

Pemodelan Regresi Nonparametrik dengan Estimator *Spline Truncated* dan Deret Fourier

Andrea Tri Rian Dani^{1*}, Narita Yuri Adrianingsih²

^{1,2} Departemen Statistika, Fakultas Sains dan Analitika Data, Institut Teknologi Sepuluh Nopember, Jl. Arif Rahman Hakim, Kota Surabaya 60111, Jawa Timur, Indonesia

* Penulis Korespondensi. Email: andrikadoko@gmail.com

ABSTRAK

Pendekatan regresi nonparametrik digunakan apabila hubungan antara variabel prediktor dan variabel respon tidak diketahui polanya. *Spline truncated* dan deret Fourier merupakan estimator dalam pendekatan nonparametrik yang terkenal, karena memiliki fleksibilitas yang tinggi dan mampu menyesuaikan terhadap sifat lokal data secara efektif. Penelitian ini bertujuan untuk mendapatkan estimator model regresi nonparametrik terbaik menggunakan *spline truncated* dan deret Fourier. Metode estimasi kurva regresi nonparametrik dilakukan dengan menyelesaikan optimasi *Ordinary Least Squares* (OLS). Kriteria kebaikan model menggunakan GCV, R^2 dan MSE. Pemodelan regresi nonparametrik diterapkan pada data *Case Fatality Rate* (CFR) akibat Demam Berdarah *Dengue* (DBD) di Indonesia. Berdasarkan hasil analisis, hasil estimasi dari pemodelan regresi nonparametrik menunjukkan bahwa estimator *spline truncated* memberikan performa yang lebih baik dibandingkan estimator deret Fourier. Hal ini ditunjukkan dengan nilai R^2 dari estimator *spline truncated* yaitu sebesar 91,80% dan MSE sebesar 0,04, sedangkan dengan estimator deret Fourier diperoleh nilai R^2 sebesar 65,44% dan MSE sebesar 0,19.

Kata Kunci:

Regresi Nonparametrik; *Spline Truncated*; Deret Fourier

ABSTRACT

The nonparametric regression approach is used when the relationship between the predictor variable and the response variable is unknown. *Spline truncated* and Fourier series are well-known estimators in the nonparametric approach because they have high flexibility and are able to adjust to the local properties of the data effectively. This study aims to obtain the best nonparametric regression model estimator using the truncated spline and the Fourier series. The nonparametric regression curve estimation method is done by completing the *Ordinary Least Squares* (OLS) optimization. The criteria for the goodness of the model use GCV, R^2 , and MSE. Nonparametric regression modeling is applied to *Case Fatality Rate* (CFR) modeling due to *Dengue Hemorrhagic Fever* (DBD) in Indonesia. Based on the analysis, the estimation results from the nonparametric regression modeling show that the truncated spline estimator provides better performance than the Fourier series estimator. This is shown by the R^2 value of the

truncated spline estimator which is 91.80% and the MSE is 0.04, while the Fourier series estimator obtained an R^2 value of 65.44% and MSE of 0.19.

Keywords:

Nonparametric Regression; Spline Truncated; Fourier Series

Format Sitasi:

A.T.R. Dani and N.Y. Adrianingsih, "Pemodelan Regresi Nonparametrik dengan Estimator *Spline Truncated* dan Deret Fourier," *Jambura J. Math.*, vol. 3, no. 1, pp. 26-36, 2021.

1. Pendahuluan

Penyakit Demam Berdarah *Dengue* (DBD) disebabkan karena infeksi *virus dengue*, sehingga menimbulkan dampak ekonomi dan sosial. Jumlah penderita DBD semakin bertambah, dikarenakan tingginya mobilitas dan jumlah penduduk. Pada tahun 1968, pertama kali demam berdarah ditemukan di Surabaya, dengan angka kematiannya 41,30%, dan menyebar ke seluruh Indonesia hingga saat ini [1]. Berdasarkan dari data WHO, kasus DBD terbesar kedua di dunia adalah Indonesia. Tahun 2014 ada 100.347 penderita kasus DBD dan banyak yang meninggal adalah 907 orang [2]. Tahun 2015 penderita DBD meningkat menjadi 126.675 orang dan yang meninggal ada 1.229 orang. Pada tahun 2016 penderita DBD melambung tinggi, peningkatannya yaitu sebanyak 204.171 orang. DBD sampai saat ini masih jadi masalah kesehatan yang mendapat perhatian dunia [3]. Beberapa penelitian mengenai kasus DBD telah banyak dilakukan seperti yang terdapat pada [4][5][6][7]. Berdasarkan beberapa penelitian sebelumnya, dapat diketahui bahwa faktor-faktor yang diduga mempengaruhi angka kematian akibat DBD adalah persentase penduduk miskin, persentase rumah tangga kumuh, persentase rumah tangga sanitasi layak, dan persentase rumah tangga menurut sumber air minum layak.

