

## PEMODELAN FAKTOR-FAKTOR YANG MEMPENGARUHI INDEKS PEMBANGUNAN MANUSIA DI INDONESIA MENGGUNAKAN REGRESI SEMPARAMETRIK *SPLINE*

Megawati Nugroho Ningrum, Neva Satyahadewi, Naomi Nessyana Debararaja

### INTISARI

*Indeks Pembangunan Manusia (IPM) dibangun melalui tiga dimensi dasar yaitu umur panjang dan hidup sehat, mempunyai pengetahuan, dan kehidupan yang layak. Ketiga dimensi tersebut memiliki pengertian sangat luas karena terkait banyak faktor yang mempengaruhi IPM. Penelitian ini bertujuan untuk memodelkan dan menganalisis pengaruh faktor pengeluaran per kapita, kepadatan penduduk, penduduk miskin, dan angka partisipasi sekolah tingkat SMA/MA terhadap IPM di Indonesia pada tahun 2017. Metode statistika yang digunakan untuk menganalisis hubungan antara IPM dengan faktor-faktor yang mempengaruhi adalah regresi semiparametrik spline. Tahapan pengerjaan metode tersebut diawali dengan membuat scatterplot antara IPM dengan faktor-faktor yang mempengaruhi, memodelkan dengan menggunakan titik-titik knot, menguji signifikansi parameter dengan uji F dan uji t, melakukan uji asumsi-asumsi klasik, menghitung MSE dan koefisien determinasi terkoreksi  $R_{adj}^2$  kemudian melakukan interpretasi serta menarik kesimpulan. Model terbaik didapatkan dari titik knot optimal berdasarkan nilai Generalized Cross Validation (GCV) terkecil. Berdasarkan penelitian ini, model regresi semiparametrik spline terbaik adalah dengan menggunakan tiga knot dan empat variabel signifikan yaitu pengeluaran per kapita, kepadatan penduduk, penduduk miskin dan angka partisipasi sekolah tingkat SMA/MA. Model tersebut memiliki  $R_{adj}^2$  sebesar 86,23%.*

**Kata Kunci:** Indeks Pembangunan Manusia, Regresi Semiparametrik, Spline.

### PENDAHULUAN

Keberhasilan pembangunan khususnya pembangunan manusia dapat dinilai secara parsial dengan melihat seberapa besar permasalahan yang paling mendasar di masyarakat tersebut dapat teratasi. Dewasa ini persoalan mengenai capaian pembangunan manusia telah menjadi perhatian para penyelenggara pemerintahan. Oleh karena itu, Badan Perserikatan Bangsa-Bangsa (PBB) menetapkan suatu ukuran standar pembangunan manusia yaitu Indeks Pembangunan Manusia atau *Human Development Index (HDI)*.

Pada tahun 2017 IPM Indonesia mencapai 70,81% meningkat 0,90% dari tahun sebelumnya. IPM tertinggi dicapai oleh Provinsi DKI Jakarta dengan IPM sebesar 80,06%, sedangkan capaian terendah diduduki oleh Provinsi Papua dengan IPM sebesar 59,09%. Oleh karena itu, diperlukan suatu penelitian yang bertujuan untuk memodelkan nilai IPM terhadap faktor-faktor yang diduga mempengaruhinya.

Salah satu metode yang dapat digunakan untuk menyelesaikan masalah pemodelan adalah analisis regresi. Analisis regresi merupakan suatu metode statistik yang digunakan untuk mengetahui pola hubungan antara variabel prediktor ( $x$ ) dan variabel respon ( $y$ ) [1]. Pola hubungan antar variabel prediktor dan variabel respon dapat didekati dengan tiga cara yaitu regresi parametrik, regresi nonparametrik, dan regresi semiparametrik.

