

PENERAPAN MODEL REGRESI SEMIPARAMETRIK DERET FOURIER UNTUK MENGIDENTIFIKASI FAKTOR PENENTU ANGKA HARAPAN HIDUP DALAM KONTEKS SDGS 3

Jesselin Paskalis Sitinjak¹, Sifriyani^{2*}, Andrea Tri Rian Dani³

^{1,2,3} Program Studi Statistika, Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Mulawarman

*e-mail: sifriyani@fmipa.unmul.ac.id

DOI: 10.14710/j.gauss.14.2.302-313

Article Info:

Received: 2025-03-05

Accepted: 2025-09-12

Available Online: 2025-09-17

Keywords:

Life Expectancy; Fourier Series; GML; Semiparametric Regression; RMSEP.

Abstract: *Semiparametric regression combines parametric and nonparametric regression, applied when some relationships between the response and predictor variables are known while others are unknown. This study employed linear regression for the parametric component and the Fourier series estimator for the nonparametric component. The data consisted of life expectancy in Indonesia and its influencing factors, relevant to Sustainable Development Goals (SDGs) 3. The relationship between life expectancy and some predictors (maternal mortality rate, poverty rate) was linear, whereas with others (open unemployment rate, average years of schooling, Gini ratio, stunting prevalence, exclusive breastfeeding) the pattern was unknown and tended to be periodic. This characteristic aligns with the strength of the Fourier series semiparametric regression, which combines linear modeling with Fourier series components for nonlinear periodic relationships. The objective was to determine the optimal number of oscillations using the Generalized Maximum Likelihood (GML) method, obtain the best model, and identify factors affecting life expectancy. Results showed the best model had three oscillations with a GML of $3,98 \times 10^{15}$, RMSEP of 0.99, and R^2 of 82.67%. Significant factors included poverty rate, open unemployment rate, average years of schooling, Gini ratio, stunting prevalence, and exclusive breastfeeding.*

1. PENDAHULUAN

Analisis regresi merupakan analisis statistik yang digunakan untuk mengidentifikasi bentuk hubungan antara variabel respon dengan variabel prediktor. (Pasarella, *et al.*, 2022), (Sifriyani, *et al.*, 2023). Ada tiga pendekatan yang dapat diterapkan untuk memperkirakan bentuk kurva regresi, yaitu pendekatan parametrik, nonparametrik, dan semiparametrik (Purnaraga, *et al.*, 2020), (Sifriyani, *et al.*, 2023). Pendekatan regresi semiparametrik menggabungkan elemen antara regresi parametrik dan regresi nonparametrik (Dani dan Adrianingsih, 2021).

Dalam beberapa kasus di kehidupan nyata yang melibatkan lebih dari satu variabel prediktor, sering kali ditemukan kondisi di mana sebagian pola hubungan antara variabel respon dan prediktor sudah diketahui, sementara sebagian lainnya belum diketahui. Oleh karena itu, regresi semiparametrik dapat menjadi alternatif yang dapat diterapkan (Dani *et al.*, 2023). Metode semiparametrik digunakan ketika bentuk kurva regresi sebagian diketahui dan sebagian polanya tidak diketahui (Sifriyani, *et al.*, 2023).

Salah satu metode estimasi yang sering digunakan dalam pendekatan nonparametrik dan semiparametrik adalah deret Fourier (Khairunnisa, *et al.*, 2020). Regresi semiparametrik deret Fourier dapat diaplikasikan pada berbagai bidang, diantaranya pendidikan, ekonomi, dan kesehatan. Pada penelitian ini mengambil kasus yang berkaitan dengan bidang kesehatan, yaitu pada data Angka Harapan Hidup di Indonesia. Peningkatan sektor

kesehatan menjadi salah satu aspek krusial dalam mencapai tujuan *Sustainable Development Goals* (SDGs).

SDGs 3 yang secara khusus bertujuan untuk menjamin kesehatan yang baik serta meningkatkan kesejahteraan bagi semua individu di setiap kelompok usia. Angka Harapan Hidup (AHH) merupakan salah satu indikator yang dapat merepresentasikan permasalahan serta isu strategis di bidang kesehatan dalam suatu wilayah. Menurut data dari Badan Pusat Statistik (BPS), AHH di Indonesia tahun 2023 sebesar 72,13 tahun. Angka tersebut meningkat 0,28 tahun dibandingkan pada tahun 2022 yaitu sebesar 71,85 tahun (Wulandari dan Nurhayati, 2024). Di Indonesia AHH berkembang secara cepat, ini menunjukkan bahwa masyarakat Indonesia memiliki tingkat kesuburan yang sangat tinggi (Pratiwi dan Mulyawan, 2023). Peningkatan AHH menunjukkan hal yang positif terhadap peningkatan kesejahteraan masyarakat. Berdasarkan studi awal penelitian, peneliti membuat *scatterplot* beberapa variabel yang diduga memengaruhi AHH, dilihat dari *scatterplot* terdapat variabel yang cenderung membentuk pola linier, yaitu AKI dan Persentase Penduduk Miskin dan terdapat variabel yang tidak memiliki pola tertentu, acak dan cenderung berulang, yaitu TPT, RLS, GR, Prevalensi Stunting, serta Persentase ASI Eksklusif. Sehingga pada penelitian ini peneliti menggunakan metode regresi semiparametrik dengan pendekatan deret Fourier. Penelitian ini bertujuan untuk memperoleh model terbaik dari regresi semiparametrik deret Fourier dengan pemilihan osilasi menggunakan GML, dan memperoleh faktor-faktor yang berpengaruh terhadap AHH sehingga dapat memberikan informasi kepada masyarakat bahkan pemerintah untuk mempertahankan bahkan meningkatkan kesejahteraan masyarakat.

