

PERBANDINGAN PEMILIHAN TITIK KNOT OPTIMAL MENGGUNAKAN METODE UNBIASED RISK (UBR) DAN GENERALIZED CROSS VALIDATION (GCV) DALAM REGRESI NONPARAMETRIK SPLINE

(Studi Kasus: Jumlah Kematian Bayi di Provinsi Jawa Tengah 2017)

Eriana Medita Purbasari¹, Tiani Wahyu Utami², M. Al Haris, S.Si, M.Si³

123Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Muhammadiyah Semarang

e-mail : erianamedita@gmail.com

ABSTRAK

Regresi nonparametrik spline adalah sebuah metode regresi nonparametrik yang mengestimasi pola data yang tidak membentuk pola tertentu dengan bantuan titik knot. Model terbaik diperoleh dari titik knot optimal. Penelitian ini membahas mengenai perbandingan model regresi nonparametrik spline dengan menggunakan metode UBR dan GCV sebagai metode pemilihan titik knot optimal pada kasus jumlah kematian bayi di Jawa Tengah. Kriteria pemilihan model terbaik yang digunakan berdasarkan nilai Mean Squared Error (MSE) dan R^2 . Kematian bayi adalah kematian yang terjadi antara saat setelah bayi dilahirkan sampai sebelum bayi telah berusia satu tahun. Faktor-faktor yang diduga berpengaruh terhadap kematian bayi antara lain persentase bayi yang diberi vitami A, rata-rata lama sekolah wanita berstatus kawin, jumlah tenaga kesehatan, persentase bayi yang tidak diberi ASI eksklusif dan persentase persalinan yang menggunakan non medis. Penelitian ini menghasilkan model regresi nonparametrik spline terbaik dari metode UBR diperoleh dengan menggunakan tiga titik knot yang menghasilkan nilai MSE sebesar 738,67 dan R^2 sebesar 85,65%. Sedangkan, model regresi nonparametrik spline terbaik dari metode GCV diperoleh dengan menggunakan tiga titik knot yang menghasilkan nilai MSE sebesar 121,43 dan R^2 sebesar 97,64%. Sehingga, dapat disimpulkan bahwa dalam penelitian ini metode yang lebih tepat digunakan untuk pemilihan titik knot optimal adalah metode GCV karena menghasilkan nilai MSE yang lebih kecil dan R^2 yang lebih besar dibandingkan dengan metode UBR.

Kata kunci : Regresi Nonparametrik, Spline, *Unbiased Risk (UBR)*, *Generalized Cross Validation (GCV)*, Kematian Bayi

ABSTRACT

Spline nonparametric regression is a nonparametric regression method that estimates data patterns that do not form certain patterns with the help of knots. The best model is obtained from the optimal knot. This study discusses about comparison of spline nonparametric regression models using the UBR and GCV methods as selecting optimal knot in cases of infant mortality in Central Java. Criteria for selecting the best model are based on the Mean Squared Error (MSE) and R^2 . Infant death is the death of the death of young children under the age of one. Factors that are thought to influence infant mortality include the percentage of infants given vitami A, the average length of schooling of married women, the number of health workers, the percentage of babies who are not exclusively breastfed and the percentage of non-medical births using. This research resulted in the best spline nonparametric regression model from the UBR method obtained using three point knots with MSE of 738.67 and R^2 of 85.65%. Meanwhile, the best spline nonparametric regression model from the GCV method was obtained using three point knots with MSE of 121.43 and R^2 of 97.64%. In conclusion, in this study the more appropriate method used for the selection of optimal point knots is the GCV method because it produces a smaller MSE and larger R^2 compared to the UBR method.

Keywords : *Nonparametric Regression, Spline, Unbiased Risk (UBR), Generalized Cross Validation (GCV), The Infant Mortality*

PENDAHULUAN

Regresi nonparametrik adalah metode yang digunakan apabila pola dari kurva regresi tidak diketahui. Regresi nonparametrik tidak terikat dengan asumsi tertentu, sehingga regresi nonparametrik ini memiliki fleksibilitas yang tinggi. Menurut Eubank, R.L (1988), pembentukan estimasi dari kurva regresi diharapkan dapat menyesuaikan data tanpa dipengaruhi faktor subyektifitas oleh peneliti. Hal tersebut adalah yang menjadi kelebihan regresi nonparametrik. Diantara beberapa model regresi nonparametrik tersebut, metode spline merupakan metode yang cukup sering digunakan.

