan model regresi nonparametrik spline dengan satu ttik knot, dua titik knot, dan tiga titik knot.
c. Menghitung niai GCV dan UBR untuk masing masing model regresi nonparametrik spline.
d. Menentukan titik knot optimal berdasarkan nilai GCV dan UBR.
e. Melakukan pengujian signifikansi parameter yang dihasilkan dari estimasi model regresi nonparametrik spline dengan metode GCV dan metode UBR
f. Melakukan uji asumsi residual yang dihasilkan dari estimasi model regresi nonparametrik spline dengan metode GCV dan metode UBR.
g. Membandingkan kebaikan model regresi nonparametrik spline dengan titik knot optimal menggunakan metode GCV dan UBR berdasarkan kriteria nilai R^2_{adj} dan nilai MSE.

3. HASIL DAN PEMBAHASAN

3.1. Metode GCV dalam Regresi Nonparametrik Spline

Metode GCV adalah metode yang sangat sering digunakan untuk memilih titik knot yang paling optimal. Asumsikan bahwa $\text{trace}\left[\mathbf{A}(\tilde{k})\right] < n$, maka kriteria GCV didefinisikan sebagai berikut:

$$GCV(\tilde{k}) = \frac{n^{-1}\text{MSE}(\tilde{k})}{\left(n^{-1}\text{trace}\left[\mathbf{I} - \mathbf{A}(\tilde{k})\right]\right)^2}$$

dimana :

$$\mathbf{A}(\tilde{k}) = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$$

$$MSE(\tilde{k}) = n^{-1}\sum_{i=1}^{n}(y_i - \hat{f}(x))^2$$

$$\hat{f}(x) = \mathbf{A}(\tilde{k})\tilde{Y} = \left[\mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\right]\tilde{Y}$$

I : matriks identitas

n : banyak pengamatan

Persamaan GCV diatas dapat ditulis juga dengan persamaan sebagai berikut :

$$\mathrm{GCV}(\tilde{k}) = n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{\mu}_{k(J)}\right)^2 \left\{\frac{1-\mathbf{A}(k_1,k_2,\dots,k_J)}{n^{-1}\,\mathrm{trace}\left[\mathbf{I}-\mathbf{A}(k_1,k_2,\dots,k_J)\right]}\right\}^2$$

Berdasarkan persamaan yang ditulis diatas dapat dilihat bahwa GCV merupakan hasil generalisasi metode CV yang telah diberi bobot. Dimana persamaan CV adalah sebagai berikut.

$$\mathrm{CV}(\tilde{k}) = n^{-1}\sum_{i=1}^{n}\left(\frac{y_i - \hat{\mu}_{k(J)}}{1-\mathbf{A}(k_1,k_2,\dots,k_J)}\right)^2$$

Untuk memperoleh titik knot optimal dengan menggunakan metode GCV, maka dilakukan optimasi sebagai berikut.

$$\underset{K_k\in R}{\mathrm{Min}}\left\{\mathrm{GCV}(k_1,k_2,\dots,k_J)\right\} = \underset{K_k\in R}{\mathrm{Min}}\left\{\frac{MSE(k_1,k_2,\dots,k_J)}{\left(n^{-1}\left(\mathrm{trace}\left[\mathbf{I}-\mathbf{A}(k_1,k_2,\dots,k_J)\right]\right)\right)^2}\right\}$$

$$= \underset{K_k\in R}{\mathrm{Min}}\left\{\frac{n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{f}(x)\right)^2}{\left(n^{-1}\left(\mathrm{trace}\left[\mathbf{I}-\mathbf{A}(k_1,k_2,\dots,k_J)\right]\right)\right)^2}\right\} = \underset{K_k\in R}{\mathrm{Min}}\left\{\frac{n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{f}(x)\right)^2}{\left\{1-n^{-1}\mathrm{trace}\left[\mathbf{A}(k_1,k_2,\dots,k_J)\right]\right\}^2}\right\}$$

$$= \underset{K_k\in R}{\mathrm{Min}}\left\{\frac{n^{-1}\sum_{i=1}^{n}\left(y_i - \hat{f}(x)\right)^2}{\left\{1-n^{-1}\mathrm{trace}\ \mathbf{A}(\underset{\sim}{k})\right\}^2}\right\} = \underset{\mathbf{K_k}\in\mathbf{R}}{\mathrm{Min}}\left\{\frac{n^{-1}\left(\tilde{Y}-\mathbf{A}[\tilde{k}]\tilde{Y}\right)^{\mathrm{T}}\left(\tilde{Y}-\mathbf{A}[\tilde{k}]\tilde{Y}\right)}{\left\{1-n^{-1}\ \mathrm{trace}\ \mathbf{A}[\tilde{k}]\right\}^2}\right\}$$

$$= \underset{K_k\in R}{\mathrm{Min}}\left\{\frac{n^{-1}\tilde{Y}^{\mathrm{T}}\left(\mathbf{I}-\mathbf{A}[\tilde{k}]\right)^{\mathrm{T}}\left(\mathbf{I}-\mathbf{A}[\tilde{k}]\right)\tilde{Y}}{\left\{1-n^{-1}\mathrm{trace}\ \mathbf{A}[\tilde{k}]\right\}^2}\right\}$$

Nilai GCV terkecil yang dihasilkan akan memberikan titik knot yang optimal.