2. TINJAUAN PUSTAKA

Regresi parametrik yaitu bentuk regresi yang memiliki pola hubungan tertentu yang telah diketahui antara variabel respon dan prediktor, misalnya pola linier, kuadratik, kubik, dan atau lainnya (Padatuan, *et al.*, 2021). Umumnya, regresi parametrik dengan estimator linier dapat dituliskan pada Persamaan (1).

$$y_i = \beta_0 + \sum_{k=1}^l \beta_k x_{ki} + \varepsilon_i ; i = 1, 2, \dots, n; k = 1, 2, \dots, l \quad (1)$$

Model pada Persamaan (1) dapat disajikan dalam representasi matriks sebagaimana ditunjukkan dalam Persamaan (2).

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon} \quad (2)$$

di mana,

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{i1} \\ 1 & x_{12} & x_{22} & \cdots & x_{i2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{in} \end{bmatrix} \quad \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_l \end{bmatrix} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (3)$$

Parameter $\boldsymbol{\beta}$ dapat ditaksir dengan metode *Maximum Likelihood Estimation* (MLE) dengan *error* yang diasumsikan identik, independen, dan berdistribusi normal dengan mean bernilai nol dan variansi σ^2 (Eubank, 1999).

Sehingga didapatkan estimasi parameter $\hat{\boldsymbol{\beta}}$ pada Persamaan (4).

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (4)$$

Regresi nonparametrik merupakan suatu metode yang digunakan untuk mengetahui pengaruh hubungan variabel prediktor terhadap variabel respon dimana pola data tidak diketahui atau tidak membentuk suatu pola tertentu (Noviani, *et al.*, 2020). Apabila pola data yang dianalisis belum diketahui dan cenderung berulang atau bersifat periodik, estimator regresi nonparametrik deret Fourier merupakan metode yang tepat digunakan.

Pada deret Fourier, bentuk kurva regresi nonparametrik dianggap tak diketahui dan terletak dalam himpunan fungsi kontinu $C(0, \pi)$ (Amalia dan Nur, 2019). Model regresi deret Fourier multivariabel ditunjukkan pada Persamaan (5).

$$y_i = \frac{1}{2}\alpha_0 + \sum_{j=1}^s \left(\gamma_j(z_{ji}) + \sum_{q=1}^Q \alpha_{qj} \cos(qz_{ji}) \right) + \varepsilon_i; j = 1, 2, \dots, s; i = 1, 2, \dots, n; q = 1, 2, \dots, Q \quad (5)$$

dengan

- y_i : variabel respon pada pengamatan ke- i
- z_{ji} : variabel prediktor ke- j pada pengamatan ke- i
- α_0 : konstanta parameter deret Fourier
- γ_j : koefisien parameter deret Fourier
- α_{qj} : koefisien parameter osilasi ke- q pada variabel prediktor ke- j
- ε_i : *error* yang diasumsikan identik, independen, berdistribusi normal dengan rata-rata bernilai nol dan variansi σ^2

Persamaan (5) dapat disajikan dalam representasi vektor dan matriks yang tuliskan pada Persamaan (6).

$$\mathbf{y} = \mathbf{D}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \quad (6)$$

Sehingga dapat diuraikan sebagai berikut

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & z_{11} & \cos z_{11} & \cdots & \cos(Qz_{11}) & \cdots & z_{s1} & \cos z_{s1} & \cdots & \cos(Qz_{11}) \\ \frac{1}{2} & z_{12} & \cos z_{12} & \cdots & \cos(Qz_{12}) & \cdots & z_{s2} & \cos z_{s2} & \cdots & \cos(Qz_{11}) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{1}{2} & z_{1n} & \cos z_{1n} & \cdots & \cos(Qz_{1n}) & \cdots & z_{sn} & \cos z_{sn} & \cdots & \cos(Qz_{11}) \end{bmatrix} \begin{bmatrix} \alpha_0 \\ \gamma_1 \\ \alpha_{11} \\ \vdots \\ \alpha_{Q1} \\ \vdots \\ \gamma_s \\ \alpha_{1s} \\ \vdots \\ \alpha_{Qs} \end{bmatrix} + \begin{bmatrix} \varepsilon_0 \\ \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{bmatrix} \quad (7)$$

Parameter $\boldsymbol{\delta}$ dapat diestimasi menggunakan metode MLE dengan *error*-nya diasumsikan $\varepsilon_i \sim IIDN(0, \sigma^2)$, maka dapat dituliskan fungsi *ln likelihood* seperti pada Persamaan (8)

$$\begin{aligned} \ell(\boldsymbol{\delta}) &= L(\boldsymbol{\delta}) \\ &= \ln \left((2\pi\sigma^2)^{-\frac{n}{2}} \exp \left(-\frac{1}{2} \pi\sigma^2 (\mathbf{y} - \mathbf{D}\boldsymbol{\delta})^T (\mathbf{y} - \mathbf{D}\boldsymbol{\delta}) \right) \right) \end{aligned} \quad (8)$$

sehingga didapatkan estimasi parameter $\boldsymbol{\delta}$ pada persamaan (9).

$$\hat{\boldsymbol{\delta}} = (\mathbf{D}^T \mathbf{D})^{-1} \mathbf{D}^T \mathbf{y} \quad (9)$$

Regresi semiparametrik merupakan perpaduan dari regresi parametrik dan nonparametrik. Dalam analisis regresi, jika terdapat komponen parametrik serta nonparametrik dengan data berpasangan (x_i, y_i, z_i) yang diasumsikan sesuai dengan model regresi semiparametrik, maka dijelaskan dalam Persamaan (10).

$$y_i = f(x_i) + g(z_i) + \varepsilon_i \quad (10)$$

Misalkan terdapat data berpasangan $(x_{1i}, x_{2i}, \dots, x_{li}, z_{1i}, z_{2i}, \dots, z_{si}, y_i)$ yang pola hubungannya diasumsikan mengikut model regresi semiparametrik. Kurva regresi semiparametrik deret Fourier dituliskan pada Persamaan (11).

$$\mu(x_i, z_i) = \sum_{k=1}^l f(x_{ki}) + \sum_{j=1}^s g(z_{ji}) \quad (11)$$

Bentuk penjabaran model regresi semiparametrik deret Fourier dalam matriks dituliskan pada Persamaan (12).