Analisis regresi merupakan analisis untuk menyelidiki pola hubungan antara variabel prediktor dengan variabel respon dan dapat digunakan untuk melakukan prediksi [8][9]. Ada 3 pendekatan dalam analisis regresi yaitu regresi parametrik, nonparametrik, dan semiparametrik. Pada regresi parametrik, harus memenuhi asumsi dan memiliki sifat yang kaku, artinya bentuk kurva regresi harus diketahui. Regresi nonparametrik, apabila kurva diasumsikan tidak diketahui, sedangkan pendekatan regresi semiparametrik merupakan kombinasi antara parametrik dan nonparametrik. Sering kali dalam prakteknya, bentuk pola hubungan antara variabel prediktor dan respon tidak diketahui, sehingga pendekatan yang disarankan adalah regresi nonparametrik.

Kurva regresi nonparametrik diasumsikan mulus (*smooth*) yaitu termuat dalam suatu fungsi tertentu. Pendekatan regresi nonparametrik memiliki fleksibilitas yang sangat tinggi, sehingga diharapkan data mampu menyesuaikan sendiri bentuk estimasi kurva regresi tanpa dipengaruhi oleh subyektifitas peneliti [10]. Pada pendekatan regresi nonparametrik, terdapat beberapa estimator yang dikembangkan oleh para peneliti, diantaranya adalah Kernel [9][10][11][12], *spline truncated* [6][8][13][14] dan deret Fourier [15][16][17].

Estimator yang paling sering digunakan adalah *spline truncated*. *Spline truncated* merupakan potongan polinomial yang sifatnya tersegmen dan kontinu. Kelebihan dari *spline truncated* yaitu model ini fleksibel, dikarenakan *spline truncated* memiliki titik-titik knot yang menunjukkan adanya perubahan pola data [8][15]. Selain itu metode yang

menjadi sorotan lainnya adalah deret Fourier yang dipopulerkan oleh Bilodeau [18]. Deret Fourier merupakan polinomial trigonometri yang punya fleksibilitas tinggi, sehingga dapat menyesuaikan sifat lokal data secara efektif [19][20]. Deret Fourier juga baik digunakan untuk kurva yang polanya berulang [21].

Pada makalah ini dilakukan pemodelan kasus *Case Fatality Rate* (CFR) akibat DBD di Indonesia dengan pendekatan regresi nonparametrik menggunakan estimator *spline truncated* dan deret Fourier. Pendekatan regresi nonparametrik dapat memberikan fleksibilitas yang tinggi, yang tidak tergantung pada asumsi bentuk kurva regresi tertentu, sementara kedua estimator yang digunakan memiliki kelebihan masing-masing. Oleh karena itu, perlu dicari estimator terbaik dalam memodelkan data kasus CFR akibat DBD di Indonesia.

2. Metode

Data yang digunakan dalam penelitian ini adalah data sekunder yang dipublikasikan oleh BPS (Badan Pusat Statistik) dan Kementerian Kesehatan Republik Indonesia (Kemenkes RI) dalam Laporan Statistik Lingkungan Hidup dan Profil Kesehatan Indonesia, tentang *Case Fatality Rate* (CFR) akibat Demam Berdarah *Dengue* (DBD) di 33 provinsi Tahun 2018 [3][22].

Variabel-variabel yang diduga mempengaruhi CFR akibat DBD di Indonesia yaitu persentase penduduk miskin, persentase rumah tangga kumuh, persentase rumah tangga yang memiliki akses sanitasi layak, dan persentase rumah tangga yang memiliki akses sumber air minum layak.

Analisis data yang digunakan adalah dengan menggunakan pendekatan regresi nonparametrik dengan estimator *spline truncated* dan deret Fourier. Untuk menyelesaikan tujuan penelitian, maka disusun langkah penelitian dengan tahapan sebagai berikut:

2.1 Mendefinisikan model regresi nonparametrik

Diberikan data berpasangan dengan variabel prediktor sebanyak 4 sehingga dapat dituliskan $(x_{1i}, x_{2i}, x_{3i}, x_{4i}, y_i)$. Selanjutnya diasumsikan pola antara masing-masing variabel prediktor $(x_{1i}, x_{2i}, x_{3i}, x_{4i})$ terhadap variabel respon (y_i) mengikuti model regresi nonparametrik. Kurva regresi yang dibentuk dari variabel prediktor diasumsikan bersifat aditif, sehingga:

$$f(x_{1i}, x_{2i}, x_{3i}, x_{4i}) = f_1(x_{1i}) + f_2(x_{2i}) + \dots + f_4(x_{4i}) = \sum_{p=1}^4 f_p(x_{pi}) \quad (1)$$

