

PENERAPAN REGRESI NONPARAMETRIK *SPLINE TRUNCATED* DALAM MEMODELKAN HUBUNGAN FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS PEMBANGUNAN GENDER (IPG) DI PROVINSI SUMATERA UTARA

Raihan Ramadhan¹, Irmeilyana^{2*}, Anita Desiani³

^{1,2,3} Universitas Sriwijaya, Fakultas MIPA Jurusan Matematika

*irmeilyana@unsri.ac.id

Abstract

The Gender Development Index (GDI) achievements of the North Sumatra Province in 2021 are still lower than the national IPG. North Sumatra Province faces challenges in realizing gender equality, including low average years of schooling and income gap between genders. These things show that the GDI in North Sumatra Province needs to be improved. This research aims to determine the best regression model and identify the significant factors influencing GDI in North Sumatra Province. This study utilizes a truncated spline nonparametric regression method because there is no discernible pattern between the response variable and the six predictor variables used in the data. The number of knot points used is 1 knot, 2 knots, 3 knots, and a combination of knots, and the order used is order 1. The optimal knot points and best model selection use Generalized Cross-Validation (GCV) based on the minimum GCV value. The best model is measured using the coefficient of determination (R^2). The results of this research are the best-truncated spline nonparametric regression model using a combination of knots (3, 3, 1, 1, 3, 3) with a minimum GCV value of 5.64129. Female labor force participation rate (X_1), maternal years of schooling (X_3), literacy rate (X_4), and life expectancy (X_6) are significant factors that influence GDI in North Sumatra Province. The best model obtained has an R^2 value of 92.79% with identical, independent residuals and normal distribution (IIDN) fulfilled.

Keywords: GDI, Truncated spline nonparametric regression, Knot points, GCV, coefficient of determination

Abstrak

Capaian Indeks Pembangunan Gender (IPG) Provinsi Sumatera Utara (Sumut) pada tahun 2021 masih berada di bawah IPG nasional. Provinsi Sumut masih dihadapkan dengan tantangan pembangunan untuk mewujudkan kesetaraan gender seperti rendahnya rata-rata lama sekolah dan tingginya kesenjangan pendapatan antar gender. Hal-hal tersebut menunjukkan bahwa IPG di Provinsi Sumut perlu ditingkatkan. Penelitian ini bertujuan untuk memperoleh model regresi terbaik dan mengetahui faktor-faktor yang berpengaruh signifikan terhadap IPG Provinsi Sumut. Penelitian ini menggunakan metode regresi nonparametrik *spline truncated* karena pola data antara variabel respon dengan 6 variabel prediktor dari data yang digunakan tidak membentuk pola tertentu. Banyaknya titik knot yang digunakan yaitu 1 knot, 2 knot, 3 knot dan kombinasi knot, serta orde yang digunakan adalah orde 1. Pemilihan titik knot optimal dan model terbaik menggunakan *Generalized Cross Validation* (GCV) berdasarkan nilai GCV minimum dan model terbaik yang diperoleh diukur menggunakan koefisien determinasi (R^2). Hasil penelitian ini yaitu model regresi nonparametrik *spline truncated* terbaik menggunakan kombinasi knot (3, 3, 1, 1, 3, 3) dengan nilai GCV minimum sebesar 5,64129. Faktor-faktor yang berpengaruh signifikan terhadap IPG di Provinsi Sumut yaitu TPAK perempuan (X_1), RLS perempuan (X_3), AMH perempuan (X_4), dan AHH perempuan (X_6). Model terbaik yang diperoleh memiliki nilai R^2 sebesar 92,79% dengan asumsi residual identik, independen, dan distribusi normal (IIDN) terpenuhi.

Kata kunci : IPG, Regresi nonparametrik *spline truncated*, Titik Knot, GCV, Koefisien Determinasi

1. Pendahuluan

Indeks Pembangunan *Gender* (IPG) adalah indikator yang menyatakan ukuran angka kemajuan pembangunan manusia berbasis *gender* berdasarkan tiga dimensi dasar kehidupan, yaitu ekonomi, pengetahuan, dan kesehatan dengan tujuan untuk melihat tingkat kesetaraan *gender* [1]. Tingkat kesetaraan *gender* di Indonesia disebabkan oleh capaian IPG pada masing-masing provinsi di Indonesia yang berbeda-beda, salah satunya adalah Provinsi Sumatera Utara (Sumut). Capaian IPG Provinsi Sumut pada tahun 2021 masih berada di bawah IPG nasional dan masih dihadapkan dengan tantangan pembangunan untuk mewujudkan kesetaraan *gender* seperti masih rendahnya rata-rata lama sekolah dan tingginya kesenjangan pendapatan perempuan dan laki-laki [2]. Hal-hal tersebut menunjukkan bahwa kesetaraan *gender* di Provinsi Sumut perlu ditingkatkan.

Untuk melihat capaian IPG yang perlu ditingkatkan, dibutuhkan model matematis yang dapat melihat pengaruh dimensi dari masing-masing faktor terhadap capaian IPG, salah satunya adalah regresi. Regresi bekerja sangat baik pada data yang telah diketahui pola datanya. Sayangnya tidak semua data diketahui pola datanya. Data yang tidak memiliki pola tertentu sebaiknya didekati dengan model regresi nonparametrik [3]. Regresi nonparametrik memiliki prinsip dasar yaitu terletak pada bentuk fungsi regresi yang diasumsikan belum diketahui pola datanya dan perlu adanya pendekatan fungsi tertentu, salah satunya dapat menggunakan pendekatan *spline* [4][5].

Spline lebih efektif dalam menyesuaikan diri terhadap karakteristik suatu data karena terdapat suatu basis fungsi, salah satunya adalah *spline truncated* [6]. *Spline truncated* merupakan model polinomial yang memiliki sifat tersegmen. Kelebihan dari basis fungsi ini yaitu memiliki fleksibilitas dan tingkat akurasi yang sangat baik dalam memodelkan pola data fluktuatif dan sebagai alternatif asumsi normalitas yang sulit terpenuhi [5].

Berdasarkan [7], beberapa hal yang perlu diperhatikan dalam membentuk model regresi nonparametrik *spline truncated* yaitu

banyaknya titik knot dan lokasi titik knot. Penentuan titik knot dan lokasi titik knot secara sebarang dapat menghasilkan model yang berbeda dan kurang tepat. Sehingga perlu dilakukan pemilihan titik knot secara optimal, salah satunya yaitu menggunakan *Generalized Cross Validation* (GCV) [8].