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \quad (12)$$

Komponen $\mathbf{X}\boldsymbol{\beta}$ merupakan bagian parametrik, sedangkan komponen $\mathbf{D}\boldsymbol{\delta}$ merupakan bagian nonparametrik (Fadlirhohim, *et al.*, 2024).

Estimasi parameter pada regresi semiparametrik deret Fourier dapat menggunakan metode MLE. Berdasarkan model regresi semiparametrik deret Fourier didapatkan model linier:

$$\begin{aligned} \mathbf{y} &= \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \\ &= (\mathbf{X}:\mathbf{D}) + \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix} + \boldsymbol{\varepsilon} \\ &= \mathbf{F}\boldsymbol{\tau} + \boldsymbol{\varepsilon} \end{aligned} \quad (13)$$

dengan $\mathbf{F} = (\mathbf{X}:\mathbf{D})$; $\boldsymbol{\tau} = \begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\delta} \end{pmatrix}$; $\boldsymbol{\varepsilon} = \mathbf{y} - \mathbf{F}\boldsymbol{\tau}$

Sehingga diperoleh estimasi untuk $\hat{\boldsymbol{\tau}}$ yang dituliskan pada Persamaan (14).

$$\hat{\boldsymbol{\tau}} = (\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T \mathbf{y} \quad (14)$$

Estimasi dari $\hat{\mathbf{y}}$ pada regresi semiparametrik deret Fourier dituliskan pada Persamaan (15).

$$\begin{aligned} \hat{\mathbf{y}} &= \mathbf{F}\hat{\boldsymbol{\tau}} \\ &= \mathbf{A}(\mathbf{F})\mathbf{y} \end{aligned} \quad (15)$$

dengan $\mathbf{A}(\mathbf{F}) = \mathbf{F}(\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T$ adalah matriks yang digunakan untuk perhitungan pada formula *Generalized Maximum Likelihood* (GML).

Metode GML dikembangkan sebagai pendekatan estimasi berbasis *likelihood* untuk parameter penghalus. Pemilihan parameter penghalus adalah sama dengan memilih jumlah osilasi yang optimal. Metode ini juga dikenal cukup stabil dan efisien, terutama dalam analisis dengan jumlah sampel kecil hingga menengah (Devi, *et al.*, 2021). Rumus dari metode GML dapat dituliskan pada Persamaan (16).

$$GML = \frac{\mathbf{y}^T (\mathbf{I} - \mathbf{A}(\mathbf{F})) \mathbf{y}}{[\det^+(\mathbf{I} - \mathbf{A}(\mathbf{F}))]^{1/(n - ((l+1) + s(1+Q)))}} \quad (16)$$

nilai $\det^+$ merupakan hasil kali nilai-nilai eigen yang bukan nol.

Model regresi terbaik yang dilihat melalui GML akan dilakukan pengujian signifikansi parameter. Pengujian parameter dalam model regresi bertujuan untuk mencari tahu apakah variabel prediktor mempengaruhi variabel respon, baik secara simultan maupun secara parsial (Sifriyani, *et al.*, 2023). Pengujian secara serentak atau simultan dilakukan untuk mencari tahu pengaruh variabel prediktor secara serentak terhadap variabel respon (Rahmania, *et al.*, 2024). Pengujian simultan dilakukan menggunakan uji F. Hipotesis, statistik uji, daerah penolakan, adalah sebagai berikut.

$$H_0: \beta_1 = \beta_2 = \dots = \beta_l = \gamma_1 = \gamma_2 = \dots = \gamma_s = \alpha_{11} = \dots = \alpha_{qs} = 0$$

$$H_1: \text{minimal ada satu } (\beta_k, \gamma_j) \neq 0 \text{ atau } \alpha_{qj} \neq 0 \text{ di mana } k = 1, 2, \dots, l; j = 1, 2, \dots, s; q = 1, 2, \dots, Q$$

Uji statistik yang digunakan pada pengujian serentak di tunjukkan oleh Persamaan (17).

$$F = \frac{MSR}{MSE} \quad (17)$$

di mana,

$$MSR = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{l + s(1 + Q)} \quad (18)$$

$$MSE = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - (l + s(1 + Q)) - 1} \quad (19)$$

H_0 ditolak apabila $F > F_{\alpha; l+s(1+Q); n-(l+s(1+Q))-1}$ atau $pvalue < \alpha$. Hal ini berarti setidaknya ada satu parameter dalam model yang signifikan atau paling tidak ada satu variabel prediktor yang memengaruhi variabel respon (Sitohang, *et al.*, 2024).

Pengujian parameter secara individu atau parsial dilakukan jika dalam pengujian parameter secara serentak terdapat paling tidak satu parameter dalam model yang berpengaruh. Berikut adalah hipotesis yang digunakan dalam pengujian parameter secara parsial.

$$H_0: \beta_k = 0; \gamma_j = 0; \alpha_{qj} = 0; k = 1, 2, \dots, l; j = 1, 2, \dots, s; q = 1, 2, \dots, Q$$

$$H_1: \text{minimal terdapat satu } \beta_k \neq 0; \gamma_j \neq 0; \alpha_{qj} \neq 0$$

Parameter model regresi nonparametrik deret Fourier dimisalkan $\hat{\delta}_w = \{\gamma_j, \alpha_{qj}\}$

Pengujian parameter parsial menggunakan statistik uji ditunjukkan pada Persamaan (20) dan (21).

$$t = \frac{\hat{\beta}_k}{SE(\hat{\beta}_k)} \quad (20)$$

$$t = \frac{\hat{\delta}_w}{SE(\hat{\delta}_w)} \quad (21)$$

H_0 ditolak jika $|t| > t_{\frac{\alpha}{2}; n-(l+s(1+Q))-1}$ atau $pvalue < \alpha$ (Fadlirhohim, *et al.*, 2024).