3.2. Metode UBR dalam Regresi Nonparametrik Spline

Metode UBR adalah digunakan untuk mencari titik knot optimal ketika ada informasi mengenai σ^2 atau σ^2 diketahui. Penggunaan metode UBR dalam pemilihan titik knot yang optimal sangat bergantung pada nilai estimasi σ^2. Artinya, ketika estimasi σ^2 yang dihasilkan cukup baik, maka metode UBR akan berkerja dengan baik pula. Berikut rumus yang digunakan untuk mencari titik knot dengan metode UBR.

$$U(\tilde{k}) = \frac{1}{n}\left[\left\|\left(\mathbf{I}-\mathbf{A}(\tilde{k})\right)\tilde{Y}\right\|^2 + \frac{\sigma^2}{n}trace\left[\mathbf{I}-\mathbf{A}(\tilde{k})\right]^2 + \frac{\sigma^2}{n}trace\left[\mathbf{A}^2(\tilde{k})\right]\right]$$

dimana :

$$\mathbf{A}(\tilde{k}) = \mathbf{X}(\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}$$

$\mathbf{I}$: matriks identitas

n : banyak pengamatan

Estimasi dari σ^2 dapat diperoleh dengan rumus sebagai berikut.

$$\hat{\sigma}^2 = \frac{\left\|\left[\mathbf{I}-\mathbf{A}(\tilde{k})\right]\tilde{Y}\right\|^2}{\mathrm{trace}\left[\mathbf{I}-\mathbf{A}(\tilde{k})\right]}$$

Selanjutnya titik knot optimal dengan metode UBR didapatkan dengan mencari nilai optimasi sebagai berikut :

$$\underset{K_k \epsilon R}{Min}\{U(k_1,k_2,k_3,\dots,k_J)\} = \underset{K_k \epsilon R}{Min}\, {}^1/_n \left[\left\|\left(\mathbf{I}-\mathbf{A}(\tilde{k})\right)\tilde{Y}\right\|^2 + {}^{\sigma^2}/_n\, trace\left[\mathbf{I}-\mathbf{A}(\tilde{k})\right]^2 + {}^{\sigma^2}/_n\, trace\mathbf{A}^2(\tilde{k})\right]$$

$$= \underset{K_k \epsilon R}{Min}\, {}^1/_n \left[\left\|\left(\mathbf{I}-\mathbf{A}(\tilde{k})\right)\tilde{Y}\right\|^2 + \frac{\left\|\left(\mathbf{I}-\mathbf{A}(\tilde{k})\right)\tilde{Y}\right\|^2}{n} trace\left[\mathbf{I}-\mathbf{A}(\tilde{k})\right] + \frac{trace\mathbf{A}^2(\tilde{k})}{n} \frac{\left\|\left(\mathbf{I}-\mathbf{A}(\tilde{k})\right)\tilde{Y}\right\|^2}{trace\left[\mathbf{I}-\mathbf{A}(\tilde{k})\right]}\right]$$

Titik knot optimal dihasilkan dengan nilai UBR yang terkecil.

3.3. Statistika Deskriptif

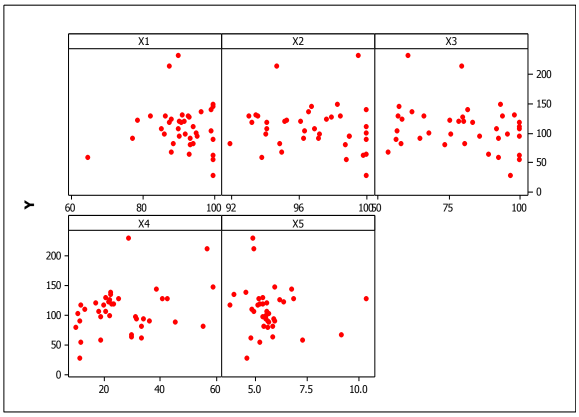
Berikut adalah statistika deskriptif dari angka kematian maternal dan variabel yang diduga berpengaruh.