Beberapa penelitian yang telah menerapkan regresi nonparametrik *spline truncated* dengan GCV pada IPG Provinsi di Indonesia. [9] menyatakan bahwa faktor-faktor yang berpengaruh terhadap IPG Jawa Barat dari setiap dimensi diantaranya dimensi kesehatan adalah Angka Keluhan Kesehatan (AKK) perempuan, dimensi ekonomi adalah Tingkat Pengangguran Terbuka (TPT) perempuan dan Tingkat Partisipasi Angkatan Kerja (TPAK) perempuan dengan nilai R^2 sebesar 99,98%. [10] menjelaskan bahwa faktor-faktor yang berpengaruh IPG di Pulau Kalimantan seperti dimensi pendidikan adalah Angka Partisipasi Sekolah (APS) SD dan SMP perempuan, dan dimensi ekonomi adalah TPAK perempuan dengan nilai R^2 sebesar 71,55%.

Penelitian ini bertujuan untuk memperoleh model regresi nonparametrik *spline truncated* terbaik dan mengetahui faktor-faktor yang berpengaruh signifikan terhadap IPG Provinsi Sumut. Data yang digunakan adalah data sekunder yang diperoleh dari BPS Provinsi Sumut pada tahun 2021 dengan variabel yang digunakan sebanyak 7 variabel yang terdiri atas 1 variabel respon yaitu IPG dan 6 variabel prediktor yaitu TPAK, TPT, RLS, AMH, AKK, dan AHH dengan objek *gender* perempuan. Banyaknya titik knot yang digunakan yaitu 1, 2, 3 knot dan kombinasi knot serta orde yang digunakan adalah orde 1. Pemilihan titik knot optimal dan model terbaik menggunakan GCV dengan melihat nilai GCV minimum dan Estimasi parameter menggunakan *Ordinary Least Square* (OLS).

2. Metoda Penelitian

Langkah-langkah yang dilakukan dalam penelitian adalah:

1. Melakukan analisis statistika deskriptif pada variabel penelitian.

2. Membuat *scatter plot* antara variabel respon dengan masing-masing variabel prediktor.

3. Melakukan pemodelan data penelitian menggunakan regresi nonparametrik *spline truncated* sebagai berikut:

a. Regresi Nonparametrik

Model regresi nonparametrik sebagai berikut [11] :

$$Y_i = f(X_i) + \varepsilon_i ; i = 1, 2, \dots, n \quad (1)$$

dimana :

- Y_i : nilai variabel respon ke- i
- $f(X_i)$: fungsi kurva regresi yang tidak diketahui bentuknya
- ε_i : nilai *error* random ke- i yang diasumsikan IIDN $(0, \sigma^2)$
- i : pengamatan
- n : banyaknya pengamatan

b. Regresi Nonparametrik *Spline Truncated Spline truncated* diperoleh dari penjumlahan antara fungsi polinomial dengan fungsi *truncated*. Model regresi nonparametrik *spline truncated* [12] :

$$Y_i = \beta_0 + \sum_{j=1}^p \beta_{j1} X_{ji} + \sum_{j=1}^p \sum_{u=1}^r \beta_{j(1+u)} (X_{ji} - K_{ju})_+ + \varepsilon_i \quad (2)$$

$$(X_{ji} - K_{ju})_+ = \begin{cases} (X_{ji} - K_{ju}) & , X_{ji} \geq K_{ju} \\ 0 & , X_{ji} < K_{ju} \end{cases}$$

dimana:

- Y_i : nilai variabel respon ke- i
- β_0 : Intersep model polinomial
- β_{j1} : parameter model polinomial *linier*
- X_{ji} : variabel prediktor ke- j pada pengamatan ke- i
- $\beta_{j(1+u)}$: parameter *truncated*
- K_{ju} : titik knot *truncated* ke- ju
- j : indeks variabel prediktor
- p : banyaknya indeks variabel prediktor
- u : titik knot
- r : banyaknya titik knot

Berdasarkan Persamaan (2), dapat dibentuk matriks sebagai berikut:

$$Y = X\beta + \varepsilon \quad (3)$$

dengan,

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \beta = \begin{bmatrix} \beta_0 \\ \beta_{11} \\ \beta_{12} \\ \vdots \\ \beta_{1(1+r)} \\ \vdots \\ \beta_{p1} \\ \beta_{p2} \\ \vdots \\ \beta_{p(1+r)} \end{bmatrix} \quad \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

$$X = \begin{bmatrix} 1 & X_{11} & (X_{11} - K_{11}) & \dots & (X_{11} - K_{1r}) & X_{p1} & (X_{p1} - K_{11}) & \dots & (X_{p1} - K_{pr}) \\ 1 & X_{12} & (X_{12} - K_{11}) & \dots & (X_{12} - K_{1r}) & X_{p1} & (X_{p2} - K_{11}) & \dots & (X_{p2} - K_{pr}) \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & X_{1n} & (X_{1n} - K_{11}) & \dots & (X_{1n} - K_{1r}) & X_{pn} & (X_{pn} - K_{11}) & \dots & (X_{pn} - K_{pr}) \end{bmatrix}$$

Keterangan :

- Y : vektor variabel respon $n \times 1$
- β : vektor parameter berukuran $(p \times (1 + r) + 1) \times 1$
- ε : vektor residual $n \times 1$
- X : matriks variabel prediktor $n \times (p \times (1 + r) + 1)$

4. Memilih titik knot optimal pada 1 knot, 2 knot, 3 knot, dan kombinasi knot.

Titik knot merupakan titik perpaduan bersama yang menunjukkan terjadinya perubahan pola perilaku data. Salah satu teknik yang dapat digunakan dalam melakukan pemilihan titik knot optimal adalah GCV [8]. Persamaan GCV yang digunakan yaitu [13] :

$$GCV(K) = \frac{MSE(K)}{[n^{-1}trace(I - A(K))]^2} = \frac{n^{-1} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2}{[n^{-1}trace(I - (X(X^T X)^{-1} X^T))]^2} \quad (4)$$

Keterangan :

- $GCV(K)$: vektor GCV dari titik-titik knot
- $MSE(K)$: error kuadrat *mean* dari titik-titik knot
- K : K_{j1}, K_{j2}, K_{j3}
- I : matriks identitas
- $A(K)$: matriks dari titik knot

5. Melakukan pemilihan model terbaik berdasarkan kriteria GCV minimum titik-titik knot optimal yang diperoleh pada langkah 5.