2.2 Menghampiri kurva regresi nonparametrik menggunakan estimator *spline truncated*.

Data berpasangan x_i dan y_i , dimana $i=1,2,\dots,n$ yang merupakan model regresi nonparametrik. Didapatkan persamaan *spline truncated* berikut:

$$y_i = \sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - K_k)_+^m + \varepsilon_i; i=1,2,\dots,n \quad (2)$$

dengan fungsi *truncated* diberikan oleh:

$$(x_i - K_k)_+^m = \begin{cases} (x_i - K_k)^m & x \geq K_k \\ 0 & x < K_k \end{cases} \quad (3)$$

Model regresi pada Persamaan (2) diubah menjadi bentuk matriks sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1^1 & x_1^2 & \cdots & x_1^m : (x_1 - K_1)_+^m & \cdots & (x_1 - K_r)_+^m \\ 1 & x_2^1 & x_2^2 & \cdots & x_2^m : (x_2 - K_1)_+^m & \cdots & (x_2 - K_r)_+^m \\ \vdots & \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ 1 & x_n^1 & x_n^2 & \cdots & x_n^m : (x_n - K_1)_+^m & \cdots & (x_n - K_r)_+^m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{m+r} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (4)$$

Maka menjadi:

$$\vec{Y} = \mathbf{X}\vec{\beta} + \vec{\varepsilon} \quad (5)$$

Selanjutnya parameter $\vec{\beta}$ dapat diestimasi dengan menggunakan *Ordinary Least Squares* (OLS) yaitu suatu metode untuk memperoleh estimator dengan cara meminimumkan jumlah kuadrat *error*. Adapun *error* dari Persamaan (5) diberikan oleh:

$$\begin{aligned} \varepsilon_i &= y_i - \left[\sum_{j=0}^m \beta_j x_i^j + \sum_{k=1}^r \beta_{m+k} (x_i - K_k)_+^m \right] \\ &= \vec{Y} - \mathbf{X}\vec{\beta} \end{aligned} \quad (6)$$

Sehingga jumlah kuadrat *error* diberikan oleh:

$$\sum_{i=1}^n \varepsilon_i^2 = \vec{\varepsilon}^T \vec{\varepsilon} = (\vec{Y} - \mathbf{X}\vec{\beta})^T (\vec{Y} - \mathbf{X}\vec{\beta}) \quad (7)$$

Metode OLS mendapatkan estimator $\hat{\beta}$ melalui penyelesaian optimasi:

$$\min \{Q(\vec{\beta})\} = \min \{(\vec{Y} - \mathbf{X}\vec{\beta})^T (\vec{Y} - \mathbf{X}\vec{\beta})\} \quad (8)$$

Untuk menyelesaikan optimasi pada Persamaan (8) dapat menggunakan derivatif parsial. Bentuk kuadrat dari $Q(\vec{\beta})$ dapat diuraikan menjadi

$$Q(\vec{\beta}) = [\vec{Y}^T \vec{Y} - 2\vec{\beta}^T \mathbf{X}^T \vec{Y} + \vec{\beta}^T \mathbf{X}^T \mathbf{X} \vec{\beta}] \quad (9)$$

Kemudian Persamaan (9) diturunkan terhadap vektor $\vec{\beta}$ dan disamadengankan nol sedemikian sehingga diperoleh:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \vec{Y} \quad (10)$$

Dengan demikian estimasi kurva regresi $\hat{f}_p(x_{pi})$ diperoleh:

$$\hat{f}_p(x_{pi}) = \mathbf{X}\hat{\beta} \rightarrow \mathbf{X}[(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\vec{Y}] \rightarrow \mathbf{Z}(K)\vec{Y} \quad (11)$$

2.3 Menghampiri kurva regresi nonparametrik menggunakan estimator deret Fourier.

Misalkan diberikan data berpasangan x_i dan y_i , dimana $i = 1, 2, \dots, n$ yang merupakan model regresi nonparametrik. Didapatkan persamaan deret Fourier sebagai berikut:

$$y_i = \frac{1}{2}\alpha_0 + \gamma x_i + \sum_{q=1}^Q \alpha_q \cos(qx_i) + \varepsilon_i; i = 1, 2, \dots, n \quad (12)$$