Regresi semiparametrik merupakan metode statistika yang digunakan apabila sebagian bentuk kurva regresi diasumsikan diketahui dan sebagian lagi tidak diketahui bentuk kurva regresinya. Salah satu regresi semiparametrik yang sering digunakan yaitu regresi *spline*. Regresi *spline* merupakan model yang sangat baik dalam memodelkan data yang memiliki pola yang berubah-ubah pada sub-sub interval tertentu. Salah satu yang perlu diperhatikan dalam membentuk regresi *spline* ialah

menentukan banyaknya titik knot serta lokasi titik knot. Titik knot merupakan titik yang menunjukkan perubahan pola data pada sub-sub interval [2]. Salah satu metode yang sering digunakan yakni *truncated*. *Truncated* memiliki kelebihan mengikuti pola data sesuai pergerakannya dengan adanya titik-titik knot [1].

Adapun tujuan dari penelitian ini adalah memodelkan nilai IPM dengan menggunakan metode regresi semiparametrik *spline* serta menganalisis karakteristik faktor-faktor yang diduga berpengaruh terhadap nilai IPM pada 34 Provinsi di Indonesia pada tahun 2017. Batasan masalah pada penelitian ini yaitu titik knot yang digunakan adalah tiga titik knot, model terbaik diperoleh dari titik knot optimal berdasarkan nilai GCV terkecil, dan model *spline* yang digunakan adalah model *spline truncated* linier. Dalam menentukan kebaikan model yakni dengan menggunakan koefisien determinasi terkoreksi ( $R_{adj}^2$ ) dan *Mean Square Error* (MSE).

Penelitian ini berupa studi literatur yang dimulai dengan mempelajari karakteristik data, menentukan pola kurva regresi, banyaknya titik knot, dan nilai titik knot. Selanjutnya menentukan nilai titik knot berdasarkan nilai GCV terkecil dan banyaknya titik knot optimal. Setelah itu menguji signifikansi parameter dan uji asumsi klasik model regresi *spline* kemudian menghitung  $R_{adj}^2$  dan MSE sebagai kriteria kebaikan model pada model regresi semiparametrik *spline* serta menginterpretasikan model dan menarik kesimpulan.

### REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED*

Regresi semiparametrik merupakan salah satu metode statistika yang digunakan untuk mengetahui pola hubungan antara variabel respon dan variabel prediktor dimana sebagian diketahui bentuk polanya dan sebagian tidak diketahui bentuknya. Model umum regresi semiparametrik adalah sebagai berikut [3]:

$$y_i = \mathbf{x}_i' \boldsymbol{\beta} + f(t_i) + \varepsilon_i; i = 1, 2, \dots, n \quad (1)$$

dimana variabel respon  $y_i$  berhubungan parametrik dengan variabel prediktor  $\mathbf{x}_i = (x_{i1}, x_{i2}, \dots, x_{im})'$  dan berhubungan nonparametrik dengan variabel prediktor  $t_i$ ,  $f(t_i)$  merupakan fungsi *spline* dengan berorde  $p$  dengan titik knot  $k_1, k_2, \dots, k_r$ .

*Spline* merupakan analisis regresi yang mampu mengestimasi data yang tidak memiliki pola tertentu dan memiliki kecenderungan mencari sendiri estimasi data dari pola yang terbentuk. Salah satu kelebihan pendekatan *spline* adalah model ini memiliki kemampuan sangat baik untuk menangani data yang perilakunya berubah-ubah pada sub-sub interval tertentu. Kelebihan ini terjadi dikarenakan di dalam *spline* terdapat titik knot. Titik knot merupakan titik perpaduan bersama yang menunjukkan terjadinya perubahan perilaku pola data.