Tabel 2. Statistika deskriptif angka kematian maternal dan variabel yang diduga berpengaruh

Variabel	Rata-rata	Varians	Jumlah	Minimum	Maximum
Y	107,51	1502,23	4085,27	26,06	230,64
X_1	91,34	52,83	3470,92	64,8	100
X_2	96,75	6,124	3676,5	91,899	100
X_3	79,99	251,52	3039,81	53,51	100
X_4	27	169,11	1026,15	10,07	59,09
X_5	5,682	1,455	2159,1	3,78	10,36

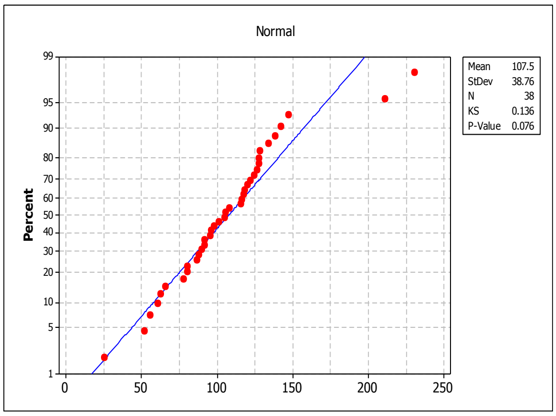
Berdasarkan Tabel 2 yang menjelaskan statistika deskriptif, terlihat bahwa variabel angka kematian maternal (Y) mempunyai nilai terendah sebesar 26,06 persen yaitu terdapat di kabupaten Madiun, sebaliknya nilai tertinggi angka kematian maternal sebesar 230,64 persen terdapat di kabupaten Probolinggo. Jumlah kabupaten keseluruhan sebanyak 38 di Provinsi Jawa Timur mempunyai rata-rata angka kematian maternal sebesar 107,51 dengan varians 1502,23. Begitu juga untuk variabel prediktor, variabel persentase ibu hamil yang mendapatkan tablet Fe1 (X_1) mempunyai nilai terendah sebesar 64,8 persen yaitu terdapat di kabupaten Pasuruan, dan nilai tertinggi sebesar 100 yaitu terdapat di kabupaten Lumajang, Jember, Bondowoso, Sumenep, Mojokerto dan Madiun. Variabel persentase ibu hamil yang mendapatkan tablet Fe_1 mempunyai rata-rata sebesar 91,34 persen dengan varians sebesar 52,83.

Berikut juga diberikan *scatter plot* dari data angka kematian maternal di provinsi Jawa Timur dengan faktor faktor yang diduga mempengaruhinya sebagai pendugaan pola data.



Gambar 1. *Scatter plot* data angka kematian maternal dan faktor-faktor yang diduga berpengaruh

Berdasarkan Gambar 1 yang menampilkan scatter plot antar variabel angka kematian maternal dengan masing masing variabel yang diduga berpengaruh dapat dilihat bahwa data pada scatter plot menyebar, tidak menunjukkan sebuah pola tertentu dan berubah-ubah pada sub-sub interval tertentu. Oleh karena itu, dalam penelitian ini, akan digunakan regresi nonparametrik. Berikut adalah hasil uji kenormalan data Angka Kematian Maternal.



Gambar 2. Uji Normalitas Angka Kematian Maternal

Berdasarkan Gambar 2 dapat dilihat bahwa Angka Kematian Maternal berdistribusi normal dengan nilai p-value yang lebih dari (0,05), yaitu 0,076.

3.4. Pemodelan Regresi dengan Menggunakan Metode *GCV* sebagai Pemilihan Titik Knot Optimal

Sebelum melakukan pemodelan, maka langkah awal yang seharusnya dilakukan adalah melakukan pemilihan titik knot optimal. Berikut adalah hasil pemilihan tiitk knot optimal dengan metode GCV dengan menggunakan, satu tiitk knot, dua titik knot, tiga titik knot dan kombinasi titik knot.

Tabel 3. Ringkasan hasil pemilihan titik knot dengan metode GCV

Titik Knot	Nilai GCV
Satu titik knot	0,02775
Dua titik knot	0,03688
Tiga titik knot	0,02022
Kombinasi titik knot (2-3-2-2-1)	0,008796

Dari tabel terlihat bahwa nilai GCV minimum kombinasi titik knot sebesar 0,008796 dengan titik knot kombinasi adalah 2-3-2-2-1. Dari hasil pemilihan titik knot optimal dengan menggunakan metode GCV diperoleh nilai GCV paling minimum yaitu sebesar 0,008796 yaitu menggunakan titik knot kombinasi 2-3-2-2-1. Sehingga untuk analisis selanjutnya menggunakan model regresi nonparametrik spline sebagai berikut.

$$\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 (x_1 - 13{,}559)_+ + \tilde{\beta}_3 (x_1 - 13{,}94)_+ + \tilde{\beta}_4 x_2 + \tilde{\beta}_5 (x_2 - 3{,}997)_+ + \tilde{\beta}_6 (x_2 - 4{,}77)_+ +$$

$$\tilde{\beta}_7 (x_2 - 4{,}963)_+ + \tilde{\beta}_8 x_3 + \tilde{\beta}_9 (x_3 - 5{,}462)_+ + \tilde{\beta}_{10} (x_3 - 5{,}714)_+ + \tilde{\beta}_{11} x_4 + \tilde{\beta}_{12} (x_4 - 5{,}943)_+ +$$

$$\tilde{\beta}_{12} (x_4 - 6{,}257)_+ + \tilde{\beta}_{14} x_5 + \tilde{\beta}_{15} (x_5 - 32{,}19)_+$$

Berikut ditampilkan hasil estimasi parameter model regresi nonparametrik spline dengan menggunakan kombinasi titik knot GCV (2-3-2-2-1).