6. Menguji Signifikansi Parameter

Estimasi parameter pada penelitian ini menggunakan *Ordinary Least Square* (OLS), yaitu meminimalkan jumlah residual kuadrat estimator yang telah diperoleh. Estimasi parameter dalam bentuk matriks yaitu [13],

$$\hat{\beta} = (X^T X)^{-1} X^T Y \quad (5)$$

Keterangan :

- $\hat{\beta}$: vektor estimasi parameter
- $(X^T X)^{-1}$: invers matriks $X^T X$
- X : matriks variabel prediktor
- X^T : *transpose* matriks variabel prediktor
- Y : matriks variabel respon

a. Uji Serentak (uji F)

Uji F merupakan pengujian parameter model regresi secara bersama-

sama, adapun hipotesis yang digunakan yaitu [12].

$$H_0: \beta_{11} = \beta_{12} = \dots = \beta_{p(1+r)} = 0$$

$$H_1: \text{minimal terdapat satu } \beta_{jl} \neq 0,$$

$$j = 1, 2, 3, \dots, p; l = 1, 2, \dots, 1 + r$$

F_{hitung}

$$= \frac{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p \times (1+r) + 1}}{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - (p \times (1+r) + 1) - 1}} \quad (6)$$

Keputusan tolak H_0 , jika $F_{hitung} > F_{(a,(p+pr+1),(n-(p+pr+1)-1))}$ atau $p\text{-value} < \alpha$. Jika H_0 ditolak berarti bahwa minimal terdapat satu parameter yang tidak sama dengan nol atau minimal ada satu variabel prediktor berpengaruh signifikan terhadap variabel respon dalam model regresi.

b. Uji Parsial (uji t)

Uji t merupakan pengujian parameter model regresi secara individu yang ertujuan untuk mengetahui parameter mana saja yang tidak berpengaruh dan berpengaruh signifikan.

$$H_0 : \beta_{jl} = 0$$

$$H_1 : \beta_{jl} \neq 0, j = 1, 2, 3, \dots, p;$$

$$l = 1, 2, \dots, 1 + r$$

$$t_{hitung} = \frac{\hat{\beta}_{jl}}{se(\hat{\beta}_{jl})} \quad (7)$$

dimana :

$\hat{\beta}_{jl}$: koefisien variabel prediktor

$se(\hat{\beta}_{jl})$: standart error dari $\hat{\beta}_{jl}$

Keputusan tolak H_0 , jika $|t_{hitung}| > t_{(\alpha/2, n-(p \times (1+r) + 1) - 1)}$ atau $p\text{-value} < \alpha$. Jika H_0 ditolak, dapat ditarik kesimpulan bahwa variabel prediktor ke- j berpengaruh signifikan terhadap variabel respon.

7. Menghitung Koefisien Determinasi (R^2)

R^2 merupakan ukuran proporsi *varians* total di sekitar \bar{y} yang dapat dijelaskan oleh model regresi. Nilai R^2 dapat diperoleh dengan rumus berikut [12].

$$R^2 = \frac{SSR}{SST} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y}_i)^2} \quad (8)$$

Keterangan :

\bar{y} : mean dari variabel respon

y_i : nilai variabel respon ke- i

\bar{y}_i : mean dari nilai variabel respon ke- i

i : pengamatan

n : banyaknya pengamatan

8. Melakukan Uji Asumsi Residual

Berdasarkan Persamaan (3), persamaan residual dapat dibentuk dalam bentuk matriks sebagai berikut,

$$\boldsymbol{\varepsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\beta} \quad (9)$$

Uji asumsi residual dilakukan untuk mengetahui apakah residual yang diperoleh telah memenuhi asumsi identik, independen, dan berdistribusi normal.

a. Uji Identik

Uji identik atau homoskedastisitas yaitu *varians* (σ^2) dari residual yang diperoleh adalah identik (sama) dan tidak terjadi heteroskedastisitas.

Heteroskedastisitas dapat dideteksi dengan uji *glejser* [9].

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

$$H_1 : \text{minimal ada satu, } \sigma_i^2 \neq \sigma^2;$$

$$i = 1, 2, 3, \dots, n$$

F_{hitung}

$$= \frac{\frac{\sum_{i=1}^n (|\hat{\varepsilon}_i| - |\bar{\varepsilon}|)^2}{p \times (1+r) + 1}}{\frac{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)^2}{n - ((p \times (1+r) + 1) - 1)}} \quad (10)$$

Keputusan tolak H_0 , apabila $F_{hitung} > F_{(a,(p(1+r)),(n-p(1+r)-1))}$. Jika H_0 ditolak, artinya terdapat kasus heteroskedastisitas.

b. Uji Independen

Uji independen atau uji autokorelasi adalah asumsi model regresi yang mengharuskan tidak adanya korelasi antar residual. Untuk mendeteksi autokorelasi dapat menggunakan uji *Durbin-Watson (DW)*.

$$H_0 : \rho = 0$$

$$H_1 : \rho \neq 0$$

$$d_{hitung} = \frac{\sum_{i=1}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2} \quad (11)$$

Keputusan tolak H_0 , apabila $d_{hitung} \leq d_{L,\alpha,n}$. Jika H_0 ditolak, maka dapat diambil kesimpulan bahwa terdapat korelasi antar residual.

c. Uji Distribusi Normal

Untuk mengetahui residual berdistribusi normal dapat menggunakan uji *Kolmogorov-Smirnov* atau *KS*.

$$H_0 : F(Z_i) = F(X_i)$$

$$H_1 : F(Z_i) \neq F(X_i)$$

$$D = \text{Sup}_X |F(Z_i) - F(X_i)| \quad (12)$$

Daerah penolakan pada uji *KS* yaitu $|D| > D_\alpha$ atau *p-value* $< \alpha$. Jika H_0 ditolak, maka residual tidak berdistribusi normal.

- Melakukan interpretasi dari model regresi nonparametrik *spline truncated* terbaik.

Taraf signifikansi (α) yang digunakan yaitu sebesar 5%. Software yang digunakan yaitu *Minitab 17* dan *R Studio*.

3. Hasil Penelitian

3.1. Statistika Deskriptif

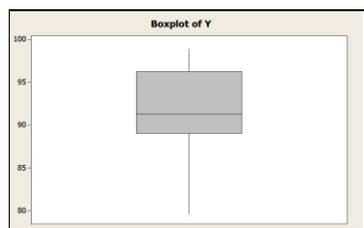
Hasil statistika deskriptif disajikan dalam Tabel 1. Setiap variabel memiliki karakteristik tingkat pengaruh faktor yang beragam seperti nilai *mean*, maksimum, dan minimum.

Tabel 1. Statistika Deskriptif Variabel Penelitian

Variabel	Mean	Maksimum	Minimum
<i>Y</i>	92,004	98,89	79,63
X_1	61,22	86,58	39,9
X_2	4,887	12,36	0,39
X_3	8,841	11,35	4,67
X_4	98,474	100	94,08
X_5	21,35	32,02	10,51
X_6	70,991	75,58	64,47

Keterangan: *Y* : IPG (*indeks*); X_1 : TPAK Perempuan (*persen*); X_2 : TPT Perempuan (*persen*); X_3 : RLS Perempuan (*tahun*); X_4 : AMH Perempuan (*persen*); X_5 : AKK Perempuan (*persen*); X_6 : AHH Perempuan (*tahun*)

Selain itu, akan dilakukan analisis statistika deskriptif menggunakan *boxplot* untuk mengetahui nilai kuartil dan melihat ada tidaknya data *outlier*. Hasil *boxplot* pada masing-masing variabel ditampilkan pada Gambar 1.