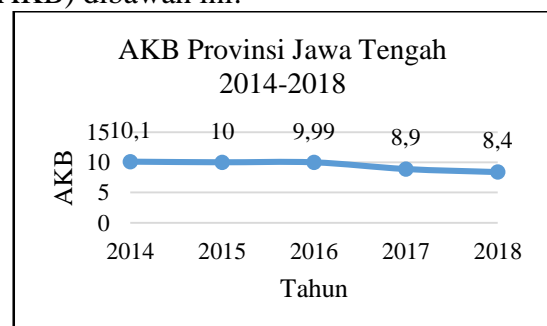
Spline merupakan jenis *piecewise* polinomial yang berarti polinomial yang memiliki sifat tersegmen. Menurut Eubank (1988), spline memiliki beberapa kelebihan diantaranya kemampuan interpretasi visual yang baik dan memiliki fleksibilitas yang tinggi. Menurut Budiantara (2009), apabila sebuah data mempunyai pola yang berubah-ubah pada interval tertentu maka sebaiknya dimodelkan dengan regresi spline. Regresi nonparametrik spline mampu mengatasi pola data yang menunjukkan naik/turun yang tajam dengan bantuan titik-titik knot sehingga kurva yang dihasilkan relatif mulus (Hardle, 1990). Titik knot merupakan titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku sebuah data (Eubank, 1988; Budiantara, 2009). Ketepatan model regresi spline yang terbentuk dipengaruhi oleh pemilihan titik knot optimal. Terdapat beberapa metode yang dapat digunakan dalam pemilihan titik knot optimal dalam regresi spline, antara lain *Cross Validation* (CV) (Craven & Wahba, 1979), *Unbiased Risk* (UBR) (Wahba, 1990; Wang, 1998), *Generalized Cross Validation* (GCV) (Wahba, 1990), dan *Generalized Maximum Likelihood* (GML) (Eubank, 1990).

Metode *Unbiased Risk* (UBR) merupakan metode pemilihan titik knot optimal yang mengharuskan nilai estimasi dari varian kesalahan diketahui (Wang, 2011). Menurut Wahba (1990 dalam Wang, 1998),

metode UBR akan menghasilkan nilai yang baik pada data *non-Gaussian* atau tidak berdistribusi normal. Sedangkan metode *Generalized Cross Validation* (GCV) akan lebih baik digunakan pada data *Gaussian* atau berdistribusi normal. Kelebihan dari GCV adalah efisien serta sederhana dalam perhitungannya dan tidak memerlukan varian.

Penelitian yang menggunakan metode UBR dan GCV telah beberapa kali dilakukan. Wulandari *et al* (2017) yang memodelkan angka morbiditas di Jawa Timur Sari (2016) yang membandingkan metode *Generalized Cross Validation* (GCV) dan *Unbiased Risk* (UBR) dalam pemilihan titik knot optimal. Liu *et al* (2014) yang membandingkan metode UBR, GML, GACV dalam pemilihan titik knot optimal. Devi (2018) yang membandingkan metode *Unbiased Risk* (UBR) dan *Cross Validation* (CV) dalam pemilihan titik knot optimal.

Studi kasus yang digunakan dalam penelitian ini adalah jumlah kematian bayi. Menurut Aulele (2012), kematian bayi adalah kematian yang terjadi antara saat setelah bayi dilahirkan sampai sebelum bayi telat berusia satu tahun. Kematian bayi adalah salah satu indikator sensitif yang dapat digunakan untuk mengukur derajat kesehatan dan derajat kesejahteraan (Kementerian Kesehatan RI, 2017). Jawa Tengah merupakan salah satu provinsi yang ada di Indonesia dengan jumlah penduduk 34.257.865. Jumlah kematian bayi di Jawa Tengah dari tahun 2013 sampai dengan tahun 2017 terus mengalami penurunan yang dapat dilihat dari grafik Angka Kematian Bayi (AKB) dibawah ini:



Gambar 1 Grafik Angka Kematian Bayi (AKB) Provinsi Jawa Tengah Tahun 2014-2018

Berdasarkan Gambar 1 terlihat bahwa

dari tahun 2014 sampai dengan 2018 AKB di Jawa Tengah terus mengalami penurunan. Walaupun angka ini telah mengalami penurunan, akan tetapi tetap perlu dilakukan penelitian guna menjaga kestabilan bahkan dapat menurunkan AKB untuk tahun-tahun berikutnya.