Tabel 4 Estimasi parameter model regresi

Parameter	Estimator
β_0	$\hat{\beta}_0 =$ 0,479159
β_1	$\hat{\beta}_1 =$ 0,004055
β_2	$\hat{\beta}_2 =$ -0,005093
β_3	$\hat{\beta}_3 =$ 0,025522
β_4	$\hat{\beta}_4 =$ 0,005456
β_5	$\hat{\beta}_5 =$ -0,016544
β_6	$\hat{\beta}_6 =$ 0,393170
β_7	$\hat{\beta}_7 =$ -0,576185
β_8	$\hat{\beta}_8 =$ 0,960686
β_9	$\hat{\beta}_9 =$ -4,889166

β_{10}	$\hat{\beta}_{10}$ = 3,597526
β_{11}	$\hat{\beta}_{11}$ = -1,082000
β_{12}	$\hat{\beta}_{12}$ = 1,383610
β_{13}	$\hat{\beta}_{13}$ = -0,188521
β_{14}	$\hat{\beta}_{14}$ = 0,109266
β_{15}	$\hat{\beta}_{15}$ = 0,01929

Berdasarkan model yang diperoleh dengan menggunakan metode GCV dalam pemilihan titik knot optimal yaitu dengan menggunakan kombinasi titik knot (2-3-2-2-1) menghasilkan nilai R^2_{adj} sebesar 92,4%. Artinya, variabel prediktor mempengaruhi angka kematian maternal sebesar 84,8%.

3.4.1 Pengujian Asumsi Residual Identik

Hipotesis yang digunakan adalah sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = ... = \sigma_n^2 = \sigma^2$

H_1 : minimal ada satu $\sigma_i^2 \neq \sigma^2$, i=1,2,...,n

Pengujian identik dilakukan dengan menggunakan uji *glejser*, dengan output sebagai berikut.

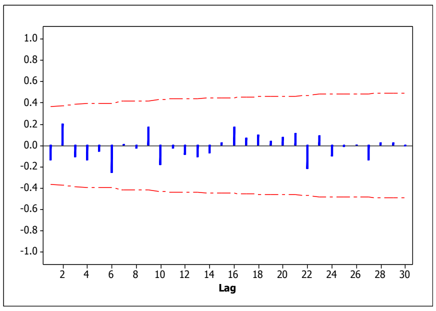
Tabel 5. ANOVA Uji Glejser

Sumber	df	SS	MS	Fhit	P value
Regresi	15	0,004654	0,0003103	0,4982	0,9056
Error	15	0,009341	0,0006227		
Total	30	0,01399			

Berdasarkan ANOVA pada tabel diatas, diperoleh nilai Fhitung sebesar 0,4982. Jika dibandingkan dengan nilai $F_{(15,15,0,05)}$ yaitu sebesar 2,58 maka diputuskan H_0 gagal ditolak. Dipertegas pula dengan nilai p-value sebesar 0,9056 yang lebih besar dari α(0,05). Sehingga dapat disimpulkan bahwa keputusannya H_0 gagal ditolak. Jadi dapat diartikan bahwa tidak terjadi heterokedastisitas. Hal ini menunjukkan bahwa residual telah memenuhi asumsi identik.

3.4.2 Pengujian Asumsi Resiudal Independen

Asumsi residual independen terpenuhi apabila tidak terdapat autokorelasi lag yang keluar dari batas signifikansi. Plot ACF adalah salah satu cara visual yang digunakan untuk mendeteksi adanya autokorelasi antar residual. Asumsi residual independen terpenuhi jika pada plot ACF tidak ada autokorelasi yang keluar dari batas signifikansi. Berikut disajikan plot ACF dari residual.



Gambar 3 *ACF plot* residual pemodelan regresi dengan metode GCV

Berdasarkan visualisasi plot ACF yang dihasilkan seperti pada Gambar 3 terlihat bahwa autokorelasi pada semua lag berada di dalam batas signifikansi atau bisa dikatakan bahwa tidak ada autokorelasi yang keluar dari batas signifikansi. Sehingga dapat disimpukan bahwa H_0 gagal ditolak, maka residual telah memenuhi asumsi independen.

3.4.3 Pengujian Asumsi Residual Distribusi Normal

Hipotesis yang digunakan untuk melakukan uji asumsi distribusi normal adalah sebagai berikut.