Model regresi pada Persamaan (12) dalam bentuk matriks adalah sebagai berikut:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_{11} & \cos x_{11} & \cos(2x_{11}) & \cdots & \cos(Qx_{11}) \\ 1 & x_{12} & \cos x_{12} & \cos(2x_{12}) & \cdots & \cos(Qx_{12}) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & \cos x_{1n} & \cos(2x_{1n}) & \vdots & \cos(Qx_{1n}) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \gamma \\ \vdots \\ \alpha_Q \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (13)$$

Maka dapat menjadi:

$$\vec{Y} = \mathbf{A}\vec{\theta} + \vec{\varepsilon} \quad (14)$$

Selanjutnya parameter $\vec{\theta}$ dapat diestimasi dengan menggunakan OLS. Adapun *error* pada Persamaan (14) diberikan oleh:

$$\begin{aligned} \varepsilon_i &= y_i - \left[\frac{1}{2}\alpha_0 + \gamma x_i + \sum_{q=1}^Q \alpha_q \cos(qx_i) \right] \\ &= \vec{Y} - \mathbf{A}\vec{\theta} \end{aligned} \quad (15)$$

Sehingga jumlah kuadrat *error* diberikan oleh:

$$\sum_{i=1}^n \varepsilon_i^2 = \vec{\varepsilon}^T \vec{\varepsilon} = (\vec{Y} - \mathbf{A}\vec{\theta})^T (\vec{Y} - \mathbf{A}\vec{\theta}) \quad (16)$$

Metode OLS mendapatkan estimator $\hat{\vec{\theta}}$ melalui penyelesaian optimasi:

$$\min \{Q(\vec{\theta})\} = \min \{(\vec{Y} - \mathbf{A}\vec{\theta})^T (\vec{Y} - \mathbf{A}\vec{\theta})\} \quad (17)$$

Untuk menyelesaikan optimasi pada Persamaan (17) dapat menggunakan derivatif parsial. Bentuk kuadratik dari $Q(\vec{\theta})$ dapat diuraikan menjadi

$$Q(\vec{\theta}) = [\vec{Y}^T \vec{Y} - 2\vec{\theta}^T \mathbf{A}^T \vec{Y} + \vec{\theta}^T \mathbf{A}^T \mathbf{A} \vec{\theta}] \quad (18)$$

Kemudian Persamaan (18) diturunkan terhadap vektor $\vec{\theta}$ dan disamadengankan nol sedemikian sehingga diperoleh:

$$\hat{\theta} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \vec{Y} \quad (19)$$

Dengan demikian estimasi kurva regresi $\hat{f}_p(x_{pi})$ diperoleh:

$$\hat{f}_p(x_{pi}) = \mathbf{A} \hat{\theta} \rightarrow \mathbf{A} [(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \vec{Y}] \rightarrow \mathbf{B}(Q) \vec{Y} \quad (20)$$

2.4 Memodelkan kasus CFR akibat DBD di Indonesia dengan estimator *spline truncated* dan deret Fourier.

Membuat *scatter plot* (diagram pencar) antara variabel respon (y) dengan masing-masing variabel prediktor (x_1, x_2, x_3, x_4). Selanjutnya dimodelkan dengan pendekatan regresi nonparametrik berdasarkan estimator *spline truncated* dan deret Fourier.

2.5 Menetapkan estimator terbaik dalam proses pemodelan kasus CFR akibat DBD di Indonesia.

Tujuan utama dalam pemodelan statistika adalah mendapatkan model yang terbaik. Adapun kriteria yang digunakan:

Kriteria *Generalized Cross Validation* (GCV):

$$GCV(K_{opt}) = \frac{n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n^{-1} \text{trace}[\mathbf{I} - [\mathbf{Z}(K)])]^2} \quad \& \quad GCV(Q_{opt}) = \frac{n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n^{-1} \text{trace}[\mathbf{I} - [\mathbf{B}(Q)])]^2} \quad (21)$$

Kriteria Koefisien Determinasi (R^2):

$$R^2 = 1 - \frac{SSE}{SST} \quad (22)$$

dengan SSE adalah $\sum_{i=1}^n (y_i - \hat{y}_i)^2$ dan SST adalah $\sum_{i=1}^n (y_i - \bar{y})^2$.

3. Hasil dan Pembahasan

Pada bagian ini akan dibahas mengenai estimasi model regresi nonparametrik *spline truncated* dan deret Fourier yang diterapkan pada data CFR akibat DBD pada 33 Provinsi di Indonesia Tahun 2018.