Penelitian terkait kasus kematian bayi telah banyak dilakukan oleh peneliti terdahulu. Sary dan Latra (2013) menggunakan pendekatan Regresi Binomial Negatif dengan variabel yang signifikan adalah persentase bayi yang mendapatkan vitamin A. Puspitasari dan Sutanto (2013) menggunakan Regresi Spline, dengan variabel yang signifikan adalah persentase bayi yang tidak diberi ASI eksklusif, persentase persalinan yang menggunakan non medis, dan jumlah tenaga kesehatan. (Syam, 2017) menggunakan *Generalized Regresi Poisson* dengan variabel yang signifikan adalah jumlah tenaga kesehatan dan persalinan yang dilakukan dengan bantuan non medis. Elyna dan Srinadi (2012) menggunakan *Pendekatan Geographically Weighted Poisson Regression* dengan variabel yang signifikan adalah variabel rata-rata lama sekolah wanita berstatus kawin. Penelitian-penelitian tersebut mendasari penulis memilih variabel persentase bayi yang mendapatkan vitamin A, persentase bayi yang tidak diberi ASI eksklusif, persentase persalinan yang menggunakan non medis, jumlah tenaga kesehatan, variabel rata-rata lama sekolah wanita berstatus kawin sebagai variabel prediktor.

Berdasarkan uraian diatas, penelitian ini berfokus pada perbandingan metode UBR dan metode GCV dalam pemilihan titik knot optimal pada regresi nonparametrik spline dengan studi kasus jumlah kematian bayi di Provinsi Jawa Tengah tahun 2017.

TINJAUAN PUSTAKA

1. Regresi Nonparametrik

Regresi nonparametrik merupakan metode statistika yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor apabila bentuk dari kurva regresi yang terbentuk tidak membentuk pola atau tidak terdapat informasi mengenai

polanya. Berikut ini adalah model regresi nonparametrik secara umum (Eubank, 1988):

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, 3, \dots, n$$

dimana

Y_i = variabel respon

x_i = variabel prediktor

$f(x_i)$ = fungsi regresi

ε_i = error yang diasumsikan berdistribusi $N(0, \sigma^2)$.

2. Regresi Spline

Spline merupakan salah satu metode yang ada di regresi nonparametrik. Spline adalah potongan-potongan (*piecewise*) polinomial yang bersifat tersegmen. Berikut ini adalah fungsi spline secara umum dengan orde-m dengan titik knot $(K_1, K_2, K_3, \dots, K_p)$:

$$f(x_i) = \beta_0 + \beta_1 x_i^1 + \dots + \beta_m x_i^m + \sum_{q=1}^p \beta_{(m+q)} (x_i - k_q)_+^m$$

dari persamaan di atas apabila dijumlahkan akan menjadi:

$$f(x_i) = \sum_{j=0}^m \beta_j x_i^j + \sum_{q=1}^p \beta_{(m+q)} (x_i - k_q)_+^m$$

3. Pemilihan Titik Knot

a. Metode UBR

Metode *Unbiased Risk* (UBR) merupakan metode yang dapat digunakan untuk memilih titik knot optimal. Metode ini dapat digunakan ketika σ^2 diketahui. berikut ini adalah perumusan metode UBR :

$$\hat{R}(K) = \frac{1}{n} \|(I - A(k))y\|^2 - \frac{\sigma^2}{n} \text{tr}[I - A(k)]^2 + \frac{\sigma^2}{n} \text{tr}[A^2](k)$$

dimana

$$\hat{\sigma}^2 = \frac{\|[I - A(k)]y\|^2}{\text{tr}[I - A(k)]}$$

$$A(k) = X(X'X)^{-1}X'$$

I : matriks identitas

n : banyak penelitian

b. Metode GCV

Metode *Generalized Cross Validation* (GCV) merupakan salah satu metode sering digunakan dalam pemilihan titik knot yang optimal. Berikut ini adalah fungsi dari GCV (Eubank, 1988) :