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Tabel 6. Uji Normalitas Shapiro-Wilk dan Anderson-Darling

Metode	P value
Shapiro-Wilk	0,7747
Anderson-Darling	0,3657

Berdasarkan uji normalitas menggunakan Shapiro-Wilk dan Anderson-Darling didapat nilai P-value yang lebih dari (0,05). Sehingga dapat disimpulkan bahwa asumsi residual berdistribusi normal telah terpenuhi.

3.4.5 Pengujian Parameter Model Secara Serentak

Berdasarkan hasil pemilihan titik knot optimal, maka selanjutnya dilakukan estimasi dan pengujian parameter secara serentak. Adapun hipotesis yang digunakan adalah sebagai berikut.

$H_0: \beta_1 = \beta_2 = \dots = \beta_{15} = 0$

$H_1:$ minimal terdapat satu $\beta_j \neq 0, \quad j = 1, 2, \dots, 15$

Hasil ANOVA untuk model regresi nonparametrik spline secara serentak disajikan pada tabel berikut.

Tabel 7. Hasil estimasi parameter model regresi

Sumber	Df	SS	MS	Fhit	P value
Regresi	15	0,3755	0,02503	12,15	$8{,}561 \times 10^{-6}$
Error	15	0,03089	0,002059		
Total	30	0,4064			

Tabel 7 adalah hasil estimasi parameter model regresi yang didapatkan dengan menggunakan metode GCV dalam pemilihan titik knot optimal. seperti yang disajikan pada tabel diatas dapat diketahui bahwa nilai F_{hitung} sebesar 12,15. Jika dibandingkan dengan nilai $F_{(15,15,0,05)}$ yaitu sebesar 2,58 maka diputuskan untuk H_0 ditolak. Dipertegas pula dengan p-value sebesar 0,000008561, nilai ini kurang dari α yaitu 0,05. Sehingga dapat disimpulkan bahwa minimal terdapat satu parameter yang signifikan terhadap angka kematian maternal. Selanjutnya, untuk mengetahui parameter mana yang memberikan pengaruh signifikan maka perlu dilakukan pengujian parameter secara parsial.

3.4.6 Pengujian Parameter Model Secara Parsial

Untuk melakukan pengujian signifikansi parameter secara parsial, maka menggunakan hipotesis sebagai berikut.

$H_0 = \beta_j = 0$

$H_0 = \beta_j \neq 0; \quad$ j=1,2,...,15

Berikut adalah hasil pengujian signifikansi parameter secara parsial.

Tabel 8. Pengujian parameter model regresi secara parsial

Variabel	Parameter	T hitung	P value	Keputusan
	β_0	2,8747	$1{,}157 x 10^{-2}$	Signifikan
X_1	β_1	0,6245	$5{,}417\ x 10^{-1}$	Tidak Signifikan
	β_2	-0,2708	$7{,}903\ x 10^{-1}$	Tidak Signifikan
	β_3	2,0887	$5{,}419\ x 10^{-2}$	Signifikan
X_2	β_4	0,6379	$5{,}332\ x 10^{-1}$	Tidak Signifikan
	β_5	-3,6127	$2{,}558\ x 10^{-3}$	Signifikan
	β_6	2,5809	$2{,}088\ x 10^{-2}$	Signifikan
	β_7	-2,8523	$1{,}211\ x 10^{-2}$	Signifikan
X_3	β_8	8,7660	$2{,}744\ x 10^{-7}$	Signifikan
	β_9	-6,6604	$7{,}605\ x 10^{-6}$	Signifikan
	β_{10}	4,0007	$1{,}158\ x 10^{-3}$	Signifikan
X_4	β_{11}	-2,9247	$1{,}046\ x 10^{-2}$	Signifikan
	β_{12}	2,6388	$1{,}860\ x 10^{-2}$	Signifikan
	β_{13}	-1,1987	$2{,}492\ x 10^{-1}$	Tidak Signifikan

X_5	β_{14}	0,5653	5,802 $x10^{-1}$	Tidak Signifikan
	β_{15}	3,7115	2,089 $x10^{-3}$	Signifikan

Berdasarkan tabel diatas terdapat beberapa parameter yang tidak signifikan terhadap angka kematian maternal. Dari 15 parameter yang ada pada model regresi nonparametrik spline, terdapat 5 parameter yang tidak signifikan pada taraf signifikansi 0,05, karena p-value lebih dari α(0,05). Meskipun 5 parameter yang tidak signifikan, namun semua variabel berpengaruh terhadap angka kematian maternal.

3.5 Pemodelan Regresi dengan Menggunakan Metode UBR sebagai Pemilihan Titik Knot Optimal

Sebelum melakukan pemodelan, maka langkah awal yang seharusnya dilakukan adalah melakukan pemilihan titik knot optimal. Berikut adalah hasil pemilihan tiitk knot optimal dengan metode UBR dengan menggunakan, satu tiitk knot, dua titik knot, tiga titik knot dan kombinasi titik knot.