Kebaikan model dapat diukur melalui nilai koefisien determinasi (R^2). Semakin tinggi nilai R^2 maka semakin baik variabel prediktor dalam model menjelaskan variabilitas variabel respon. Nilai koefisien determinasi dapat dihitung menggunakan Persamaan (22).

$$\begin{aligned} R^2 &= 1 - \frac{SSR}{SST} \\ &= 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \end{aligned} \quad (22)$$

dengan SSR merupakan *Sum of Squared Regression* dan SST merupakan *Sum of Squared Total*.

Root Mean Square Error of prediction (RMSEP) merupakan salah satu ukuran yang dapat digunakan sebagai ukuran kebaikan hasil prediksi seperti dinyatakan dalam Persamaan (23).

$$RMSEP = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n}} \quad (23)$$

dengan n adalah banyak sampel, semakin kecil RMSEP, semakin baik prediksi model yang dihasilkan

(Rosanti dan Budiantara, 2020).

3. METODE PENELITIAN

Data dalam penelitian ini adalah data sekunder yang diperoleh dari laman resmi Badan Pusat Statistik (BPS) serta Studi Status Gizi Indonesia (SSGI). Berikut adalah rincian dari variabel yang digunakan dalam penelitian ini:

- y : Angka harapan hidup (tahun)
- x_1 : Angka kematian ibu (jiwa)
- x_2 : Persentase penduduk miskin (persen)
- x_3 : Tingkat pengangguran terbuka (persen)
- x_4 : Rata-rata lama sekolah (tahun)
- x_5 : *Gini ratio* (indeks)

x_6 : Prevalensi *stunting* (perse)

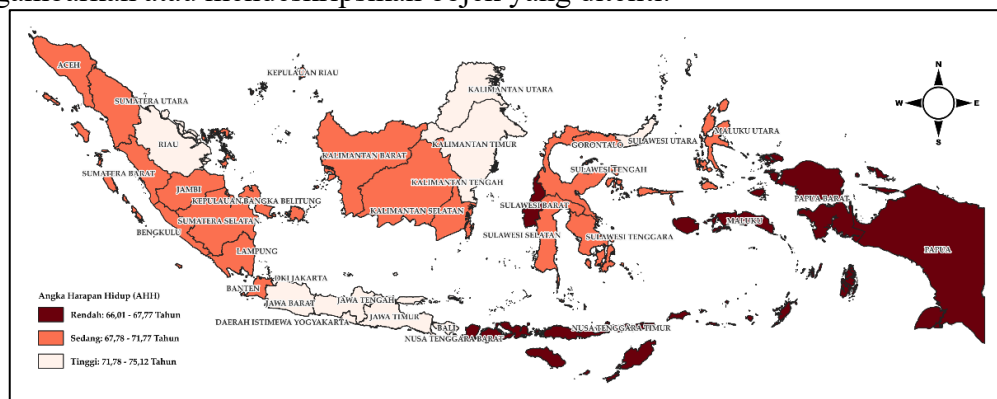
x_7 : Persentase bayi usia < 6 bulan yang mendapatkan ASI eksklusif (persen)

Analisis data dilakukan dengan bantuan *software* R Studio. Tahapan analisis dalam penelitian dituliskan dengan rincian berikut:

1. Mengumpulkan data sekunder mengenai Angka Harapan Hidup di Indonesia Tahun 2023 beserta faktor-faktor yang memengaruhinya.
2. Melakukan analisis statistika deskriptif dengan *spatial mapping* untuk variabel respon dan menghitung nilai rata-rata, nilai minimum, nilai maksimum untuk seluruh variabel.
3. Melakukan pengecekan multikolinieritas
4. Menganalisis pola yang terbentuk antara variabel respon dengan setiap variabel prediktor menggunakan *scatter plot*.
5. Menetapkan variabel prediktor yang dimodelkan dengan estimator parametrik dan nonparametrik.
6. Melakukan pemodelan pada data menggunakan regresi semiparametrik deret Fourier dengan 1 osilasi, 2 osilasi, dan 3 osilasi pada Persamaan (24)
7. Menentukan jumlah osilasi optimal menggunakan metode GML pada Persamaan (16) dan dengan mempertimbangkan nilai R^2 pada Persamaan (22) dan RMSE pada Persamaan (22)
8. Melakukan uji signifikansi parameter, secara simultan menggunakan Persamaan (17) dan parsial pada Persamaan (20) dan (21).
9. Membuat kesimpulan.

4. HASIL DAN PEMBAHASAN

Statistika deskriptif merupakan serangkaian teknik analisis yang digunakan untuk mengumpulkan dan menyajikan data yang berguna memberikan informasi bermanfaat serta menggambarkan atau mendeskripsikan objek yang diteliti.



Gambar 1. Peta Persebaran AHH di Indonesia Tahun 2023

Merujuk pada Gambar 1 dapat dilihat terdapat 6 provinsi dengan kategori AHH rendah dengan interval 66,01-67,77 tahun. Kemudian pada kelompok kedua terdapat 18 provinsi dengan kategori AHH sedang dengan interval 67,78-71,77 tahun. Pada kelompok ketiga terdapat 10 provinsi dengan kategori AHH tinggi dengan interval 71,78-75,12 tahun. Kemudian penelitian ini menggunakan statistika deskriptif seperti yang ditunjukkan oleh Tabel 1.