Secara umum,  $f(t_i)$  merupakan fungsi *spline* berorde  $p$  dengan titik knot  $k_1, k_2, k_3, \dots, k_r$  dan dapat dituliskan persamaannya sebagai berikut [3]:

$$f(t_i) = \sum_{j=0}^p \gamma_j t_i^j + \sum_{l=1}^r \gamma_{p+l} (t_i - k_l)_+^p \quad (2)$$

dimana  $\gamma$  adalah parameter model,  $p$  adalah orde *spline*,  $k$  adalah banyak knot *spline*, dan  $(t_i - k_l)_+^p$  merupakan fungsi *truncated* (potongan) yang dapat dijabarkan sebagai berikut [3]:

$$(t_i - k_l)_+^p = \begin{cases} (t_i - k_l)^p, & t_i \geq k_l \\ 0, & t_i < k_l \end{cases} \quad (3)$$

Apabila persamaan (2) disubstitusikan ke dalam persamaan (1), diperoleh persamaan regresi sebagai berikut [3]:

$$y_i = \sum_{j=0}^p \gamma_j t_i^j + \sum_{l=1}^r \gamma_{(p+l)} (t_i - k_l)_+^p + \varepsilon_i; i = 1, 2, \dots, n \quad (4)$$

**PEMILIHAN TITIK KNOT OPTIMAL**

Pemilihan titik knot yang optimal merupakan hal penting dalam regresi semiparametrik *spline*. Metode yang digunakan dalam memilih titik knot optimal pada penelitian ini adalah GCV. Model *spline* terbaik dengan titik knot optimal yang diperoleh dari nilai GCV terkecil [4]. Bentuk umum dari metode GCV dapat ditulis sebagai berikut [5]:

$$GCV(k) = \frac{MSE(k)}{\left[ n^{-1} \text{trace}(\mathbf{I} - \mathbf{A}(k)) \right]^2} \quad (5)$$

dimana  $\mathbf{I}$  adalah matriks identitas,  $n$  adalah jumlah pengamatan,  $k$  merupakan titik-titik knot, *trace* merupakan hasil penjumlahan diagonal utama matriks,  $MSE(k) = n^{-1} \sum_{i=1}^n (y_i - \hat{y})^2$ , dan  $\mathbf{A}(k) = \mathbf{x}(k)(\mathbf{x}(k)' \mathbf{x}(k))^{-1} \mathbf{x}(k)'$ .

**PENGUJIAN HIPOTESIS TERHADAP PARAMETER REGRESI**

Dalam melakukan pengujian hipotesis terdiri dari dua tahapan, yaitu pengujian secara serentak (simultan) dan pengujian secara parsial (individu). Berikut ini tahapan dalam pengujian hipotesis parameter regresi:

1. Pengujian Parameter Secara Serentak (Uji F)

Hipotesis yang digunakan yakni:

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_k = 0$$

$$H_1 : \text{minimal satu } \beta_i \neq 0$$

dimana  $i = 1, 2, \dots, k$

Statistik uji yang digunakan untuk uji F sebagai berikut [6]:

$$F = \frac{JKR / (k - 1)}{JKS / (n - k)}$$

dengan JKR adalah jumlah kuadrat regresi, JKS adalah jumlah kuadrat sisaan,  $k$  adalah jumlah variabel prediktor, dan  $n$  adalah jumlah pengamatan. Jika  $F_{statistik} > F_{(\alpha, k-1, n-k)}$  dengan  $\alpha$  adalah 0,05 maka keputusan tolak  $H_0$  dan dapat disimpulkan bahwa semua variabel prediktor secara serentak berpengaruh signifikan terhadap variabel respon.