Tabel 9. Ringkasan hasil pemilihan titik knot dengan metode UBR

Titik Knot	Nilai UBR
Satu titik knot	1,00 x10^{-15}
Dua titik knot	2,19 x10^{-15}
Tiga titik knot	3,65 x10^{-15}
Kombinasi titik knot (2-3-2-2-1)	1,85 x10^{-15}

Dari tabel terlihat bahwa nilai UBR minimum adalah sebesar 1,00 x10^{-15} dengan satu titik knot. Sehingga untuk analisis selanjutnya menggunakan satu titik knot dengan model regresi nonparametrik spline sebagai berikut.

$$\tilde{y} = \tilde{\beta}_0 + \tilde{\beta}_1 x_1 + \tilde{\beta}_2 (x_1 - 13{,}94)_+ + \tilde{\beta}_3 x_2 + \tilde{\beta}_4 (x_2 - 4{,}77)_+ + \tilde{\beta}_5 x_3 + \tilde{\beta}_6 (x_3 - 5{,}714)_+ + \tilde{\beta}_7 x_4 + \tilde{\beta}_8 (x_4 - 6{,}257)_+ + \tilde{\beta}_9 x_5 + \tilde{\beta}_{10} (x_5 - 70{,}66)_+$$

Berikut ditampilkan hasil estimasi parameter model regresi nonparametrik spline dengan menggunakan titik knot UBR paling minimum yaitu dengan menggunakan satu titik knot.

Tabel 10. Estimasi Parameter Model

Parameter	Estimator
β_0	$\hat{\beta}_0$ = -0,032371
β_1	$\hat{\beta}_1$ = -0,003134
β_2	$\hat{\beta}_2$ = 0,074941
β_3	$\hat{\beta}_3$ = 0,00338
β_4	$\hat{\beta}_4$ = -0,005877
β_5	$\hat{\beta}_5$ = 0,002066
β_6	$\hat{\beta}_6$ = 0,024672
β_7	$\hat{\beta}_7$ = -0,490557
β_8	$\hat{\beta}_8$ = -0,003914
β_9	$\hat{\beta}_9$ = -0,023601
β_{10}	$\hat{\beta}_{10}$ = -0,02353

Berdasarkan model regresi yang diperoleh dengan menggunakan metode UBR dalam pemilihan titik knot optimal yaitu dengan menggunakan satu titik knot menghasilkan nilai R^2_{adj} sebesar 52,15%. Artinya, variabel prediktor mempengaruhi angka kematian maternal sebesar 52,15%.

3.5.1 Pengujian Asumsi Residual Identik

Hipotesis yang digunakan untuk pengujian residual identik adalah sebagai berikut.

$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2 = \sigma^2$

H_1 : minimal ada satu $\sigma_i^2 = \sigma^2$, i=1,2,...,n

Pengujian identik dilakukan untuk mengetahui apakah residual memiliki varians yang sama (homogen) atau tidak. Pengujian identik dilakukan dengan menggunakan uji *glejser*, dengan output sebagai berikut.

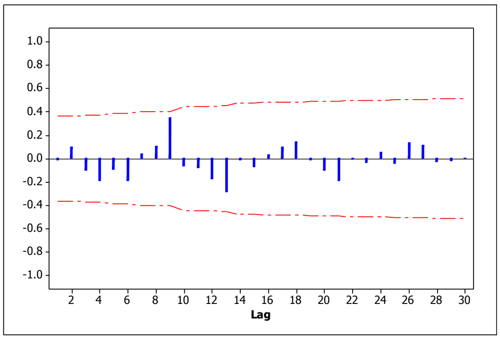
Tabel 11. ANOVA Uji Glejser

Sumber	Df	SS	MS	Fhit	P Value
Regresi	10	0,003588	0,0003588	0,6897	0,7225
Error	20	0,01041	0,0005203		
Total	30	0,01399			

Berdasarkan ANOVA pada tabel diatas, diperoleh nilai Fhitung sebesar 0,6897. Jika dibandingkan dengan nilai $F_{(15,15,0,05)}$ yaitu sebesar 2,58 maka diputuskan H_0 gagal ditolak. Dipertegas pula dengan nilai p-value sebesar 0,7225 yang lebih besar dari α(0,05). Sehingga dapat disimpulkan bahwa keputusannya H_0 gagal ditolak. Jadi dapat diartikan bahwa tidak terjadi heterokedastisitas. Hal ini menunjukkan bahwa residual telah memenuhi asumsi identik.

3.5.2 Pengujian Asumsi Residual Independen

Asumsi residual independen terpenuhi apabila tidak terdapat autokorelasi lag yang keluar dari batas signifikansi. Plot ACF adalah salah satu cara visual yang digunakan untuk mendeteksi adanya autokorelasi antar residual. Asumsi residual independen terpenuhi jika pada plot ACF tidak ada autokorelasi yang keluar dari batas signifikansi. Berikut disajikan plot ACF dari residual.