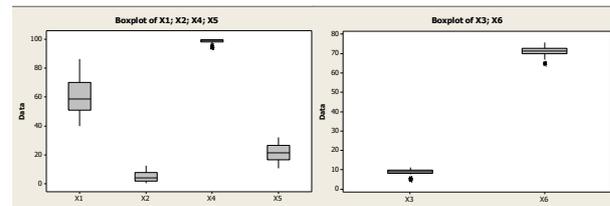


Gambar 1. *Boxplot* Variabel Respon

Berdasarkan Gambar 1, dapat dilihat bahwa data variabel *Y* memiliki panjang *box* yang ditentukan *interquartile range* (IQR)

cenderung semakin tinggi. Artinya data IPG memiliki *varians* relatif besar atau datanya bervariasi semakin menyebar. Selanjutnya, dilihat dari jarak median ke Q_1 dan jarak median ke Q_3 yang panjangnya tidak sama. Artinya data IPG tidak berbentuk simetris. Selain itu, tidak ada data yang terpisah jauh dengan yang lainnya atau dapat dikatakan tidak terdapat *outlier*.

Sementara itu, grafik *boxplot* untuk variabel prediktor divisualisasikan berdasarkan satuan persen dan tahun. Grafik *boxplot* menurut satuan persen yang terdiri atas data variabel X_1, X_2, X_4 , dan X_5 , sedangkan grafik *boxplot* menurut satuan tahun yang terdiri atas data variabel X_3 dan X_6 secara visual disajikan dalam Gambar 2.



Gambar 2. *Boxplot* (a) Variabel Prediktor (dalam %) dan (b) Variabel Prediktor (dalam Satuan Tahun)

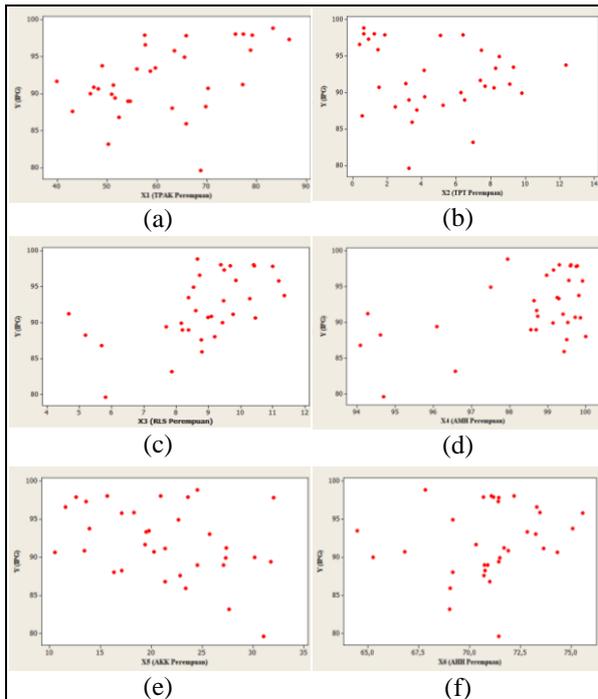
Berdasarkan Gambar 2. *boxplot* X_1, X_2 dan X_5 memiliki bidang IQR yang relatif panjang. Artinya data TPAK perempuan, TPT perempuan, dan AKK perempuan memiliki data yang bervariasi atau semakin menyebar. Sedangkan *boxplot* X_4 memiliki bidang IQR yang relatif pendek. Artinya RLS perempuan, AMH perempuan, dan AHH perempuan memiliki data yang kurang bervariasi atau relatif tidak menyebar. Sementara itu, *boxplot* X_5 memiliki bentuk simetris. Hal ini dapat dilihat dari jarak median ke Q_1 dan jarak median ke Q_3 yang cenderung sama panjangnya dan tidak terdapat *outlier*.

Selanjutnya, grafik *boxplot* menurut satuan tahun yang terdiri atas data variabel X_3 dan X_6 . *boxplot* X_3 dan X_6 memiliki bidang IQR yang relatif pendek. Artinya RLS perempuan dan AHH perempuan memiliki data yang kurang bervariasi atau relatif tidak menyebar. Jarak median ke Q_1 dan jarak median ke Q_3 yang tidak sama panjangnya, Artinya RLS perempuan dan AHH perempuan tidak berbentuk simetris. Selain itu, RLS perempuan dan AHH perempuan

memiliki *outlier* pada bagian bawah menandakan data cenderung ke kiri.

3.2. Scatterplot antara Variabel Respon dengan Variabel Prediktor

Hasil *scatter plot* antara IPG dengan masing-masing variabel prediktor ditampilkan pada Gambar 3. Berdasarkan Gambar 3, dapat diketahui bahwa data pada *scatter plot* menyebar, tidak menunjukkan sebuah pola tertentu dan berubah-ubah pada sub-sub interval tertentu. Dengan demikian, dalam penelitian ini sebaiknya menggunakan pendekatan regresi nonparametrik.



Gambar 3. *Scatter Plot* (a) TPAK Perempuan, (b) TPT Perempuan, (c) RLS Perempuan, (d) AMH Perempuan, (e) AKK Perempuan, dan (f) AHH Perempuan Terhadap IPG

3.3. Pemodelan Regresi Nonparametrik Spline Truncated

Pada penelitian ini terdapat 1 variabel respon (Y) dan 6 variabel prediktor (X_1, X_2, \dots, X_6) pada 33 kabupaten/kota di Provinsi Sumut ($n = 33$). Model regresi nonparametrik secara umum berdasarkan Persamaan (1). Selanjutnya, hampiran fungsi $f(X_i)$ dengan fungsi *spline truncated* berorde m dan titik knot *truncated* K_{ju} yang diberikan oleh Persamaan (2). Misalkan diberikan $p = 6$ pada orde *linier* ($m = 1$) dengan $r = 3$, sehingga diperoleh model regresi

nonparametrik *spline truncated multivariable* sebagai berikut :

$$Y_i = \beta_0 + \sum_{j=1}^6 \beta_{j(c)} X_{j(i)} + \sum_{j=1}^6 \sum_{u=1}^3 \beta_{j(1+u)} (X_{j(i)} - K_{ju})_+ + \varepsilon_i$$

$$(X_{j(i)} - K_{ju})_+ = \begin{cases} (X_{j(i)} - K_{ju}) & , X_{j(i)} \geq K_{ju} \\ 0 & , X_{j(i)} < K_{ju} \end{cases}$$

Untuk dapat membentuk model regresi nonparametrik *spline truncated*, diperlukan informasi mengenai lokasi penempatan knot dan banyaknya titik knot yang akan digunakan. Informasi tersebut dapat diperoleh dengan melakukan pemilihan titik optimal.

3.4. Pemilihan Titik Knot Optimal

Pemilihan titik knot optimal dalam pembahasan ini menggunakan *Generalized Cross Validation* (GCV) yang diberikan oleh Persamaan (3).