2. Pengujian Parameter Secara Parsial (Uji t)

Hipotesis yang digunakan yakni:

$$H_0 : \beta_i = 0, \quad i = 1, 2, \dots, k$$

$$H_1 : \beta_i \neq 0, \quad i = 1, 2, \dots, k$$

Statistik uji yang digunakan untuk uji t sebagai berikut [6]:

$$t = \frac{b_i - \beta_i}{SE_{b_i}}$$

merupakan suatu variabel acak berdistribusi  $t$  dengan  $\beta_i$  adalah *slope*,  $b_i$  adalah rata-rata distribusi sampling dari  $\beta_i$  dan  $SE_{b_i}$  adalah penduga *standard error* dari  $b_i$ . Jika hipotesis  $H_0$  benar, artinya  $\beta_i = 0$ , maka nilai statistik uji  $t$  berubah menjadi

$$t = \frac{b_i}{SE_{b_i}}$$

Nilai statistik  $t$  akan mengikuti distribusi  $t$  dengan derajat bebas sebesar  $n - k$ . Jika  $t_{hit} > t_{(\alpha/2, n-k)}$  maka keputusan tolak  $H_0$  dan dapat disimpulkan bahwa variabel prediktor secara parsial mempunyai pengaruh signifikan terhadap variabel respon.

### UJI ASUMSI KLASIK

Berikut ini asumsi klasik yang harus dipenuhi dalam regresi *spline* adalah sebagai berikut [7]:

#### 1. Uji Normalitas

Uji normalitas digunakan untuk mengetahui nilai residual terstandarisasi pada model regresi berdistribusi normal atau tidak. Uji yang digunakan adalah *Kolmogorov-Smirnov*. Hipotesis yang digunakan yakni:

$H_0$  : *error* berdistribusi normal

$H_1$  : *error* tidak berdistribusi normal

Statistik uji yang digunakan dari uji *Kolmogorov-Smirnov* adalah sebagai berikut:

$$KS_{hitung} = \max |F(z_i) - S(z_i)| \quad (6)$$

dimana  $z_i = \frac{x_i - \bar{x}}{s}$  adalah  $\bar{x}$  dan  $s$  masing-masing merupakan rata-rata dan simpangan baku

sampel,  $F(z_i) = P(z \leq z_i)$  dan  $S(z_i) = \frac{z_1, z_2, \dots, z_n \leq z_i}{n}$ .

#### 2. Uji Autokorelasi

Uji autokorelasi bertujuan untuk mengetahui apakah ada korelasi antara anggota serangkaian data observasi yang diuraikan menurut waktu (*time series*) atau ruang (*cross section*) [8]. Uji yang digunakan ialah *Durbin-Watson*. Hipotesis yang digunakan yakni:

$H_0$  : Tidak ada autokorelasi

$H_1$  : Ada autokorelasi

Berikut statistik uji yang digunakan [8]:

$$d = \frac{\sum_{i=2}^n (\varepsilon_i - \varepsilon_{i-1})^2}{\sum_{i=1}^n \varepsilon_i^2} \quad (7)$$

dimana  $d$  adalah nilai *Durbin-Watson*,  $\varepsilon_i$  adalah nilai *error* dan  $\varepsilon_{i-1}$  adalah nilai *error* pada satu periode sebelumnya. Jika nilai  $d < d_L$  maka  $H_0$  ditolak yang berarti adanya autokorelasi dalam model regresi.

#### 3. Uji Heteroskedastisitas

Uji heteroskedastisitas digunakan untuk menguji ada atau tidaknya perbedaan varians dari nilai residual suatu periode pengamatan ke periode pengamatan yang lain. Untuk mendeteksi adanya heteroskedastisitas menggunakan uji *Glejser*. Hipotesis yang digunakan yakni:

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

$$H_1 : \text{paling sedikit ada satu } \sigma_1^2 \neq \sigma^2$$

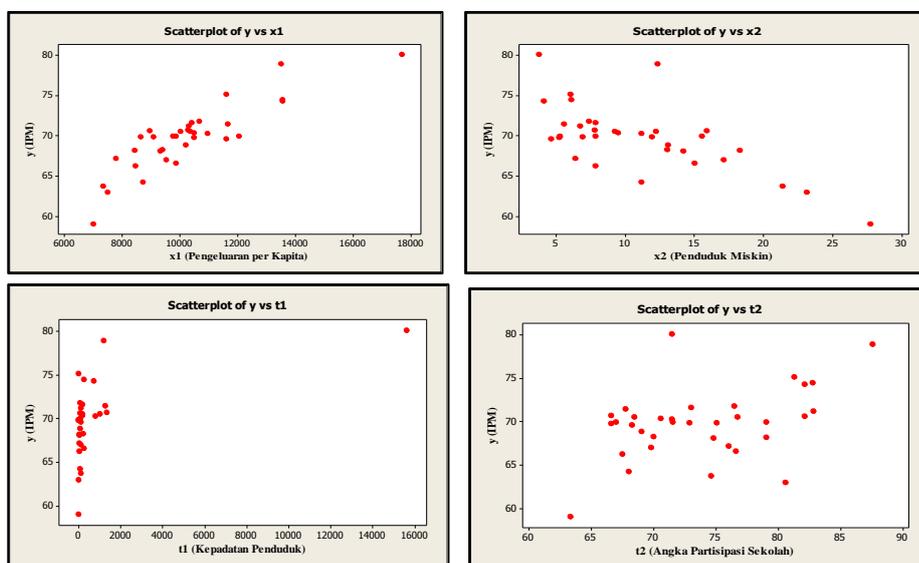
Statistik uji dari uji *Glejser* adalah sebagai berikut [9]:

$$F_{hitung} = \frac{\sum_{i=1}^n (|\hat{\varepsilon}_i| - |\hat{\varepsilon}_i|)^2 / k-1}{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|) / n-k} \tag{8}$$

dimana  $|\varepsilon_i|$  adalah nilai *error* mutlak. Jika  $\beta$  signifikan maka terdapat pengaruh variabel prediktor terhadap nilai residual mutlak sehingga dinyatakan bahwa terdapat gejala heteroskedastisitas [8].

### STUDI KASUS

Data dalam penelitian ini merupakan data sekunder yang diambil dari Badan Pusat Statistik (BPS) Nasional tahun 2017. Penelitian ini menganalisis data dari 34 provinsi di Indonesia dengan variabel-variabel yang digunakan dalam penelitian ini adalah variabel respon ( $y$ ) berupa Indeks Pembangunan Manusia (%). Kemudian terdapat pula 4 variabel prediktor ( $x$ ) yaitu pengeluaran per kapita ( $x_1$ ), penduduk miskin ( $x_2$ ), kepadatan penduduk ( $t_1$ ), dan angka partisipasi sekolah tingkat SMA/MA ( $t_2$ ). Langkah awal dalam penentuan model hubungan antara IPM dengan masing-masing variabel yaitu membuat *scatter plot* antara variabel respon dan masing-masing variabel prediktor. Berikut merupakan *scatter plot* IPM antarvariabel:



**Gambar 1.** *Scatterplot* Antarvariabel Prediktor dengan Variabel Respon

Berdasarkan Gambar 1, terdapat dua variabel komponen parametrik yaitu pengeluaran per kapita ( $x_1$ ) dan penduduk miskin ( $x_2$ ) serta dua variabel komponen nonparametrik yaitu kepadatan penduduk ( $t_1$ ) dan angka partisipasi sekolah tingkat SMA/MA ( $t_2$ ). Titik knot optimal ditentukan berdasarkan nilai GCV yang terkecil. Nilai GCV terkecil pada pemilihan titik knot optimal dengan satu titik knot, dua titik knot, dan tiga titik knot. Berikut nilai GCV terkecil dan jumlah titik knot dari masing-masing variabel prediktor terhadap variabel respon.