Gambar 4. *ACF plot* residual pemodelan regresi dengan metode UBR

Berdasarkan visualisasi plot ACF yang dihasilkan seperti pada gambar 4 terlihat bahwa autokorelasi pada semua lag berada di dalam batas signifikansi atau bisa dikatakan bahwa tidak ada autokorelasi yang keluar dari batas signifikansi. Sehingga dapat disimpukan bahwa H_0 gagal ditolak, maka residual telah memenuhi asumsi independen.

3.5.3 Pengujian Asumsi Residual Distribusi Normal

Hipotesis yang digunakan untuk pengujian asumsi distribusi normal adalah sebagai berikut.

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Tabel 12. Uji Normalitas Shapiro-Wilk dan Anderson-Darling

Metode	P value
Shapiro-Wilk	0,01674
Anderson-Darling	0,01257

Berdasarkan uji normalitas menggunakan Shapiro-Wilk dan Anderson-Darling didapat nilai P-value yang kurang dari (0,05). Sehingga dapat disimpulkan bahwa asumsi residual berdistribusi normal tidak terpenuhi.

Berdasarkan uji asumsi residual pada model regresi nonparametrik spline dengan menggunakan metode UBR sebagai metode pemilihan titik knot optimum didapatkan residual tidak memenuhi salah satu asumsi, yaitu residual tidak berdistribusi normal sehingga tidak dapat dilanjutkan pada uji signifikansi parameter.

3.6 Perbandingan Model Regresi dengan Menggunakan Metode GCV dan UBR

Berikut ditampilkan tabel perbandingan metode GCV dan UBR dalam pemodelan regresi nonparametrik spline pada data angka kematian maternal provinsi Jawa Timur.

Tabel 13. Perbandingan Metode GCV dan UBR

	GCV	UBR
MSE	0,002059	0,01315

R_{adj}	92,4%	52,15%
Asumsi Residual Identik	Terpenuhi	Terpenuhi
Asumsi Residual Independen	Terpenuhi	Terpenuhi
Asumsi Residual Distribusi Normal	Terpenuhi	Tidak terpenuhi

Berdasarkan tabel 13 dapat dilihat bahwa nilai dengan menggunakan metode GCV dalam pemilihan titik knot, didapat titik knot optimal dengan menggunakan kombinasi titik knot. Sementara, dengan menggunakan metode UBR titik knot optimal didapat dengan menggunakan satu titik knot. Berdasarkan nilai MSE yang didapatkan dalam pemodelan dapat disimpulkan bahwa menggunakan metode GCV dalam pemilihan titik knot optimal pada data angka kematian maternal adalah lebih baik dibandingkan dengan menggunakan metode UBR. Dan untuk pengujian asumsi residual dengan menggunakan metode UBR diperoleh uji asumsi distribus normal yang tidak terpenuhi sementara dengan metode GCV semua asumsi residual terpenuhi. Hal ini dapat menjadi pembuktian bahwa metode GCV adalah lebih baik digunakan pada data *Gaussian* (berdistribusi normal).

4. KESIMPULAN

Berdasarkan analisis dan pembahasan yang telah dilakukan pada bab 4, dapat ditarik kesimpulan sebagai berikut :

1. Pemilihan titik knot optimum menggunakan metode GCV yaitu dengan optimasi sebagai berikut :

$$\underset{K_k \in R}{\text{Min}}\left\{\text{GCV}(k_1, k_2, \ldots, k_J)\right\} = \underset{K_k \in R}{\text{Min}}\left\{\frac{MSE\left(k_1, k_2, \ldots, k_J\right)}{\left(n^{-1}\left(\text{trace}\left[\mathbf{I} - \mathbf{A}\left(k_1, k_2, \ldots, k_J\right)\right]\right)\right)^2}\right\}$$

2. Pemilihan titik knot optimum menggunakan metode UBR yaitu diperoleh dengan optimasi sebagai berikut :

$$\underset{K_k \in R}{\text{Min}}\left\{\mathbf{U}(k_1, k_2, \ldots, k_J)\right\} = \underset{K_k \in R}{\text{Min}}\frac{1}{n}\left[\left\|\left(\mathbf{I} - \mathbf{A}\left(\underset{\sim}{k}\right)\right)\tilde{Y}\right\|^2 + \frac{\sigma^2}{n}\text{trace}\left[\mathbf{I} - \mathbf{A}(\underset{\sim}{k})\right]^2 + \frac{\sigma^2}{n}\text{trace}\,\mathbf{A}^2(\underset{\sim}{k})\right]$$

3. Pada aplikasi data angka kematian maternal, model regresi nonparametrik yang dihasilkan dengan metode GCV yang paling minimum adalah sebagai berikut.