Tabel 1. Statistika Deskriptif Variabel Respon dan Prediktor

Variabel	Rata-rata	Minimum	Maksimum
y	70,69	66,01	75,12
x_1	210,50	48,00	565,00

x_2	10,09	4,25	26,03
z_1	4,61	2,27	7,52
z_2	9,30	7,34	11,42
z_3	0,34	0,25	0,45
z_4	16,08	5,50	26,20
z_5	70,41	55,11	82,45

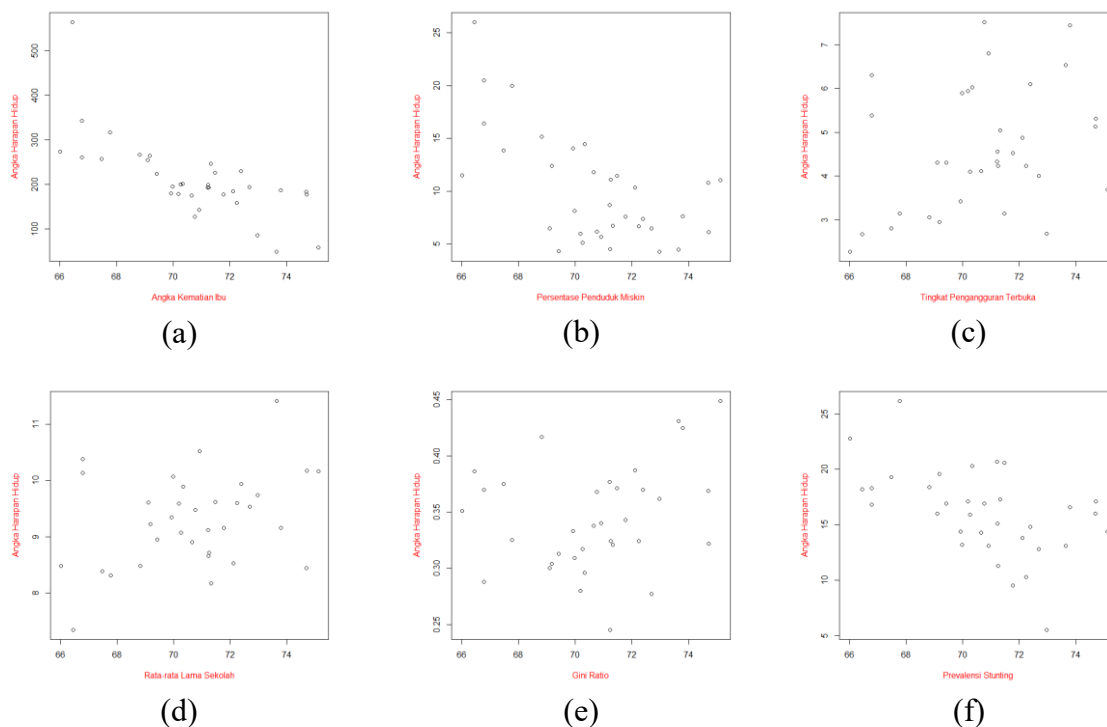
Setelah melakukan analisis statistika deskriptif dilanjutkan dengan pengecekan multikolinieritas. Multikolinieritas adalah kondisi di mana terjadi korelasi antara variabel prediktor atau variabel prediktor tidak bersifat saling bebas. Pengecekan multikolinieritas dilakukan dengan *Variance Inflation Factor* (VIF).

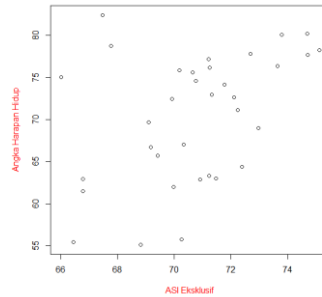
Tabel 2. Pengecekan Multikolinieritas Menggunakan VIF

Variabel Prediktor	VIF
Angka Kematian Ibu	5,03
Persentase Penduduk Miskin	3,12
Tingkat Pengangguran Terbuka	1,50
Rata-rata Lama Sekolah	2,31
<i>Gini Ratio</i>	1,24
Prevalensi <i>Stunting</i>	1,51
Persentase Bayi Usia < 6 Bulan yang Mendapatkan ASI Eksklusif	1,58

Berdasarkan Tabel 2, seluruh nilai VIF berada dibawah nilai 10 yang dimana berarti tidak terdapat multikolinieritas antar variabel prediktor.

Langkah berikutnya dalam memodelkan faktor-faktor yang memengaruhi AHH adalah dengan memeriksa bentuk atau pola hubungan antara AHH dengan variabel prediktor menggunakan *scatter plot* yang visualisasinya ditampilkan pada Gambar 2.





(g)

Gambar 2. *Scatter plot* antara AHH dengan (a) Angka Kematian Ibu, (b) Persentase Penduduk Miskin, (c) Tingkat Pengangguran terbuka, (d) Rata-rata Lama Sekolah, (e) *Gini Ratio*, (f) Prevalensi *Stunting*, (g) Persentase ASI Eksklusif

Gambar 2 menunjukkan bahwa pola data antara AHH dengan (a) dan (b) cenderung membentuk pola linier, sedangkan AHH dengan (c), (d), (e), (f), dan (g) tidak diketahui polanya dan cenderung berulang atau periodik maka data akan dimodelkan menggunakan dengan metode regresi semiparametrik deret Fourier. Hasil identifikasi variabel prediktor ditampilkan pada Tabel 3.

Tabel 3. Identifikasi Variabel Prediktor

Variabel	Estimator	Deskripsi
x_1	Regresi Linier	Angka Kematian Ibu
x_2		Persentase Penduduk Miskin
z_1		Tingkat Pengangguran Terbuka
z_2		Rata-rata Lama Sekolah
z_3	Deret Fourier	<i>Gini Ratio</i>
z_4		Prevalensi <i>Stunting</i>
z_5		Persentase Bayi Usia < 6 Bulan yang Mendapatkan ASI Eksklusif

Model matematis dari regresi semiparametrik deret Fourier pada Angka Harapan Hidup di Indonesia tahun 2023 dituliskan pada Persamaan (24).

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \sum_{j=1}^5 \left(\gamma_j(z_{ji}) + \sum_{q=1}^3 \alpha_{qj} \cos(qz_{ji}) \right) + \varepsilon_i; i = 1, 2, \dots, 34 \quad (24)$$

Model regresi semiparametrik deret Fourier sangat bergantung pada parameter osilasi (q). Pemilihan osilasi optimal dilakukan dengan menghitung nilai GML yang minimum. Namun, pada penelitian ini pemilihan osilasi optimal juga mempertimbangkan nilai RMSEP dan koefisien determinasi. Hasil analisis untuk nilai osilasi 1 sampai 3 ditunjukkan Tabel 4.