3.1. Statistika Deskriptif

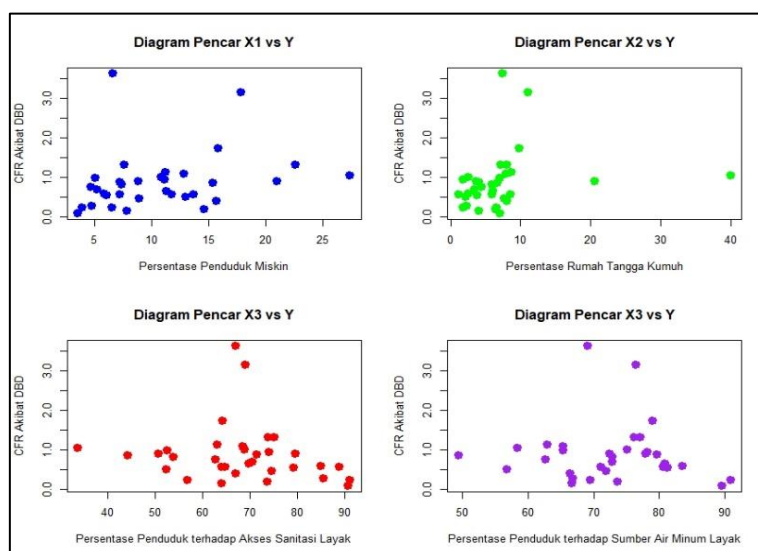
Sebelum dilakukan proses pemodelan dengan menggunakan estimator *spline truncated* dan estimator deret Fourier pada kasus DBD di Indonesia, ditampilkan statistika deskriptif guna memberikan informasi awal dari masing-masing variabel yang digunakan pada Tabel 1.

Statistika deskriptif yang ditampilkan pada Tabel 1 memperlihatkan bahwa masing-masing Provinsi di Indonesia memiliki karakteristik yang berbeda-beda untuk semua variabel prediktornya. Provinsi Maluku Utara merupakan daerah yang memiliki CFR akibat DBD tertinggi yaitu sebesar 3,64%, sedangkan DKI Jakarta merupakan daerah yang memiliki CFR akibat DBD terendah yaitu hanya sebesar 0,07%.

Tabel 1. Statistika deskriptif variabel penelitian

Variabel	N	Minimum	Maksimum	Rata-Rata	Standar Deviasi
Y	33	0,07	3,64	0,88	0,75
x_1	33	3,55	27,43	10,72	5,75
x_2	33	1,13	40,01	7,10	6,95
x_3	33	33,75	91,14	68,25	13,13
x_4	33	49,37	90,90	72,49	8,98

Selanjutnya hubungan variabel respon dengan masing-masing variabel prediktor dapat dilihat dari diagram pencar. Hasil diagram pencar untuk masing-masing variabel respon terhadap variabel prediktor ditampilkan pada Gambar 1.



Gambar 1. Diagram pencar antara variabel prediktor terhadap variabel respon

Diagram pencar pada Gambar 1 menunjukkan pola hubungan antara CFR Akibat DBD (Y) terhadap ke empat variabel prediktor. Diagram pencar yang ditampilkan pada Gambar 1 memperlihatkan tidak adanya kecenderungan membentuk suatu pola tertentu. Sehingga dengan pola data yang demikian, kurang tepat apabila didekati dengan menggunakan regresi parametrik.

Selanjutnya pola hubungan antara antara CFR Akibat DBD (Y) terhadap ke empat variabel prediktor akan di dekati dengan model regresi nonparametrik. Estimator yang digunakan pada penelitian ini ada 2 yaitu *spline truncated* dan deret Fourier. Akan di pilih model yang terbaik dari kedua estimator yang digunakan.

Penelitian ini menggunakan jumlah titik knot dengan jumlah yang sama untuk setiap variabel prediktor yang digunakan yaitu 1 sampai dengan 4 titik knot untuk estimator *spline truncated*, sedangkan untuk estimator deret Fourier akan dicobakan parameter osilasi dari 1 sampai dengan 4.

3.2. Pemodelan Regresi Nonparametrik dengan Estimator *Spline Truncated*

Pada pemodelan regresi nonparametrik dengan estimator *spline truncated*, langkah awal yang dilakukan adalah menentukan banyaknya titik knot serta lokasi dimana titik knot

Pemodelan Regresi Nonparametrik dengan Estimator Spline Truncated ...

yang optimal berada sedemikian sehingga akan menghasilkan nilai GCV yang minimum. Metode estimasi yang digunakan adalah OLS. Adapun hasil yang didapatkan dari proses pemodelan CFR Akibat DBD (Y) ditampilkan pada Tabel 2.