**Tabel 1** Nilai GCV Terkecil Setiap Pemilihan Titik Knot Optimal

Model	GCV
1 Knot	2,48978
2 Knot	2,04697
<b>3 Knot</b>	<b>1,89169</b>

Sumber: Pengolahan menggunakan *Software R-3.5.1*

Tabel 1 menunjukkan nilai GCV terkecil dimiliki oleh model dengan 3 titik knot yaitu sebesar 1,89169. Berdasarkan hasil nilai GCV terkecil, diperoleh model terbaik yaitu model regresi semiparametrik *spline* dengan 3 titik knot. Berikut hasil estimasi parameter model regresi semiparametrik *spline* terbaik.

$$\begin{aligned} \hat{y}_i = & 0,040929 + 0,001411 x_{i1} - 0,195368 x_{i2} - 0,007931 t_{i1} \\ & + 0,016900 (t_{i1} - 547,448)_+ - 0,012424 (t_{i1} - 2701,24)_+ \\ & + 0,865451 (t_{i1} - 4316,59)_+ - 1,068562 t_{i2} + 1,685176 (t_{i2} - 64,1866)_+ - 1,414836 (t_{i2} - 67,5328)_+ \end{aligned}$$

### UJI SIGNIFIKANSI PARAMETER

Berdasarkan hasil Analisis variansi (Anava), berikut ini hasil pengujian parameter secara serentak:

**Tabel 2** Uji Serentak

Sumber	Df	SS	MS	FHit	P-Value	Keputusan
Regresi	9	512,86	56,985	39,3298	3,2395	Tolak $H_0$
Error	24	34,774	1,4489			
Total	33	547,64				

Sumber: Pengolahan menggunakan *Software R-3.5.1*

Berdasarkan hasil Anava pada Tabel 2 dapat diketahui p-value 3,2395. Menggunakan taraf signifikan ( $\alpha$ ) sebesar 0,05, sehingga dalam hal ini dapat disimpulkan bahwa  $H_0$  ditolak, maka terdapat pengaruh antara variabel prediktor terhadap variabel respon secara serentak. Untuk mengetahui parameter yang berpengaruh dan signifikan terhadap variabel respon dilakukan pengujian secara parsial. Berikut hasil pengujian terhadap parameter pada setiap variabel prediktor:

**Tabel 3** Uji Parsial

Variabel	Parameter	Estimasi	Standar Error	P-Value	Keputusan
-	$\beta_0$	0,040929	0,013419	0,004487	Tolak $H_0$
$x_1$	$\beta_1$	0,001411	0,00024	1,40680	Tolak $H_0$
$x_2$	$\beta_2$	-0,195368	0,059656	0,002486	Tolak $H_0$
$t_1$	$\gamma_{11}$	-0,007931	0,002845	0,008729	Tolak $H_0$
	$\gamma_{21}$	0,016900	0,004818	0,001132	Tolak $H_0$
	$\gamma_{31}$	-0,012424	0,002987	0,000213	Tolak $H_0$
	$\gamma_{41}$	0,865451	0,050004	4,031143	Tolak $H_0$
$t_2$	$\gamma_{12}$	-1,068562	0,529168	0,051636	Gagal tolak $H_0$
	$\gamma_{22}$	1,685176	0,780058	0,038107	Tolak $H_0$
	$\gamma_{32}$	-1,414836	0,427544	0,002269	Tolak $H_0$

Sumber: Pengolahan menggunakan *Software R-3.5.1*

Tabel 3 menunjukkan bahwa dari keseluruhan parameter model yang diperoleh ternyata parameter  $\beta_1$  pada  $x_1$ , parameter  $\beta_2$  pada  $x_2$ , parameter  $\gamma_{11}, \gamma_{21}, \gamma_{31}$  dan  $\gamma_{41}$  pada  $t_1$ , serta parameter  $\gamma_{22}$  dan  $\gamma_{32}$  pada  $t_2$  yang berpengaruh signifikan terhadap model regresi *spline* linier.