Tabel 4. Nilai GML, RMSEP, dan Koefisien Determinasi

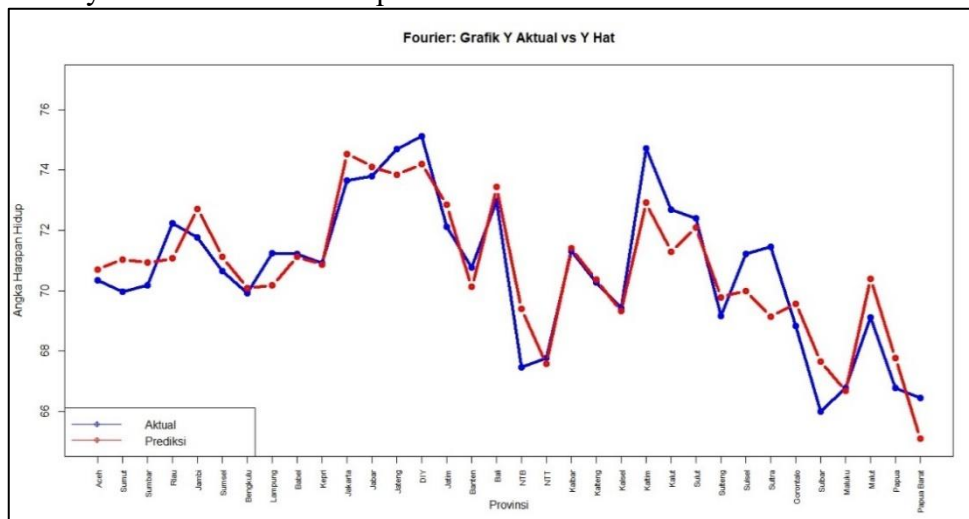
Osilasi (q)	GML	RMSEP	R^2
1	$1,03 \times 10^{15}$	1,23	73,11%
2	$1,48 \times 10^{14}$	1,08	79,41%
3	$3,98 \times 10^{15}$	0,99	82,67%

Melalui Tabel 4 diketahui bahwa nilai GML minimum didapatkan pada jumlah osilasi 2. Namun, dengan mempertimbangkan Nilai RMSEP dan R^2 diputuskan bahwa osilasi 3 menjadi jumlah osilasi yang optimal. Hal ini juga mempertimbangkan hasil pengujian signifikansi parameter baik secara simultan maupun parsial.

Berdasarkan pemilihan osilasi yang optimal yang telah dilakukan, diperoleh model terbaik dengan model yang dituliskan pada Persamaan (22).

$$\hat{y}_i = 2,30 \times 10^2 + 3,71 \times 10^{-4}x_{1i} - 2,75 \times 10^{-1}x_{2i} + 4,64 \times 10^{-2}z_{1i} - 1,37 \times 10^2 \cos z_{1i} + 9,19 \times 10^1 \cos 2z_{1i} - 1,11 \times 10^1 \cos 3z_{1i} + 5,26 \times 10^{-1}z_{2i} + 5,40 \times 10^1 \cos z_{2i} + 1,23 \times 10^2 \cos 2z_{2i} + 1,55 \times 10^2 \cos 3z_{2i} - 5,58 \times 10^2 z_{3i} - 2,17 \times 10^3 \cos z_{3i} - 1,82 \times 10^4 \cos 2z_{3i} + 5,31 \times 10^3 \cos 3z_{3i} - 7,69 \times 10^{-2}z_{4i} + 2,31 \times 10^2 \cos z_{4i} + 4,82 \cos 2z_{4i} - 9,71 \times 10^1 \cos 3z_{4i} + 1,31 \times 10^{-1}z_{5i} + 1,37 \times 10^2 \cos z_{5i} + 1,14 \times 10^2 \cos 2z_{5i} - 6,73 \times 10^1 \cos 3z_{5i} \quad (22)$$

Setelah didapatkan model terbaik regresi semiparametrik deret Fourier maka dapat diperoleh nilai prediksi $\hat{y}$ dan divisualisasikan pada Gambar 3.



Gambar 3. Grafik Perbandingan Data Aktual dengan Data Prediksi

Melalui Gambar 3 terlihat data prediksi umumnya cenderung mengikuti pola data aktual di sebagian besar provinsi, hal ini menandakan bahwa model regresi semiparametrik deret Fourier cukup baik dalam memprediksi tren AHH.

Langkah berikutnya adalah menguji signifikansi parameter secara simultan untuk mengetahui pengaruh seluruh variabel prediktor secara simultan terhadap AHH.

Tabel 5. Analisis Variansi

Sumber Variasi	Derajat Bebas	Jumlah Kuadrat	Rataan Kuadrat	F hitung	$F_{0,05;22;11}$
Regresi	22	157,81	7,17	7,37	2,63
Error	11	33,07	0,97		
Total	33	190,88			

Berdasarkan Tabel 5 diperoleh nilai $F_{hitung} = 7,27$ dan $F_{0,05;22;11}$ sebesar 2,63, maka dikarenakan $F_{hitung} > F_{0,05;22;11}$ diperoleh keputusan bahwa H_0 ditolak, sehingga diperoleh kesimpulan bahwa secara serentak variabel prediktor memiliki pengaruh terhadap variabel AHH. Setelah melakukan uji secara simultan dilanjutkan pengujian parsial menggunakan uji statistik dengan derajat bebas $t_{\frac{0,05}{2};11}$. Hasil pengujian signifikansi model parsial disajikan pada Tabel 6.