$$GCV(k) = n^{-1} \frac{\sum_{i=1}^n (y_i - \hat{f}(x_i))^2}{[1 - n^{-1} \text{trace}(A(k))]^2}$$

dimana:

$$\hat{f}(x_i) = A(k)y = [X(X'X)^{-1}X']y$$

$$A(k) = [X(X'X)^{-1}X']$$

n = banyak pengamatan

4. Kriteria Pemilihan Model Terbaik

Nilai *Mean Square Error* (MSE) adalah salah satu kriteria yang dapat digunakan untuk menentukan model terbaik. Menurut Draper dan Smith (1992), rumus untuk menghitung nilai MSE adalah sebagai berikut:

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - m - 1}$$

Untuk mengetahui seberapa jauh model yang tersebut mampu menerangkan variasi dari variabel respon dapat menggunakan koefisien determinasi (R^2). Berikut ini adalah rumus koefisien determinasi (R^2) :

$$R^2 = \frac{SSR}{SST}$$

dimana :

$$SSR = \hat{\beta}' X'y - n\bar{y}^2$$

$$SST = y'y - n\bar{y}^2$$

5. Kematian bayi

Kematian bayi adalah kematian yang terjadi antara saat setelah bayi lahir sampai bayi belum tepat berusia satu tahun. Terdapat banyak faktor yang menyebabkan kematian bayi. Dari sisi penyebabnya, kematian bayi dibagi menjadi dua macam yaitu endogen dan eksogen (Dinkes Tengah, 2017) . Berikut ini adalah faktor yang diduga berpengaruh terhadap kematian bayi:

- Persentase bayi yang mendapat vitamin A
- Rata-rata lama sekolah wanita yang berstatus kawin
- Jumlah tenaga kesehatan

- Persentase bayi yang tidak diberi ASI eksklusif
- Persentase persalinan menggunakan tenaga non medis

METODE PENELITIAN

1. Sumber Penelitian

Data yang digunakan dalam penelitian ini adalah data sekunder yang bersumber dari publikasi Profil Kesehatan Provinsi Jawa Tengah tahun 2017 oleh Dinas Kesehatan Provinsi Jawa Tengah dan Badan Pusat Statistik (BPS).

2. Variabel Penelitian

Variabel respon yang digunakan dalam penelitian ini adalah jumlah kematian bayi di provinsi Jawa Tengah pada tahun 2017. Sedangkan variabel prediktor yang digunakan adalah Persentase bayi yang mendapat vitamin A, Rata-rata lama sekolah wanita berstatus kawin, Jumlah tenaga kesehatan, Persentase bayi yang tidak diberi ASI eksklusif dan Persentase persalinan yang menggunakan non medis

3. Langkah-langkah penelitian

Adapun langkah-langkah penelitian ini adalah:

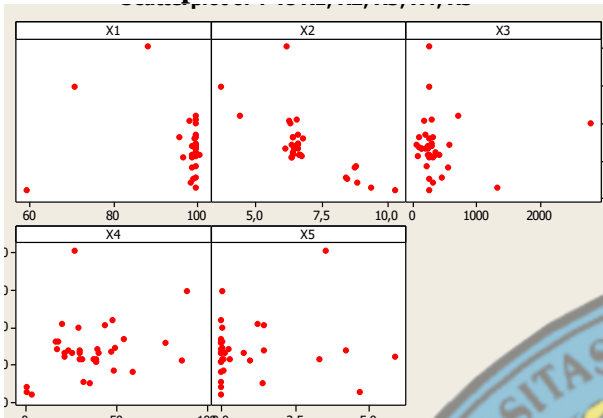
- Membuat *scatter plot* antara variabel respon (y) dengan masing-masing variabel prediktor (x).
- Memodelkan data dengan regresi nonparametrik spline menggunakan satu titik knot, dua titik knot, dan tiga titik knot dengan metode UBR.
- Memilih titik knot optimal menggunakan metode UBR.
- Memodelkan data dengan regresi nonparametrik spline menggunakan titik knot optimal dari metode UBR.
- Memodelkan data dengan regresi nonparametrik spline menggunakan satu titik knot, dua titik knot, dan tiga titik knot dengan metode GCV.
- Memilih titik knot optimal menggunakan metode GCV.
- Memodelkan data dengan regresi nonparametrik spline menggunakan titik knot optimal dari metode GCV.
- Membandingkan model regresi nonparametrik spline dengan titik knot

optimal yang diperoleh dari metode UBR dan metode GCV.