3.4.1. 1 Knot

Pemilihan titik knot optimal diawali dengan melakukan pemilihan titik knot optimal dengan 1 knot terlebih dahulu. Hasil perhitungan nilai GCV berdasarkan Persamaan (4) pada 1 knot menghasilkan 48 iterasi dan diketahui bahwa nilai GCV minimum pada 1 knot sebesar 12,5676 yang terletak pada iterasi ke-9. Titik-titik knot optimal 1 knot pada masing-masing variabel prediktor sebagai berikut :

$$\begin{matrix} K_{11} = 48,473 & K_{21} = 2,59 & K_{31} = 5,897 \\ K_{41} = 95,167 & K_{51} = 14,46 & K_{61} = 66,51 \end{matrix}$$

3.4.2. 2 Knot

Hasil perhitungan nilai GCV berdasarkan Persamaan (4) pada 2 knot menghasilkan 1.225 iterasi. Nilai GCV minimum 2 knot adalah 8,8698 yang terdapat pada iterasi ke-486 dengan banyaknya titik knot sebanyak 12 dengan lokasi penempatan:

$$\begin{matrix} K_{11} = 50,38 & K_{12} = 52,28 & K_{21} = 3,08 \\ K_{22} = 3,57 & K_{31} = 6,17 & K_{32} = 6,44 \\ K_{41} = 95,41 & K_{42} = 95,65 & K_{51} = 15,34 \\ K_{52} = 16,22 & K_{61} = 66,96 & K_{62} = 67,42 \end{matrix}$$

3.4.3. 3 Knot

Hasil perhitungan nilai GCV berdasarkan Persamaan (4) pada 3 knot menghasilkan 17.296 iterasi. Nilai GCV minimum 3 knot adalah 5,955544 yang

terletak pada iterasi ke-8897 dengan titik knot optimal pada masing-masing variabel prediktor sebagai berikut:

$$\begin{aligned}
 K_{11} &= 50,38 & K_{12} &= 52,28 & K_{13} &= 53,24 \\
 K_{21} &= 3,08 & K_{22} &= 3,57 & K_{23} &= 3,81 \\
 K_{31} &= 6,17 & K_{32} &= 6,44 & K_{33} &= 6,58 \\
 K_{41} &= 95,41 & K_{42} &= 95,65 & K_{43} &= 95,77 \\
 K_{51} &= 15,34 & K_{52} &= 16,22 & K_{53} &= 16,65 \\
 K_{61} &= 66,96 & K_{62} &= 67,41 & K_{63} &= 67,64
 \end{aligned}$$

3.4.4. Kombinasi Knot

Selanjutnya, dilakukan pemilihan titik knot optimal dengan kombinasi knot antara 1 knot, 2 knot, dan 3 knot. Hasil perhitungan nilai GCV berdasarkan Persamaan (4) pada kombinasi knot menghasilkan 729 iterasi dan diketahui bahwa nilai GCV minimum kombinasi knot adalah 5,641295 yang terdapat pada iterasi ke-657. Titik knot optimal untuk variabel X_1 dan X_2 masing-masing sebanyak 3 titik knot, untuk variabel X_3 dan X_4 masing-masing sebanyak 1 titik knot, untuk variabel X_5 dan X_6 masing-masing sebanyak 3 titik knot atau kombinasi knot (3, 3, 1, 1, 3, 3). Adapun titik knot optimal dari hasil kombinasi knot (3, 3, 1, 1, 3, 3) pada masing-masing variabel prediktor sebagai berikut:

$$\begin{aligned}
 K_{11} &= 50,38 & K_{12} &= 52,28 & K_{13} &= 53,24 \\
 K_{21} &= 3,08 & K_{22} &= 3,57 & K_{23} &= 3,81 \\
 K_{31} &= 5,897 & K_{41} &= 95,17 & K_{51} &= 15,34 \\
 K_{52} &= 16,22 & K_{53} &= 16,65 & K_{61} &= 66,96 \\
 K_{62} &= 67,41 & K_{63} &= 67,64
 \end{aligned}$$

3.5. Pemilihan Model Terbaik

Hasil GCV minimum dalam pemilihan model terbaik disajikan dalam Tabel 2.

Tabel 2. GCV Minimum

Titik Knot Optimal	GCV Minimum	Banyaknya Titik Knot	Banyaknya Parameter
1 Knot	12,5676	6	13
2 Knot	8,8698	12	19
3 Knot	5,9555	18	25
Kombinasi Knot	5,64129	14	21

Keterangan: angka yang dicetak tebal adalah nilai GCV minimum

Berdasarkan Tabel 2, diketahui bahwa kombinasi knot (3, 3, 1, 1, 3, 3) memiliki nilai GCV paling minimum yaitu 5,64129 dengan 14 titik knot dan 21 parameter yang juga termasuk parameter β_0 . Model regresi nonparametrik *spline truncated* terbaik yang

terbentuk adalah menggunakan kombinasi knot (3, 3, 1, 1, 3, 3).

3.6. Pengujian Signifikansi Parameter

Estimasi model terbaik berdasarkan Persamaan (5), sehingga diperoleh hasil pada Tabel 3.

Tabel 3. Estimasi Parameter Model Terbaik

Variabel	Parameter	Koefisien
Intersep	β_0	-46,3853
X_1	$\beta_{1(1)}$	-0,79128
	$\beta_{1(2)}$	6,505965
	$\beta_{1(3)}$	-9,15213
	$\beta_{1(4)}$	3,445731
X_2	$\beta_{2(1)}$	-2,22627
	$\beta_{2(2)}$	-36,3932
	$\beta_{2(3)}$	-7,60046
	$\beta_{2(4)}$	46,90338
X_3	$\beta_{3(1)}$	-10,7529
	$\beta_{3(2)}$	13,32853
X_4	$\beta_{4(1)}$	8,482316
	$\beta_{4(2)}$	-8,9211
X_5	$\beta_{5(1)}$	-0,09121
	$\beta_{5(2)}$	-8,67153
	$\beta_{5(3)}$	22,26999
	$\beta_{5(4)}$	-13,3014
X_6	$\beta_{6(1)}$	-2,35242
	$\beta_{6(2)}$	10,34215
	$\beta_{6(3)}$	-1,25605
	$\beta_{6(4)}$	-7,05387

3.6.1. Uji Serentak (Uji F)

Uji F melibatkan data penelitian dan koefisien $\beta_{j(l)}$ pada Tabel 3 untuk mencari nilai F_{hitung} menggunakan tabel ANOVA. Hasil ANOVA model secara serentak disajikan pada Tabel 5.

Tabel 5 ANOVA Pengujian Parameter Secara Serentak

Sumber Variasi	Df	Sum of Square (SS)	Mean Square (MS)	F_{hitung}
Regresi	20	651,7675	32,58838	7,7278
Error	12	50,60422	4,217019	
Total	32	702,3718		

Berdasarkan Tabel 5, perhitungan nilai F_{hitung} yang diberikan oleh Persamaan (6):

$$F_{hitung} = \frac{32,58838}{4,217019} = 7,727824$$

Nilai $F_{hitung} = 7,727824 > F_{(0,05,20,12)} = 2,5435$ dan nilai $p - value = 0,000396 <$

0,05, sehingga dapat diambil keputusan tolak H_0 yang artinya minimal terdapat satu parameter variabel prediktor yang berpengaruh signifikan terhadap IPG Provinsi Sumut.