$$\tilde{y} = 0{,}479 + 0{,}004x_1 - 0{,}005(x_1 - 13{,}559)_+ + 0{,}025(x_1 - 13{,}94)_+ + 0{,}005x_2 - 0{,}02(x_2 - 3{,}997)_+ + 0{,}393(x_2 - 4{,}77)_+ - \tilde{\beta}_7(x_2 - 4{,}963)_+ + 0{,}960x_3 - 4{,}889(x_3 - 5{,}462)_+ + 3{,}597(x_3 - 5{,}714)_+ + 1{,}082x_4 + 1{,}384(x_4 - 5{,}943)_+ - 0{,}188(x_4 - 6{,}257)_+ + 0{,}109x_5 + 0{,}019(x_5 - 32{,}19)_+$$

Selanjutnya, model regresi nonparametrik yang dihasilkan dengan metode UBR yang paling minimum adalah sebagai berikut.

$$\tilde{y} = -0{,}003 - 0{,}003x_1 + 0{,}075(x_1 - 13{,}94)_+ + 0{,}003x_2 - 0{,}006(x_2 - 4{,}77)_+ + 0{,}002x_3 + 0{,}025(x_3 - 5{,}714)_+ - 0{,}490x_4 - 0{,}004(x_4 - 6{,}257)_+ - 0{,}024x_5 - 0{,}023(x_5 - 70{,}66)_+$$

Berdasarkan hasil analisis data menunjukkan bahwa model spline dengan pemilihan titik knot optimal menggunakan metode GCV memberikan hasil yang lebih baik dibandingkan model spline dengan pemilihan titik knot optimal menggunakan metode UBR. Kesimpulan tersebut dengan memperhatikan kriteria kebaikan model, yaitu MSE yang relatif lebih kecil dan R^2_{adj} yang lebih besar serta dengan melihat hasil dari pengujian asumsi IIDN (0, σ²) dari residual. Hal ini dapat menjadi pembuktian bahwa metode GCV adalah lebih baik digunakan pada data *Gaussian* (berdistribusi normal).

5. Saran

Berdasarkan penelitian yang telah dilakukan, permasalahan yang dikaji dapat dikembangkan untuk penelitian-penelitian berikutnya, antara lain:

1. Perlu dikembangkan pemodelan mengenai angka kematian maternal dengan menggunakan model regresi nonparametrik spline kuadratik dan kubik.
2. Perlu dilakukan kajian metode pemilihan titik knot yang lain dan aplikasi pada data angka kematian maternal, seperti GML agar bisa dilakukan perbandingan yang lebih komprehensif.
3. Perlu dikembangkan titik knot yang digunakan dalam penelitian, 4 titik knot atau 5 titik knot.

REFERENSI

[1] Hardle, W,. 1990. *Applied Nonparametric Regression*.Cambridge University Press. New York.
[2] Budiantara, I.N.. 2009. Spline dalam Regresi Nonparametrik dan Semiparametrik Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang, ITS Press, Institut Teknologi Sepuluh Nopember Surabaya.
[3] Eubank, R.L,. 1988. *Spline Smoothing and Nonparametric Regression*. Marcel Dekker.Inc. New York.
[4] Budiantara, I.N. 2001. *Regresi Nonparametrik dan Semiparametrik Serta Perkembangannya*. Makalah Pembicara Utama pada Seminar Nasional Alumni Pasca Sarjana Matematika Universitas Gajah Mada, Yogyakarta.
[5] Wahba, G. 1990. *Spline Models For Observasion Data*. SIAM Pensylvania
[6] Wang, Y. 1998. "Smoothing Spline Models With Correlated Random Errors", Journal of The American Statistical Association. Vol. 93. No. 441. Hal. 341-348
[7] Safrudin dan Hamidah. 2009. *Kebidanan Komunitas*. Jakarta: EGC.
[8] Budiantara, I.N. 2001. *Regresi Nonparametrik dan Semiparametrik Serta Perkembangannya*. Makalah Pembicara Utama pada Seminar Nasional Alumni Pasca Sarjana Matematika Universitas Gajah Mada, Yogyakarta.
[9] Budiantara, I.N. 2005. *Model Keluarga Spline Polinomial Truncated dalam Regresi Semiparametrik*. Surabaya: Berkala MIPA, Institut Teknologi Sepuluh Nopember
[10] Drapper, N.R. dan Smith, H,. 1992. *Analisis Regresi Terapan*. Edisi Kedua. PT GramediaPustakaUtama. Jakarta.
[11] Wei, W.W.S., 1990. *Time Series Univariate and Multivariate Methods*. Canada: Addison Wesley Publishing Company, Inc.
[12] Kusumawati, D A. 2012. Gambaran Faktor Penyebab Kematian Maternal di Wilayah Kerja Dinas Kesehatan Kabupaten Magetan Tahun 2010. Magetan: Politeknik Kesehatan Kemenkes Surabaya
[13] Arfan, N. 2013. Pemodelan regresi nonparametrik spline pada kasus angka kematian maternal di Jawa Timur. Surabaya: Institut Teknologi Sepuluh Nopember