9. Memilih model terbaik dengan kriteria berdasarkan R^2 dan MSE.
10. Menginterpretasikan hasil analisis yang didapatkan.
11. Menarik kesimpulan.

HASIL DAN PEMBAHASAN

1. Scatter plot variabel respon dengan variabel prediktor



Gambar 2 Scatter Plot

Gambar 2 menunjukkan bahwa pola hubungan variabel respon dan prediktor tidak membentuk sehingga dapat didekati menggunakan pendekatan regresi nonparametrik.

2. Pemodelan Jumlah Kematian Bayi di Provinsi Jawa Tengah Tahun 2017 Menggunakan Metode UBR

Langkah awal sebelum melakukan pemodelan adalah menentukan titik knot optimal. Berikut adalah hasil pemilihan titik knot optimal dengan metode UBR dengan menggunakan, satu titik knot, dua titik knot, tiga titik knot:

Tabel 1 Ringkasan Nilai UBR

Titik Knot	Nilai UBR
Satu Titik Knot	$8,4 \times 10^{-22}$
Dua Titik Knot	$2,05 \times 10^{-22}$
Tiga Titik Knot	$1,96 \times 10^{-22}$

Tabel 1 terlihat bahwa nilai UBR minimum sebesar $8,4 \times 10^{-22}$. Sehingga titik knot optimal diperoleh dari tiga titik knot. Berikut ini adalah hasil estimasi parameter dari model regresi nonparametrik spline dengan tiga titik knot:

Tabel 2 Estimasi Parameter

Parameter	Estimator
β_0	$\widehat{\beta}_0 = 1,22$
β_1	$\widehat{\beta}_1 = 4,16$
β_2	$\widehat{\beta}_2 = 11,09$
β_3	$\widehat{\beta}_3 = -33,47$
β_4	$\widehat{\beta}_4 = -8,38$
β_5	$\widehat{\beta}_5 = -8,98$
β_6	$\widehat{\beta}_6 = -35,61$
β_7	$\widehat{\beta}_7 = 18,13$
β_8	$\widehat{\beta}_8 = 2,25$
β_9	$\widehat{\beta}_9 = 0,03$
β_{10}	$\widehat{\beta}_{10} = -0,04$
β_{11}	$\widehat{\beta}_{11} = 0,07$
β_{12}	$\widehat{\beta}_{12} = 0,003$
β_{13}	$\widehat{\beta}_{13} = 2,83$
β_{14}	$\widehat{\beta}_{14} = -2,71$
β_{15}	$\widehat{\beta}_{15} = 0,20$
β_{16}	$\widehat{\beta}_{16} = -51,30$
β_{17}	$\widehat{\beta}_{17} = 17,96$
β_{18}	$\widehat{\beta}_{18} = -18,18$
β_{19}	$\widehat{\beta}_{19} = 11,90$
β_{20}	$\widehat{\beta}_{20} = -448,22$

Berdasarkan tabel diatas, maka hasil pemodelan regresti nonparametrik spline adalah sebagai berikut:

$$\begin{aligned} \hat{y} = & 1,22 + 4,16x_1 + 11,09(x_1 - 64,32)_+ \\ & -33,47(x_1 - 82,15)_+ - 8,38(x_1 - 99,98)_+ \\ & -8,98x_2 - 35,61(x_2 - 4,52)_+ + 18,13 \\ & (x_2 - 7,34)_+ + 2,25(x_2 - 10,17)_+ + \\ & 0,03x_3 - 0,04(x_3 - 393,39)_+ + 0,07 \\ & (x_3 - 1567,24)_+ + 0,003(x_3 - 2741,10)_+ \\ & + 2,81x_4 - 2,71(x_4 - 10,88)_+ + 0,20 \\ & (x_4 - 48,97)_+ - 51,30(x_4 - 87,07)_+ + 17,96x_5 \\ & - 18,18(x_5 - 0,72)_+ + 11,90(x_5 - 3,25)_+ \\ & - 448,22(x_5 - 5,77)_+ \end{aligned}$$