3.6.2. Uji Parsial (Uji t)

Uji t dilakukan berdasarkan Persamaan (7) dengan menggunakan data penelitian dan koefisien $\beta_{j(l)}$ pada Tabel 3. Hasil perhitungan t_{hitung} untuk masing-masing variabel prediktor disajikan pada Tabel 6.

Berdasarkan Tabel 6, terdapat 7 parameter yang signifikan yaitu $\beta_{1(1)}$ dan $\beta_{1(2)}$ untuk X_1 , $\beta_{3(1)}$ dan $\beta_{3(2)}$ untuk X_3 , $\beta_{4(1)}$ dan $\beta_{4(2)}$ untuk X_4 , dan $\beta_{6(2)}$ untuk X_6 , serta 13 parameter lainnya tidak signifikan. Apabila terdapat satu parameter saja yang signifikan, maka variabel prediktor tersebut berpengaruh signifikan terhadap variabel respon [14]. Sehingga, dapat disimpulkan bahwa dari 6 variabel prediktor yang digunakan terdapat 4 variabel prediktor yang berpengaruh signifikan terhadap IPG Provinsi Sumut yaitu X_1, X_3, X_4 , dan X_6 .

Tabel 6. Uji Signifikansi Secara Parsial

Var	Para	t_{hitung}	Keputusan	Simpulan
X_1	β_0	-2,38281	Tolak H_0	Sig.
	$\beta_{1(1)}$	-2,74966	Tolak H_0	
	$\beta_{1(2)}$	2,932881	Tolak H_0	Sig.
	$\beta_{1(3)}$	-1,74509	Gagal Tolak H_0	
	$\beta_{1(4)}$	1,015548	Gagal Tolak H_0	
X_2	$\beta_{2(1)}$	-1,7542	Gagal Tolak H_0	
	$\beta_{2(2)}$	-1,09046	Gagal Tolak H_0	Tidak Sig.
	$\beta_{2(3)}$	-0,98889	Gagal Tolak H_0	
	$\beta_{2(4)}$	1,241837	Gagal Tolak H_0	
X_3	$\beta_{3(1)}$	-2,67282	Tolak H_0	Sig.
	$\beta_{3(2)}$	3,101504	Tolak H_0	
X_4	$\beta_{4(1)}$	2,329396	Tolak H_0	Sig.
	$\beta_{4(2)}$	-2,26195	Tolak H_0	
X_5	$\beta_{5(1)}$	-0,12824	Gagal Tolak H_0	
	$\beta_{5(2)}$	-1,48746	Gagal Tolak H_0	Tidak Sig.
	$\beta_{5(3)}$	1,645206	Gagal Tolak H_0	
	$\beta_{5(4)}$	-1,52793	Gagal Tolak H_0	
X_6	$\beta_{6(1)}$	-1,67411	Gagal Tolak H_0	
	$\beta_{6(2)}$	2,93738	Tolak H_0	Sig.
	$\beta_{6(3)}$	-1,68713	Gagal Tolak H_0	
	$\beta_{6(4)}$	-1,92173	Gagal Tolak H_0	

Keterangan: angka yang dicetak tebal merupakan parameter dan variabel signifikan. Sign. = Signifikan

3.7. Koefisien Determinasi (R^2)

Nilai R^2 diperoleh dari Tabel 5 yang disubstitusikan dalam Persamaan (8) sebagai berikut :

$$R^2 = \frac{651,7675}{702,3718} \times 100\% = 92,79\%$$

Hasil R^2 yang diperoleh sebesar 92,79%. Hal ini berarti secara simultan semua variabel prediktor yang digunakan dalam penelitian ini mampu menjelaskan model sebesar 92,79%, dan sisanya sebesar 7,21% dijelaskan oleh variabel lain yang tidak termasuk dalam model.

3.8. Pengujian Asumsi Residual

Nilai residual (ϵ_i) diperoleh dengan melakukan perhitungan data penelitian pada nilai-nilai titik knot optimal kombinasi knot, dan koefisien $\beta_{j(l)}$ pada Tabel 3, dengan berdasarkan Persamaan (9). Hasil perhitungan ϵ_i disajikan dalam Tabel 7.

Tabel 7. Output Residual

ϵ_i	Nilai Residual	ϵ_i	Nilai Residual
1	-0,74853	18	1,715928
2	-0,21516	19	-1,55533
3	-0,11105	20	0,063982
4	1,197655	21	-3,4904
5	0,785396	22	-0,20478
6	-0,13771	23	1,598952
7	2,302086	24	-0,38618
8	-0,50034	25	0,105171
9	1,16109	26	2,297448
10	1,620388	27	0,32621
11	0,038824	28	0,256281
12	-0,89568	29	-2,05518
13	-1,6309	30	-0,92444
14	1,31732	31	-0,06943
15	-1,29124	32	-0,05558
16	0,372126	33	-0,17189
17	-0,71505		

Pengujian asumsi residual dilakukan untuk mengetahui apakah residual yang diperoleh telah memenuhi asumsi IIDN (identik, independen, dan berdistribusi normal) sebagai berikut.

3.8.1. Uji Identik

Uji identik diperoleh dengan melakukan perhitungan residual pada Tabel ,

sehingga diperoleh tabel ANOVA pada Tabel 8. Perhitungan nilai F_{hitung} berdasarkan Persamaan (10) sebagai berikut.

$$F_{hitung} = \frac{0,8639496}{0,4559818} = 1,894702$$

Tabel 8. ANOVA Uji Identik

Sumber Variasi	Df	SS	MS	F_{hitung}
Regresi	20	17,27899	0,8639496	1,894702
Error	12	5,471782	0,4559818	
Total	32	22,75077		

Nilai $F_{hitung} = 1,894702 < F_{(0,05,20,12)} = 2,5435$, sehingga dapat diambil keputusan gagal tolak H_0 yang artinya tidak terjadi heteroskedastisitas. Dengan demikian, hal ini menunjukkan bahwa asumsi residual identik terpenuhi.

3.8.2. Uji Independen

Uji independen diperoleh dengan melakukan perhitungan residual pada Tabel 7 menggunakan uji DW berdasarkan Persamaan (11) adalah:

$$d_{hitung} = \frac{\sum_{i=1}^{33} (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^{33} \varepsilon_i^2} = \frac{109,3286}{50,60422} = 2,160464$$

Nilai d_{hitung} yang diperoleh yaitu sebesar 2,160464, dilihat dari tabel DW diketahui nilai $d_{L,0,025,33}$ sebesar 1,28. Apabila nilai $d_{hitung} = 2,160464 \geq d_{L,0,025,33} = 1,28$, sehingga dapat diambil keputusan gagal tolak H_0 yang artinya tidak terdapat korelasi antar residual atau tidak terjadi autokorelasi. Dengan demikian, hal ini menunjukkan bahwa asumsi residual independen terpenuhi.