Tabel 2. Nilai GCV dan R² berdasarkan banyaknya dan letak titik knot

Banyaknya Titik Knot	Lokasi Knot Optimal				GCV	Koefisien Determinasi
	X ₁	X ₂	X ₃	X ₄		
1	22,88	32,60	80,20	82,98	0,78	18,27%
2	17,19	23,34	66,54	73,10	0,71	47,60%
	18,33	25,19	69,27	75,08		
3	16,05	21,49	63,81	71,12	0,53	71,05%
	17,19	23,34	66,54	73,10		
	18,33	25,19	69,27	75,08		
4	8,09	8,53	44,68	57,28	0,25	91,80%
	13,78	17,79	58,34	67,17		
	14,92	19,64	61,07	69,14		
	16,05	21,49	63,81	71,12		

Berdasarkan hasil analisis, nilai GCV minimum dan koefisien determinasi terbesar didapatkan saat banyaknya titik knot adalah 4. Berikut disajikan persamaan regresi nonparametrik *spline truncated* dengan 4 titik knot.

$$\hat{y} = 0,01 + 0,14x_1 + 0,00x_2 - 0,94x_3 + 0,84x_4 - 0,15(x_1 - 8,09)_+ - 0,73(x_1 - 13,78)_+ + 1,57(x_1 - 14,92)_+ - 0,76(x_1 - 16,05)_+ + 0,83(x_2 - 8,53)_+ - 4,68(x_2 - 17,79)_+ + 0,30(x_2 - 19,64)_+ + 2,88(x_2 - 21,49)_+ + 0,09(x_3 - 44,68)_+ + 3,17(x_3 - 58,34)_+ - 3,20(x_3 - 61,07)_+ + 0,88(x_3 - 63,81)_+ - 0,78(x_4 - 57,28)_+ + 1,50(x_4 - 67,17)_+ - 3,04(x_4 - 69,14)_+ + 1,47(x_4 - 71,12)_+$$

3.3. Pemodelan Regresi Nonparametrik dengan Estimator Deret Fourier

Pada pemodelan regresi nonparametrik dengan estimator deret Fourier, maka langkah awal yang dilakukan adalah menentukan banyaknya osilasi (gelombang) sedemikian sehingga akan menghasilkan nilai GCV yang minimum. Osilasi akan disimbolkan dengan Q , dimana nilai Q merupakan bilangan bulat positif. Q yang akan dicobakan dari 1 hingga 4. Selanjutnya dilakukan metode estimasi dengan OLS. Adapun hasil yang didapatkan dari proses pemodelan CFR Akibat DBD (Y) ditampilkan pada Tabel 3.

Tabel 3. Nilai GCV dan R² berdasarkan banyaknya osilasi

Nilai Q	GCV	Koefisien Determinasi
$Q = 1$	7,53	19,44%
$Q = 2$	2,97	28,63%
$Q = 3$	0,93	60,35%
$Q = 4$	0,52	65,44%

Berdasarkan hasil analisis, nilai GCV minimum dan koefisien determinasi terbesar didapatkan saat banyaknya osilasi adalah 4. Sebenarnya banyaknya nilai Q yang dicobakan bisa lebih dari 4, akan tetapi akan berdampak terhadap banyaknya parameter yang akan diestimasi. Sebagai contoh dengan $Q = 5$ maka parameter yang diestimasi

ada sebanyak 25. Model yang dihasilkan tidak parsimoni (sederhana). Berikut disajikan persamaan regresi nonparametrik deret Fourier dengan banyaknya osilasi adalah 4.

$$\hat{y} = -1.07 + 0,08x_{1i} + 0,12 \cos x_{1i} - 0,09 \cos 2x_{1i} - 0,12 \cos 3x_{1i} - 0,01 \cos 4x_{1i} - 0,00x_{2i} + 0,01 \cos x_{2i} - 0,36 \cos 2x_{2i} - 0,74 \cos 3x_{2i} - 0,27 \cos 4x_{2i} - 0,01x_{3i} - 0,28 \cos x_{3i} + 0,11 \cos 2x_{3i} + 0,17 \cos 3x_{3i} + 0,14 \cos 4x_{3i} + 0,02x_{4i} + 0,35 \cos x_{4i} + 0,24 \cos 2x_{4i} + 0,49 \cos 3x_{4i} - 0,29 \cos 4x_{4i}$$