3. Pemodelan Jumlah Kematian Bayi di Provinsi Jawa Tengah Tahun 2017 Menggunakan Metode GCV

Langkah awal sebelum melakukan pemodelan adalah menentukan titik knot optimal. Berikut adalah hasil pemilihan titik knot optimal dengan metode GCV dengan

menggunakan, satu titik knot, dua titik knot, tiga titik knot:

Tabel 3 Ringkasan Nilai GCV

Titik Knot	Nilai GCV
Satu Titik Knot	1500,61
Dua Titik Knot	972,82
Tiga Titik Knot	514,70

Tabel 3 terlihat bahwa nilai GCV minimum sebesar 514,70. Sehingga titik knot optimal diperoleh dari tiga titik knot. Berikut ini adalah hasil estimasi parameter dari model regresi nonparametrik spline dengan tiga titik knot:

Tabel 4 Estimasi Parameter

Parameter	Estimator
β_0	$\widehat{\beta}_0 = -3,10$
β_1	$\widehat{\beta}_1 = -7,23$
β_2	$\widehat{\beta}_2 = 165,84$
β_3	$\widehat{\beta}_3 = -215,03$
β_4	$\widehat{\beta}_4 = 56,91$
β_5	$\widehat{\beta}_5 = 101,02$
β_6	$\widehat{\beta}_6 = -17,21$
β_7	$\widehat{\beta}_7 = -120,60$
β_8	$\widehat{\beta}_8 = -5,34$
β_9	$\widehat{\beta}_9 = 0,02$
β_{10}	$\widehat{\beta}_{10} = 2,10$
β_{11}	$\widehat{\beta}_{11} = -5,70$
β_{12}	$\widehat{\beta}_{12} = 10,62$
β_{13}	$\widehat{\beta}_{13} = 1,56$
β_{14}	$\widehat{\beta}_{14} = -5,49$
β_{15}	$\widehat{\beta}_{15} = 6,99$
β_{16}	$\widehat{\beta}_{16} = -13,56$
β_{17}	$\widehat{\beta}_{17} = 45,19$
β_{18}	$\widehat{\beta}_{18} = -100,12$
β_{19}	$\widehat{\beta}_{19} = 58,05$
β_{20}	$\widehat{\beta}_{20} = -34,60$

Berdasarkan tabel diatas, maka hasil pemodelan regresti nonparametrik spline adalah sebagai berikut:

$$\hat{y} = -3,10 - 7,23x_1 + 165,84(x_1 - 67,71)_+ - 215,03(x_1 - 74,51)_+ + 56,91(x_1 - 93,19)_+ + 101,02x_2 - 17,21(x_2 - 5,05)_+ - 120,60(x_2 - 6,13)_+ - 5,34(x_2 - 9,09)_+ + 0,02x_3 + 2,10(x_3 - 616,98)_+ - 5,70(x_3 - 1064,16)_+ + 10,62$$

$$(x_3 - 2293,92)_+ + 1,56x_4 - 5,49(x_4 - 18,84)_+ + 6,99(x_4 - 32,65)_+ - 13,56(x_4 - 72,56)_+ + 45,19x_5 - 100,12(x_5 - 1,20)_+ + 58,05(x_5 - 2,16)_+ - 34,60(x_5 - 4,81)_+$$

4. Perbandingan Metode UBR dan Metode GCV dalam Pemodelan Regresi Nonparametrik Spline

Berikut ditampilkan tabel perbandingan metode UBR dan GCV dalam pemodelan regresi nonparametrik spline pada data Jumlah Kematian Bayi:

Tabel 5 Perbandingan Metode UBR dan GCV

	MSE	R ²
Metode UBR	738,67	85,65
Metode GCV	121,43	97,64

Berdasarkan tabel diatas terlihat bahwa metode GCV menghasilkan nilai MSE yang lebih kecil dari metode UBR serta menghasilkan nilai R² yang lebih besar dari metode UBR, maka metode GCV merupakan metode yang lebih baik digunakan untuk pemilihan titik knot optimal dibandingkan metode UBR dalam pemodelan regresi nonparametrik spline pada kasus jumlah kematian bayi di Provinsi Jawa Tengah tahun 2017.