3.8.3. Uji Distribusi Normal

Uji distribusi normal diperoleh dengan melakukan perhitungan residual pada Tabel 7 berdasarkan Persamaan (12) sebagai berikut :

$$D = \text{Sup}_x |F(Z_{24}) - F(X_{24})| = |0,616354 - 0,727273| = |-0,1109| \approx 0,111$$

Nilai D sebesar 0,111. Sementara itu, nilai D_α pada tabel KS sebesar 0,231. Karena nilai $D = 0,111 < D_{0,05} = 0,231$, dapat diambil keputusan gagal tolak H_0 yaitu residual berdistribusi normal. Hal ini menunjukkan

bahwa asumsi residual berdistribusi normal terpenuhi.

3.9. Interpretasi Model Terbaik

Berdasarkan pemilihan model terbaik pada Tabel 2 diperoleh model terbaik pada kombinasi knot (3, 3, 1, 1, 3, 3) dengan koefisien $\beta_{j(l)}$ pada Tabel 3, sehingga terbentuk model sebagai berikut.

$$\hat{Y}_i = -46,3853 - 0,791X_{1(i)} + 6,506(X_{1(i)} - 50,38)_+ - 9,152(X_{1(i)} - 52,29)_+ + 3,446(X_{1(i)} - 53,24)_+ - 2,226X_{2(i)} - 7,600(X_{2(i)} - 3,57)_+ + 46,903(X_{2(i)} - 3,81)_+ - 10,753X_{3(i)} + 36,393(X_{2(i)} - 3,08)_+ - 13,328(X_{3(i)} - 5,897)_+ + 8,482X_{4(i)} - 8,921(X_{4(i)} - 95,17)_+ - 0,091X_{5(i)} - 8,671(X_{5(i)} - 15,34)_+ + 22,270(X_{5(i)} - 16,22)_+ - 13,301(X_{5(i)} - 16,66)_+ - 2,352X_{6(i)} + 10,342(X_{6(i)} - 66,96)_+ - 1,256(X_{6(i)} - 67,42)_+ - 7,054(X_{6(i)} - 67,64)_+$$

Model di atas mempunyai empat variabel prediktor yang signifikan terhadap IPG di Provinsi Sumut. Variabel-variabel tersebut yaitu TPAK perempuan (X_1), RLS perempuan (X_3), AMH perempuan (X_4), dan AHH perempuan (X_6) dengan nilai R^2 sebesar 92,79%. Berikut adalah intepretasi model untuk setiap variabel yang signifikan terhadap IPG di Provinsi Sumut.

1. Hubungan antara TPAK Perempuan (X_1) terhadap IPG di Provinsi Sumut dengan asumsi apabila variabel X_2, X_3, X_4, X_5 , dan X_6 dianggap konstan atau tetap adalah sebagai berikut.

$$\hat{Y}_i = -0,791X_{1(i)} + 6,506(X_{1(i)} - 50,38)_+ - 9,152(X_{1(i)} - 52,28)_+ + 3,446(X_{1(i)} - 53,24)_+ = \begin{cases} -0,791X_{1(i)} & ; & X_{1(i)} < 50,38 \\ 5,715X_{1(i)} - 327,766 & ; & 50,38 \leq X_{1(i)} < 52,28 \\ -3,437X_{1(i)} + 150,737 & ; & 52,28 \leq X_{1(i)} < 53,24 \\ 0,009X_{1(i)} - 32,718 & ; & X_{1(i)} \geq 53,24 \end{cases}$$

Apabila kabupaten/kota di Provinsi Sumut dengan TPAK perempuan kurang dari $50,379 \approx 50$ orang dan TPAK perempuan bertambah 1 orang, maka IPG Provinsi Sumut akan turun 0,791%. Jika TPAK perempuan berjumlah antara $50,379 \approx 50$ orang sampai $52,284 \approx 52$ orang dan TPAK perempuan bertambah 1 orang, maka akan mengakibatkan IPG naik 5,715%. Apabila nilai TPAK perempuan berkisar antara $52,284 \approx 52$ orang sampai $53,237 \approx 53$

orang dan TPAK perempuan bertambah 1 orang, maka akan mengakibatkan IPG turun 3,437%. Apabila TPAK perempuan lebih dari sama dengan 53,237 \approx 53 orang dan TPAK perempuan bertambah 1 orang, maka IPG cenderung naik 0,009%.

2. Hubungan antara RLS Perempuan (X_3) terhadap IPG di Provinsi Sumut dengan asumsi apabila variabel X_1, X_2, X_4, X_5 , dan X_6 dianggap konstan atau tetap adalah sebagai berikut.

$$\hat{Y}_i = -10,753X_{3(i)} + 13,328(X_{3(i)} - 5,897)_+ \\ = \begin{cases} -10,753X_{3(i)} & ; X_{3(i)} < 5,897 \\ 2,575X_{3(i)} - 78,595 & ; X_{3(i)} \geq 5,897 \end{cases}$$

Apabila kabupaten/kota di Provinsi Sumut dengan RLS perempuan kurang dari 5,897 \approx 6 tahun dan RLS perempuan mengalami peningkatan 1 tahun, maka IPG Provinsi Sumut cenderung turun 10,753%. Sedangkan apabila RLS perempuan lebih dari sama dengan 5,897 \approx 6 tahun dan RLS perempuan bertambah 1 tahun, maka IPG cenderung naik 2,575%.

3. Hubungan antara AMH Perempuan (X_4) terhadap IPG di Provinsi Sumut dengan asumsi apabila variabel X_1, X_2, X_3, X_5 , dan X_6 dianggap konstan atau tetap adalah sebagai berikut.

$$\hat{Y}_i = 8,482X_{4(i)} - 8,921(X_{4(i)} - 95,167)_+ \\ = \begin{cases} 8,482X_{4(i)} & ; X_{4(i)} < 95,167 \\ -0,439X_{4(i)} + 848,985 & ; X_{4(i)} \geq 95,167 \end{cases}$$

Apabila kabupaten/kota di Provinsi Sumut dengan AMH perempuan kurang dari 95,167 \approx 95 persen dan setiap kenaikan 1 persen AMH perempuan, maka IPG Provinsi Sumut akan naik 8,482%. Apabila AMH perempuan lebih dari 95,167 \approx 95 persen dan setiap kenaikan 1 persen AMH perempuan, maka IPG akan turun 0,439%.