Tabel 6. Pengujian Parameter secara Parsial

Variabel	Parameter	Estimasi	t_{hitung}	Keterangan
-	$\hat{\beta}_0$	$2,30 \times 10^2$	2,49*	Berpengaruh
x_1	$\hat{\beta}_1$	$3,71 \times 10^{-4}$	0,06	Tidak Berpengaruh
x_2	$\hat{\beta}_2$	$-2,75 \times 10^{-1}$	-3,06*	Berpengaruh
	$\hat{\gamma}_1$	$4,64 \times 10^{-2}$	0,25	Tidak Berpengaruh
z_1	$\hat{\alpha}_{11}$	$-1,37 \times 10^2$	-2,51*	Berpengaruh
	$\hat{\alpha}_{21}$	$9,19 \times 10^1$	1,59	Tidak Berpengaruh

Variabel	Parameter	Estimasi	t_{hitung}	Keterangan
z_2	$\hat{\alpha}_{31}$	$-1,11 \times 10^1$	-0,17	Tidak Berpengaruh
	$\hat{\gamma}_2$	$5,26 \times 10^{-1}$	1,11	Tidak Berpengaruh
	$\hat{\alpha}_{12}$	$5,40 \times 10^1$	0,82	Tidak Berpengaruh
	$\hat{\alpha}_{22}$	$1,23 \times 10^2$	2,22	Tidak Berpengaruh
	$\hat{\alpha}_{32}$	$1,55 \times 10^2$	2,34*	Berpengaruh
z_3	$\hat{\gamma}_3$	$-5,58 \times 10^2$	-1,75	Tidak Berpengaruh
	$\hat{\alpha}_{13}$	$-2,17 \times 10^3$	-1,81	Tidak Berpengaruh
	$\hat{\alpha}_{23}$	$-1,82 \times 10^4$	-1,76	Tidak Berpengaruh
	$\hat{\alpha}_{33}$	$5,31 \times 10^3$	2,01*	Berpengaruh
	$\hat{\gamma}_4$	$-7,69 \times 10^{-2}$	-1,02	Tidak Berpengaruh
z_4	$\hat{\alpha}_{14}$	$2,31 \times 10^2$	3,41*	Berpengaruh
	$\hat{\alpha}_{24}$	4,82	0,07	Tidak Berpengaruh
	$\hat{\alpha}_{34}$	$-9,71 \times 10^1$	0,07	Tidak Berpengaruh
	$\hat{\gamma}_5$	$1,31 \times 10^{-1}$	3,28*	Berpengaruh
	$\hat{\alpha}_{15}$	$1,37 \times 10^2$	1,66	Tidak Berpengaruh
z_5	$\hat{\alpha}_{25}$	$1,14 \times 10^2$	1,69	Tidak Berpengaruh
	$\hat{\alpha}_{35}$	$-6,73 \times 10^1$	-1,11	Tidak Berpengaruh

Catatan: * adalah parameter pada variabel prediktor yang signifikan

Berdasarkan Tabel 6. Terdapat 7 parameter yang signifikan dari 22 parameter. Pada pengujian parameter secara parsial apabila hanya ada satu parameter dalam variabel prediktor yang signifikan, maka variabel prediktor tersebut sudah dapat dikatakan signifikan. Parameter tersebut dinyatakan signifikan apabila $|t_{hitung}| > t_{tabel} = 2,20$. Secara parsial, variabel yang berpengaruh terhadap AHH adalah persentase penduduk miskin, tingkat pengangguran terbuka, rata-rata lama sekolah, *gini ratio*, prevalensi *stunting*, dan persentase bayi usia kurang dari 6 bulan yang mendapatkan ASI Eksklusif. Oleh karena itu, upaya peningkatan AHH perlu difokuskan pada penurunan angka kemiskinan dan pengangguran, perluasan akses pendidikan untuk meningkatkan rata-rata lama sekolah, serta pengurangan ketimpangan sosial ekonomi. Di sisi kesehatan, percepatan penurunan *stunting* melalui intervensi gizi terpadu dan peningkatan cakupan ASI eksklusif menjadi kunci untuk memperbaiki kualitas hidup sejak usia dini. Tindakan ini sebaiknya dilaksanakan secara lintas sektor dengan dukungan kebijakan pembiayaan yang berpihak pada kelompok rentan, penguatan layanan kesehatan primer, serta program berbasis komunitas yang berorientasi pada pencegahan. Dengan demikian, intervensi yang terarah pada faktor-faktor tersebut diharapkan mampu mendorong peningkatan AHH sekaligus mempercepat pencapaian target SDGs 3 di Indonesia.

5. KESIMPULAN

Melalui hasil analisis dan pembahasan yang telah dilakukan, dapat disimpulkan bahwa model regresi semiparametrik deret Fourier terbaik yang diperoleh yaitu dengan 3 osilasi dengan persamaan sebagai berikut:

$$\begin{aligned} \hat{y}_i = & 2,30 \times 10^2 + 3,71 \times 10^{-4}x_{1i} - 2,75 \times 10^{-1}x_{2i} + 4,64 \times 10^{-2}z_{1i} - 1,37 \times 10^2 \\ & \cos z_{1i} + 9,19 \times 10^1 \cos 2z_{1i} - 1,11 \times 10^1 \cos 3z_{1i} + 5,26 \times 10^{-1}z_{2i} + 5,40 \times 10^1 \\ & \cos z_{2i} + 1,23 \times 10^2 \cos 2z_{2i} + 1,55 \times 10^2 \cos 3z_{2i} - 5,58 \times 10^2 z_{3i} - 2,17 \times 10^3 \\ & \cos z_{3i} - 1,82 \times 10^4 \cos 2z_{3i} + 5,31 \times 10^3 \cos 3z_{3i} - 7,69 \times 10^{-2}z_{4i} + 2,31 \times 10^2 \\ & \cos z_{4i} + 4,82 \cos 2z_{4i} - 9,71 \times 10^1 \cos 3z_{4i} + 1,31 \times 10^{-1}z_{5i} + 1,37 \times 10^2 \cos z_{5i} \\ & + 1,14 \times 10^2 \cos 2z_{5i} - 6,73 \times 10^1 \cos 3z_{5i} \end{aligned}$$

Model regresi semiparametrik deret Fourier yang didapatkan menghasilkan nilai GML sebesar $3,98 \times 10^{15}$ dengan nilai RMSEP sebesar 0,99 dan diperoleh nilai R^2 sebesar 82,67 yang berarti sebesar 82,67% variasi yang terjadi pada Angka Harapan Hidup disebabkan oleh variabel prediktor. Hal ini menandakan bahwa ada sekitar 17,33% variabel lain yang belum diketahui dan tidak dimasukkan kedalam model yang memengaruhi Angka Harapan Hidup.