KESIMPULAN

1. Model regresi nonparametrik spline terbaik menggunakan metode UBR diperoleh dari tiga titik knot. Model yang terbentuk adalah sebagai berikut:

$$\hat{y} = 1,22 + 4,16x_1 + 11,09(x_1 - 64,32)_+ - 33,47(x_1 - 82,15)_+ - 8,38(x_1 - 99,98)_+ - 8,98x_2 - 35,61(x_2 - 4,52)_+ + 18,13(x_2 - 7,34)_+ + 2,25(x_2 - 10,17)_+ + 0,03x_3 - 0,04(x_3 - 393,39)_+ + 0,07(x_3 - 1567,24)_+ + 0,003(x_3 - 2741,10)_+ + 2,81x_4 - 2,71(x_4 - 10,88)_+ + 0,20(x_4 - 48,97)_+ - 51,30(x_4 - 87,07)_+ + 17,96x_5 - 18,18(x_5 - 0,72)_+ + 11,90(x_5 - 3,25)_+$$

$$-448,22(x_5 - 5,77)_+$$

2. Model regresi nonparametrik spline terbaik menggunakan metode GCV diperoleh dari tiga titik knot. Model yang terbentuk adalah sebagai berikut:

$$\begin{aligned} \hat{y} = & -3,10 - 7,23x_1 + 165,84(x_1 - 67,71)_+ \\ & -215,03(x_1 - 74,51)_+ + 56,91(x_1 - 93,19)_+ \\ & +101,02x_2 - 17,21(x_2 - 5,05)_+ - 120,60 \\ & (x_2 - 6,13)_+ - 5,34(x_2 - 9,09)_+ + 0,02x_3 + 2,10 \\ & (x_3 - 616,98)_+ - 5,70(x_3 - 1064,16)_+ + 10,62 \\ & (x_3 - 2293,92)_+ + 1,56x_4 - 5,49(x_4 - 18,84)_+ \\ & +6,99(x_4 - 32,65)_+ - 13,56(x_4 - 72,56)_+ \\ & +45,19x_5 - 100,12(x_5 - 1,20)_+ + 58,05 \\ & (x_5 - 2,16)_+ - 34,60(x_5 - 4,81)_+ \end{aligned}$$

3. Model terbaik dengan metode UBR menghasilkan nilai MSE sebesar 738,67 dengan R^2 sebesar 85,65%, sedangkan model terbaik dengan metode GCV menghasilkan nilai MSE sebesar 121,43 dengan R^2 sebesar 97,64%. Sehingga dapat disimpulkan bahwa dalam penelitian ini metode GCV menghasilkan model yang lebih baik dibandingkan dengan metode UBR karena menghasilkan nilai MSE yang lebih kecil dan nilai R^2 yang lebih besar.