**PENGUJIAN ASUMSI KLASIK**

Berdasarkan model yang telah diperoleh, untuk melihat kelayakan model dilakukan uji asumsi klasik. Untuk mengetahui normalitas *error* pada model menggunakan uji *Kolmogorov-Smirnov*. Dari hasil yang diperoleh, nilai *p-value* > 0,8502 dengan menggunakan signifikansi  $\alpha$  sebesar 0,05, uji *Kolmogorov-Smirnov* memberikan keputusan gagal tolak  $H_0$  karena *p-value* >  $\alpha$  sehingga *error* pada model berdistribusi normal.

Selanjutnya dilakukan pengujian asumsi klasik autokorelasi untuk mengetahui ada atau tidaknya korelasi antar residual dengan uji *Durbin-Watson*. Dari hasil yang diperoleh nilai *Durbin-Watson* sebesar 1,510669. Nilai dL dan dU untuk  $k = 4$  dan  $n = 34$  adalah 1,2078 dan 1,7277 sehingga  $dU < d < 4 - dU$ .

Selanjutnya dilakukan pengujian asumsi klasik heteroskedastisitas. Uji ini dilakukan dengan melihat hasil uji *Glejser*. Berikut hasil uji *Glejser*.

**Tabel 4 Uji Glejser**

Sumber	db	SS	MS	Fhit	P-Value	Keputusan
Regresi	10	4,728	0,5254	2,1163	0,0069	Gagal Tolak $H_0$
Error	20	5,958	0,2483			
Total	30	10,68				

Sumber: Pengolahan menggunakan *Software R-3.5.1*

Berdasarkan Tabel 4 diperoleh *p-value* < 0,05 hal ini menunjukkan bahwa  $H_0$  tidak ditolak sehingga tidak terdapat heteroskedastisitas atau asumsi klasik terpenuhi.

**INTERPRETASI MODEL REGRESI SEMIPARAMETRIK SPLINE**

Berdasarkan hasil estimasi parameter yang diperoleh, maka model dari regresi semiparametrik *spline* pada indeks pembangunan manusia dengan setiap variabel prediktor dapat dituliskan dalam persamaan regresi semiparametrik *spline* sebagai berikut:

$$\hat{y}_i = 0,040929 + 0,001411 x_{i1} - 0,195368 x_{i2} - 0,007931 t_{i1} + 0,016900(t_{i1} - 547,448)_+ - 0,012424(t_{i1} - 2701,24)_+ + 0,865451(t_{i1} - 4316,59)_+ - 1,068562 t_{i2} + 1,685176(t_{i2} - 64,1866)_+ - 1,414836(t_{i2} - 67,5328)_+$$

Berdasarkan hasil estimasi regresi semiparametrik *spline* diperoleh nilai  $R^2_{adj}$  sebesar 86,23%. Angka tersebut berarti bahwa model terbaik yang didapatkan mampu menjelaskan variasi variabel respon berupa IPM Indonesia sebesar 86,23%. Sedangkan sisa koefisien determinasi terkoreksi 13,77% dijelaskan oleh variabel prediktor lain yang tidak termasuk ke dalam model.

**KESIMPULAN**

Model regresi semiparametrik *spline* terbaik untuk memodelkan IPM di Indonesia adalah dengan tiga titik knot. Dimana titik knot berjumlah tiga untuk variabel angka partisipasi sekolah tingkat SMA/MA dan berjumlah tiga pula untuk variabel kepadatan penduduk. Berikut adalah model akhir yang diperoleh.

$$\begin{aligned} \hat{y}_i = & 0,040929 + 0,001411 x_{i1} - 0,195368 x_{i2} - 0,007931 t_{i1} \\ & + 0,016900 (t_{i1} - 547,448)_+ - 0,012424 (t_{i1} - 2701,24)_+ \\ & + 0,865451 (t_{i1} - 4316,59)_+ - 1,068562 t_{i2} + 1,685176 (t_{i2} - 64,1866)_+ - 1,414836 (t_{i2} - 67,5328)_+ \end{aligned}$$