4. Hubungan antara AHH Perempuan (X_6) terhadap IPG di Provinsi Sumut dengan asumsi apabila variabel X_1, X_2, X_3, X_4 , dan X_5 dianggap konstan atau tetap adalah sebagai berikut.

$$\hat{Y}_i = -2,352X_{6(i)} + 10,342(X_{6(i)} - 66,96)_+ - \\ 1,256(X_{6(i)} - 67,42)_+ \\ - 7,054(X_{6(i)} - 67,64)_+$$

$$= \begin{cases} -2,352X_{6(i)} & ; X_{6(i)} < 66,96 \\ 7,989X_{6(i)} - 692,542 & ; 66,96 \leq X_{6(i)} < 67,42 \\ 6,733X_{6(i)} - 607,865 & ; 67,42 \leq X_{6(i)} < 67,64 \\ -0,32X_{6(i)} - 130,70 & ; X_{6(i)} \geq 67,64 \end{cases}$$

Apabila kabupaten/kota di Provinsi Sumut dengan AHH perempuan kurang dari 66,964 \approx 66 tahun dan AHH perempuan bertambah 1 tahun, maka akan mengakibatkan IPG Provinsi Sumut turun 2,352%. Jika AHH perempuan memiliki nilai berkisar antara 66,964 \approx 66 tahun hingga 67,418 \approx 67 tahun dan AHH perempuan bertambah 1 tahun, maka akan mengakibatkan IPG naik 7,989%. Apabila nilai AHH perempuan berkisar antara 67,418 \approx 67 tahun hingga 67,644 \approx 68 tahun dan AHH perempuan bertambah 1 tahun, maka akan mengakibatkan IPG naik 6,733%. Jika AHH perempuan lebih dari sama dengan 67,644 \approx 68 tahun dan AHH perempuan bertambah 1 tahun, maka akan mengakibatkan IPG turun 0,32%.

4. Kesimpulan

Pemodelan IPG di Provinsi Sumatera Utara tahun 2021 menggunakan metode regresi nonparametrik *spline truncated* menghasilkan model terbaik pada kombinasi knot (3, 3, 1, 1, 3, 3) dengan nilai GCV minimum sebesar 5,64129. Hasil uji signifikansi parameter model terbaik, terdapat 4 variabel yang berpengaruh signifikan terhadap IPG di Provinsi Sumatera Utara yaitu TPAK perempuan (X_1), RLS perempuan (X_3), AMH perempuan (X_4), dan AHH perempuan (X_6). Model terbaik yang diperoleh memiliki nilai R^2 sebesar 92,79% dan asumsi residual Identik, Independen, dan Distribusi Normal (IIDN) yang telah terpenuhi. Sehingga dapat disimpulkan bahwa perlu perhatian khusus pada variabel X_1, X_3, X_4 , dan X_6 untuk memperbaiki pembangunan *gender* di Provinsi Sumatera Utara dengan melakukan upaya peningkatan variabel X_1 dan X_3 , serta cenderung mempertahankan nilai variabel X_4 dan X_6 supaya *gap* antar *gender* tidak terlalu jauh dan tetap seimbang.

2. Saran

Penelitian selanjutnya sebaiknya bisa mengkaji beberapa faktor IPG lainnya,

menambah kemungkinan titik knot seperti 4, 5, dan 6 knot, dan menambah orde lain seperti orde 2 (*kuadratik*) dan orde 3 (*kubik*) agar bisa dikomparasikan untuk mendapatkan model yang terbaik. Selain itu, diharapkan dapat menggunakan metode lain seperti metode regresi nonparametrik *B-spline*.

6. Daftar Pustaka

- [1] Kemen PPPA, *Pembangunan Manusia Berbasis Gender*, 2022.
- [2] BPS Sumatera Utara, *Analisis Pembangunan Manusia Berbasis Gender Provinsi Sumatera Utara*, 2021.
- [3] N. Fajriyyah and I. Nyoman, "Pemodelan Indeks Pembangunan Gender dengan Pendekatan Regresi Nonparametrik Spline di Indonesia," vol. 4, no. 2, 2015.
- [4] A. Rahayu, R. Akbarita, and R. Narendra, "Analisis Pengaruh Gender Terhadap Indeks Pembangunan Gender Menggunakan Regresi Campuran Nonparametrik Spline Linier Truncated dan Fungsi Kernel," vol. 16, no. 1, pp. 33–39. 2022.
- [5] Andrianzah, *Regresi Nonparametrik Spline Truncated Untuk Memodelkan Inflasi di Indonesia*, Skripsi Program Studi Matematika, UIN Maulana Malik Ibrahim. Malang, 2023.
- [6] T. Lyche and K. Morken, *Spline Method Draft*, Norwegia: Department of Informatics Centre of Mathematics for Applications University of Oslo, <http://www.uio.no/studier/emner/matnat/ifi/INFMAT5340/v11/undervisningsmaterial/e/book.pdf>, 2008.
- [7] E. L. Montoya, N. Ulloa, and V. Miller, "A Simulation Study Comparing Knot Selection Methods With Equally Spaced Knots in a Penalized Regression Spline," vol. 3, no. 3, pp. 96–110, 2014.
- [8] N. C. S. Dewi and I. N. Budiantara, "Faktor-Faktor yang Mempengaruhi Angka Kecelakaan Lalu Lintas di Provinsi Jawa Timur Menggunakan Regresi Nonparametrik Spline Truncated," *J. Sains dan Seni ITS*, vol. 7, no. 2, 2018.
- [9] N. F. Rahayu, L. Wachidah, "Regresi Nonparametrik Spline untuk Memodelkan Faktor-faktor yang Mempengaruhi Indeks Pembangunan Gender (IPG) di Jawa Barat Tahun 2020," pp. 273–281, 2020.
- [10] D. Nikita and M. Ratna, "Pemodelan Indeks Pembangunan Gender di Pulau Kalimantan Menggunakan Metode Regresi Nonparametrik Spline Truncated," *J. Sains dan Seni ITS*, vol. 12, no. 1, 2023.
- [11] S. U. R. Sari, "Aplikasi Metode Regresi Nonparametrik Spline Multivariabel untuk Pemodelan Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Gender di Provinsi Jawa Barat," *Jurnal Statistika*, vol. 6, no. 2, 2018.
- [12] C. N. Sholicha, I. N. Budiantara, M. Ratna, D. Statistika, F. Matematika, and S. Data, "Regresi Nonparametrik Spline Truncated untuk Memodelkan Persentase Unmet Need di Kabupaten Gresik," *J. Sains dan Seni ITS*, vol. 7, no. 2, 2018.
- [13] A. Tri, R. Dani, and L. Ni, "Penerapan Keluarga Model Spline Truncated Polinomial pada Regresi Nonparametrik," *Inferensi*, vol. 5, no. March, pp. 37–44, 2022, doi: 10.12962/j27213862.v5i1.12537.
- [14] A. F. D. Rositawati and I. N. Budiantara, "Pemodelan Indeks Kebahagiaan Provinsi di Indonesia Menggunakan Regresi Nonparametrik Spline Truncated," *J. Sains dan Seni ITS*, vol. 8, no. 2, 2020, doi: 10.12962/j23373520.v8i2.45160.