DAFTAR PUSTAKA

- Aulele, S. N. (2012). Pemodelan Jumlah Kematian Bayi Di Provinsi Maluku Tahun 2010 dengan Menggunakan Regresi Poisson. *Jurnal Berekeng*, 5(2), 23–27. Retrieved from https://ejournal.unpatti.ac.id/ppr_iteminfo_ink.php?id=491
- Budiantara, I.N. (2009). Spline dalam Regresi Nonparametrik dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang. In *Pidato Pengukuhan untuk Jabatan Guru Besar dalam Bidang Ilmu Matematika Statistika dan Probabilitas*. Institut Teknologi Sepuluh Nopember: Departemen Statistika.
- Craven, P., & Wahba, G. (1979). Smoothing Noisy Data with Spline Function - Estimating the Correct Degree of Smoothing by the Method of Generalized Cross Validation. *Numerische Mathematik*, 31, 377–403. <https://doi.org/10.1007/BF01404567>
- Devi, A. R. (2018). *Metode Unbiased Risk (UBR) dan Cross Validation (CV) Untuk Pemilihan Titik Knot Optimal dalam Regresi Nonparametrik Multivariabel Spline Truncated (Studi Kasus : Data Tingkat Pengangguran Terbuka di Provinsi Jawa Tengah Tahun 2015)*. Institut Teknologi Sepuluh Nopember.
- Draper, N., & Smith, H. (1992). *Analisis Regresi Terapan Edisi Kedua* (Kedua). Jakarta: Gramedia Pustaka Utama.
- Elyna, M. A. K., & Srinadi, I. G. A. M. (2012). Pemodelan Angka Kematian Bayi Dengan Pendekatan Geographically Weighted Poisson Regression Di Provinsi Bali. *E-Jurnal Matematika*, 1(1), 94–98. Retrieved from <http://oaji.net/articles/2015/2656-1449846067.pdf>
- Eubank, R.L. (1988). *Nonparametric Regression and Spline Smoothing* (Second Edi). New York: Marcel Dekker, Inc.
- Hardle, W. (1990). Applied Nonparametric Regression. In *Journal of the Royal Statistical Society. Series A (Statistics in Society)*. <https://doi.org/10.2307/2982873>
- Kementrian Kesehatan RI (2017). *Profil Kesehatan Indonesia Tahun 2017*. Jakarta: Kementrian Kesehatan Republik Indonesia.
- Liu, A., Tong, T., & Wang, Y. (2014). Smoothing Spline Estimation of Variance Functions. *Journal of Computational and Graphical Statistics*, 16(2), 312–329. <https://doi.org/10.1198/106186007X204>

- Puspitasari, E., & Sutanto, H. T. (2013). Model Regresi Spline Knot Optimal Untuk Mengetahui Faktor – Faktor yang Mempengaruhi Jumlah Kematian Di Jawa Timur. *Jurusan Matematika, Fakultas Mipa, Universitas Negeri Surabaya*, 1–7. Retrieved from <https://media.neliti.com/media/publications/16096-ID-pemodelan-jumlah-kematian-bayi-di-provinsi-jawa-timur-tahun-2011-dengan-pendekat.pdf>
- Rakhmawati, M., Susongko, P., & Rohman, M. S. (2019). Pengaruh Indeks Sosial Ekonomi Budaya (ESCS) dan Dukungan Guru terhadap Prestasi Belajar Matematika Siswa Indonesia Berdasarkan Survey PISA Tahun 2015. *Jurnal Pendidikan MIPA Pancasakti*, 3(1), 11–23.
- Sari, S. U. R. (2016). *Perbandingan Model Regresi Nonparametrik Spline Multivariabel dengan Menggunakan Metode Generalized Cross Validation (GCV) dan Unbiased Risk (UBR) dalam Pemilihan Titik Knot Optimal (Studi Kasus Data Angka Kematian Maternal di Jawa Timur)*. Institut Teknologi Sepuluh Nopember.
- Sary, S. A., & Latra, I. N. (2013). Pemodelan Jumlah Kematian Bayi di Provinsi Jawa Timur Tahun 2011 dengan Pendekatan Regresi Binomial Negatif. *Jurnal Sains Dan Seni Pomits*, 2(2), 2337–3520. Retrieved from <https://media.neliti.com/media/publications/16096-ID-pemodelan-jumlah-kematian-bayi-di-provinsi-jawa-timur-tahun-2011-dengan-pendekat.pdf>
- Syam, A. R. (2017). *Pemodelan Generalized Regresi Poisson Pada Faktor- Faktor yang Mempengaruhi Angka Kematian Bayi di Provinsi Sulawesi Selatan Tahun 2014* (Universitas Islam Negeri (UIN) Alauddin Makassar). Retrieved from <http://repositori.uin->
- Wahba, G. (1990). Spline Models for Observational Data. *Society for Industrial and Applied Mathematics, Philadelphia*, 59, 502.
- Wang, Y. (1998). Smoothing Spline Models With Correlated Random Errors. *Journal of the American Statistical Association*, 93(441), 341–348. <https://doi.org/10.1080/01621459.1998.10474115>
- Wulandari, K., Budiantara, I. N., & Ratna, M. (2017). Pemodelan Faktor-Faktor yang Mempengaruhi Angka Morbiditas di Jawa Timur Menggunakan Regresi Nonparametrik Spline. *Jurnal Sains Dan Seni ITS*, 6(1). <https://doi.org/10.12962/j23373520.v6i1.22477>