Dimana  $x_1$  adalah pengeluaran per kapita,  $x_2$  adalah penduduk miskin,  $t_1$  adalah kepadatan penduduk, dan  $t_2$  adalah angka partisipasi sekolah tingkat SMA/MA. Model menunjukkan bahwa semua variabel berpengaruh signifikan terhadap IPM di Indonesia. Terdapat 1 provinsi yang terkategori IPM kelompok rendah, 18 provinsi yang terkategori dalam IPM kelompok sedang, terdapat 14 provinsi di Indonesia yang terkategori dalam IPM kelompok tinggi, dan terdapat 1 provinsi yang dikategorikan ke dalam IPM kelompok sangat tinggi. Pada tahun 2017 Indonesia memiliki nilai IPM sebesar 70,81. Nilai IPM terendah yaitu Provinsi Papua yaitu sebesar 59,09. Nilai IPM tertinggi yaitu Provinsi DKI Jakarta dengan nilai sebesar 80,06. Berdasarkan *scatter plot* antara IPM di Indonesia dengan setiap variabel yang diduga mempengaruhinya, terdapat dua variabel komponen parametrik (pengeluaran per kapita dan penduduk miskin) dan dua komponen nonparametrik (kepadatan penduduk dan angka partisipasi sekolah tingkat SMA/MA).

#### DAFTAR PUSTAKA

- [1]. Astuti EP. Pemilihan Titik Knot Optimal Dalam Regresi Nonparamterik *Spline Truncated* Pada Data Longitudinal. 2017. Tesis: Institut Teknologi Sepuluh November, Surabaya.
- [2]. Fadhilah NK, Suparti, Tarno. Pemodelan Regresi *Spline Truncated* Untuk Data Longitudinal (Studi Kasus: Harga Saham Bulanan pada Kelompok Saham Perbankan Periode Januari 2009 – Desember 2015), *Jurnal Gaussian*. 2016: 5(3);447–454.
- [3]. Pratiwi DA, Budiantara IB, Wibowo W. Pendekatan Regresi Semiparametrik Spline Untuk Memodelkan Rata-Rata Umur Kawin Pertama (UKP) di Jawa Timur, *Jurnal Sains dan Seni ITS*. 2017: 6(1);129-136.
- [4]. Arfan N, Budiantara IB. Pendekatan *Spline* untuk Estimasi Kurva Regresi Nonparametrik (Studi Kasus: Data Angka Kematian Maternal di Jawa Timur, *Jurnal Sains dan Seni Pomits*. 2014: 3(1);13–17.
- [5]. Sugiantari PA, Budiantara IB. Analisis Faktor-Faktor yang Mempengaruhi Angka Harapan Hidup di Jawa Timur Menggunakan Regresi Semiparametrik *Spline*, *Jurnal Sains dan Seni Pomits*. 2013: 2(1);37 - 41.
- [6]. Kusnandar D. *Metode Statistik dan Aplikasinya dengan MINITAB dan Excel*. Yogyakarta: Madyan Press;2004
- [7]. Gunawan I. *Pengantar Statistika Inferensial*. Malang: Rajawali Press;2015.
- [8]. Suliyanto. *Ekonometrika Terapan Teori dan Aplikasi dengan SPSS*. Yogyakarta: CV. Andi Offser;2011.
- [9]. Kurniasari W, Kusnandar D, Sulistianingsih E. Estimasi Parameter Regresi *Spline* Dengan Metode *Penalized Spline*, *Bimaster*. 2019:8(2);175-184.

MEGAWATI NUGROHO NINGRUM : Jurusan Matematika FMIPA Untan, Pontianak,  
mningrum@gmail.com  
NEVA SATYAHADEWI : Jurusan Matematika FMIPA Untan, Pontianak,  
neva.satya@math.untan.ac.id  
NAOMI NESSYANA DEBATARAJA : Jurusan Matematika FMIPA Untan, Pontianak,  
naominessyana@math.untan.ac.id

---