DAFTAR PUSTAKA

- Amalia, S.H. dan Nur, I.M. (2019) "Pemodelan Regresi Nonparametrik Deret Fourier pada Kasus Tingkat Kemiskinan di Provinsi Sumatera Utara," in *Prosiding Mahasiswa Seminar Nasional Unimus*, hal. 334–343.
- Dani, A.T.R., Putra, F.B., Zen, M.A., Sifriyani, Fauziyah, M., Ratnasari, V. dan Adrianingsih, N.Y. (2023) "Pemodelan Kadar Hemoglobin pada Pasien Demam Berdarah di Kota Samarinda Menggunakan Regresi Semiparametrik Spline Truncated," *Jambura Journal of Probability and Statistics*, 4(2), hal. 56–64.
- Dani, A.T.R. dan Adrianingsih, N.Y. (2021) "Pemodelan Regresi Nonparametrik dengan Estimator Spline Truncated dan Deret Fourier," *Jambura Journal of Mathematics*, 3(1), hal. 26–36.
- Devi, A.R., Budiantara, I.N. dan Ratnasari, V. (2021) "Unbiased Risk and Cross-Validation Method for Selecting Optimal Knots in Multivariable Nonparametric Regression Spline Truncated (Case Study: Unemployment Rate in Central Java, Indonesia, 2015)," in *AIP Conference Proceedings*.
- Eubank, R.L. (1999) *Nonparametric Regression and Spline Smoothing*. Second Edition. New York: Taylor & Francis.
- Fadlirhohim, R.D., Sifriyani dan Dani, A.T.R. (2024) "Modeling Stunting Prevalence in Indonesia Using Spline Truncated Semiparametric Regression," *BAREKENG: Journal of Mathematics and Its Applications*, 18(3), hal. 2015–2028.
- Khairunnisa, L.R., Prahutama, A. dan Santoso, R. (2020) "Pemodelan Regresi Semiparametrik dengan Pendekatan Deret Fourier (Studi Kasus: Pengaruh Indeks Dow Jones dan BI Rate Terhadap Indeks Harga Saham Gabungan)," *Jurnal Gaussian*, 9(1), hal. 50–63.
- Noviani, W., Sifriyani dan Purnamasari, I. (2020) "Pemodelan Regresi Nonparametrik Spline Linear Persentase Penduduk Miskin di Kalimantan," *Jurnal Siger Matematika*, 1(2), hal. 35–41.
- Padatuan, A.B., Sifriyani dan Prangga, S. (2021) "Pemodelan Angka Harapan Hidup dan Angka Kematian Bayi di Kalimantan dengan Regresi Nonparametrik Spline Birespon," *BAREKENG: Journal of Mathematics and Its Applications*, 15(2), hal. 283–296.
- Pasarella, M.D., Sifriyani dan Amijaya, F.D.T. (2022) "Nonparametric Regression Model Estimation With the Fourier Series the Fourier Series Approach and Its Application To the Accumulative Covid-19 Data in Indonesia," *BAREKENG: Journal of Mathematics and Its Applications*, 16(4), hal. 1167–1174.
- Pratiwi, Y. dan Mulyawan (2023) "Implementasi Algoritma K-Means untuk Menentukan Angka Harapan Hidup berdasarkan Tingkat Provinsi," *Blend Sains Jurnal Teknik*, 1(4), hal. 284–294.
- Purnaraga, T., Sifriyani dan Prangga, S. (2020) "Regresi Nonparametrik Spline pada Data Laju Pertumbuhan Ekonomi di Kalimantan," *BAREKENG: Journal of Mathematics and Its Applications*, 14(3), hal. 343–356.
- Rahmania, Sifriyani dan Dani, A.T.R. (2024) "Modeling Open Unemployment Rate in

- Kalimantan Island Using Nonparametric Regression With Fourier Series Estimator,” *BAREKENG: Journal of Mathematics and Its Applications*, 18(1), hal. 245–254.
- Rosanti, I.W. dan Budiantara, I.N. (2020) “Pemodelan Faktor-Faktor yang Mempengaruhi Morbiditas di Jawa Tengah Menggunakan Regresi Nonparametrik Spline Truncated,” *Inferensi*, 3(2), hal. 107–114.
- Sifriyani, Sari, A.R.M., Dani, A.T.R. dan Jalaluddin, S. (2023) “Bi-Response Truncated Spline Nonparametric Regression With Optimal Knot Point Selection Using Generalized Cross-Validation in Diabetes Mellitus Patient’S Blood Sugar Levels,” *Communications in Mathematical Biology and Neuroscience*, 2023(48), hal. 1–18.
- Sifriyani, Dani, A.T.R., Fauziyah, M., Hayati, M.N., Wahyuningsih, S. dan Prangga, S. (2023) “Spline and Kernel Mixed Estimators in Multivariable Nonparametric Regression for Dengue Hemorrhagic Fever Model,” *Communications in Mathematical Biology and Neuroscience*, 2023(11), hal. 1–15.
- Sitohang, F.K.S., Sifriyani dan Mahmuda, S. (2024) “Application of Truncated Spline Nonparametric Regression for Modeling the Height of Yeop Chagi Kicks of Taekwondo Athletes in Samarinda City,” *BAREKENG: Journal of Mathematics and Its Applications*, 18(2), hal. 657–666.
- Wulandari, E. dan Nurhayati (2024) “Determinan Umur Harapan Hidup di Indonesia,” *Jurnal Ekonomi Trisakti*, 4(1), hal. 217–224.