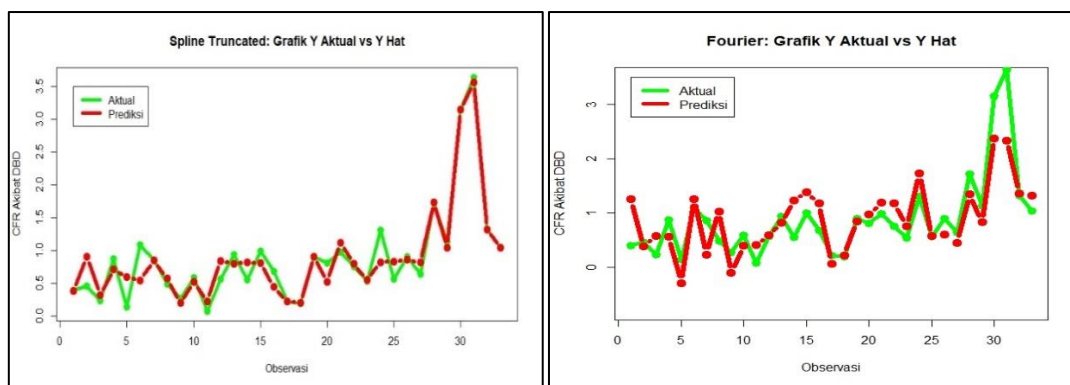
3.4. Pemilihan Model Regresi Nonparametrik Terbaik

Setelah dilakukan pemodelan dengan menggunakan pendekatan regresi nonparametrik dengan 2 estimator yang dicobakan, yaitu *spline truncated* dan deret Fourier maka langkah selanjutnya adalah mendapatkan model yang terbaik.

Tabel 4. Pemilihan model terbaik

Estimator	Koefisien Determinasi	MSE
Spline Truncated	91,80%	0,04
Deret Fourier	65,44%	0,19

Berdasarkan Tabel 4, terlihat estimator *spline truncated* memiliki nilai GCV dan MSE yang lebih kecil serta nilai koefisien determinasi yang lebih tinggi dibandingkan estimator deret Fourier, sehingga estimator terbaik untuk memodelkan kasus CFR akibat DBD di Indonesia adalah *spline truncated*. Pada *spline truncated* banyaknya titik knot yang digunakan adalah 4. Adapun hasil visualisasi dengan menggunakan diagram garis berdasarkan nilai y dengan \hat{y} dari kedua estimator di ditampilkan pada Gambar 2.



Gambar 2. Grafik aktual dan prediksi dari kedua estimator

Terlihat nilai \hat{y} yang didapatkan dari pemodelan regresi nonparametrik menggunakan estimator *spline truncated* cenderung mengikuti pola data aktual.

4. Kesimpulan

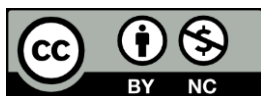
Hasil pemodelan *Case Fatality Rate* (CFR) akibat Demam Berdarah *Dengue* (DBD) di 33 provinsi Indonesia menggunakan pendekatan regresi nonparametrik, menunjukkan bahwa estimator *spline truncated* memiliki performa dan akurasi yang lebih baik dibandingkan estimator deret Fourier. Hal ini ditunjukkan dengan nilai koefisien determinasi dari estimator *spline truncated* adalah 91,80% dan nilai MSE sebesar 0,04.

Pada estimator deret Fourier, diperoleh nilai koefisien determinasi sebesar 65,44% dan nilai MSE sebesar 0,19.

Referensi

- [1] Kemenkes RI, "Buletin Demam Berdarah," 2014. [Online]. Available: <https://www.kemkes.go.id/folder/view/01/structure-publikasi-pusdatin/buletin.html>. [Accessed: 12-09-2020].
- [2] Kemenkes RI, *Profil Kesehatan Indonesia 2015*, Jakarta: Kementerian Kesehatan RI, 2016.
- [3] Kemenkes RI, *Profil Kesehatan Indonesia 2017*, Jakarta: Kementerian Kesehatan RI, 2018.
- [4] S. Sifriyani, "Simultaneous Hypothesis Testing of Multivariable Nonparametric Spline Regression in the GWR Model," *International Journal of Statistics and Probability.*, vol. 8, no. 4, pp. 32-46, 2019.
- [5] A. Fitriani, I. G. A. M. Srinadi, and M. Susilawati, "Estimasi Model Regresi Semiparametrik Menggunakan Estimator Kernel Uniform (Studi Kasus: Pasien DBD di RS Puri Raharja)," *E-Jurnal Matematika*, vol. 4, no. 04, pp. 176-180, 2015.
- [6] E. Yunita, K. Kuzairi, and R. Mubarak, "Model Proyeksi Kematian Penderita Demam Berdarah Dengue (DBD) Berdasarkan Estimator Spline," *Prosiding National Conference on Mathematics, Science, and Education (NACOSME)*, pp. 195-204, 2018.
- [7] N. W. Yani, I. G. A. M. Srinadi, and I. W. Sumarjaya, "Aplikasi Model Regresi Semiparametrik Spline Truncated (Studi Kasus: Pasien Demam Berdarah Dengue (DBD) di Rumah Sakit Puri Raharja)," *E-Jurnal Matematika*, vol. 6, pp. 65-73, 2017.
- [8] I. N. Budiantara, "Spline Dalam Regresi Nonparametrik Dan Semiparametrik: Sebuah Pemodelan Statistika Masa Kini dan Masa Mendatang," *Pidato Pengukuhan Untuk Jabatan Guru Besar Dalam Bidang Ilmu: Matematika Statistika dan Probabilitas, Pada Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Insititut Teknologi Sepuluh November, Surabaya*, 2009.
- [9] N. A. K. Rifai, "Pendekatan Regresi Nonparametrik dengan Fungsi Kernel untuk Indeks Harga Saham Gabungan," *E-Jurnal Statistika*, vol. 19, no. 01, pp. 53-61, 2019.
- [10] R. L. Eubank, *Spline Smoothing and Nonparametric Regression*, New York: Marcel Dekker, 1988.
- [11] W. Hardle, *Applied Nonparametric Regression*, New York: Cambridge University Press, 1990.
- [12] A. P. Putri, R. Santoso, and Sugito, "Analisis Regresi Nonparametrik Kernel Menggunakan Metode Jacknife Sampel Terhapus-1 dan Sample Terhapus-2 (Studi Kasus Pemodelan Tingkat Inflasi Terhadap Nilai Tukar Rupiah di Indonesia Periode 2004-2016)," *Jurnal Gaussian*, vol. 6, no. 1, pp. 01-10, 2017.
- [13] V. Ratnasari, I. N. Budiantara, M. Ratna, and I. Zain, "Estimation of Nonparametric Regression Curve using Mixed Estimator of Multivariable

- Truncated Spline and Multivariable Kernel," *Global Journal of Pure and Applied Mathematics*. vol. 10, pp. 5047-5057, 2016.
- [14] G. Wahba, *Spline Models for Observational Data*, Pennsylvania: SIAM, 1990.
- [15] R. L. Eubank, *Nonparametric Regression and Spline Smoothing, 2 Edition*. New York: Marcel Dekker, 1999.
- [16] D. R. S. Saputro, A. Sukmayanti, and P. Widyaningsih, "The Nonparametric Regression Model Using Fourier Series Approximation and Penalized Least Squares (Case on Data Poverty in East Java)," in *The Sixth Seminar Nasional Pendidikan Matematika Universitas Ahmad Dahlan, IOP Conference Series: Journal Physics* 1188, 2019.
- [17] A. Prahutama, S. Suparti, and T. W. Utami, "Modelling Fourier Regression for Time Series Data-a Case Study: Modelling Inflation in Foods Sector in Indonesia," in *International Conference on Mathematics: Pure, Applied and Computation, IOP Conference Series: Journal Physics* 974, 2018.
- [18] M. Bilodeau, "Fourier Smoother and Additive Models", *The Canadian Journal of Statistics*, vol. 03, pp. 257-269, 1992.
- [19] L. J. Asrini and I. N. Budiantara, "Fourier Series Semiparametric Regression Models (Case Study: The Production of Low Land Rice Irrigation in Central Java)," *ARNP Journal of Engineering and Applied Sciences*, vol. 09, pp. 1501-1506, 2014.
- [20] R. Pane, I. N. Budiantara, I. Zain, and B. W. Otok, "Parametric and Nonparametric Estimators in Fourier Series Semiparametric Regression and Their Characteristics," *Applied Mathematical Sciences*, vol. 102, no. 08, pp. 5053-5064, 2012.
- [21] L. J. Asrini "Regresi Semiparametrik Deret Fourier", in *Prosiding Seminar Nasional FMIPA Universitas Negeri Surabaya*, pp. 77-80, 2012.
- [22] BPS, *Statistik Lingkungan Hidup*, Jakarta: Badan Pusat Statistik, 2016.



This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution-NonCommercial 4.0 International License](https://creativecommons.org/licenses/by-nc/4.0/). Editorial of JJoM: Department of Mathematics, Universitas Negeri Gorontalo, Jln. Prof. Dr. Ing. B.J. Habibie, Moutong, Tilongkabila, Kabupaten Bone Bolango, Provinsi Gorontalo 96119, Indonesia.