

PEMODELAN INDEKS PEMBANGUNAN MANUSIA DI KEPULAUAN MALUKU DENGAN PENDEKATAN ESTIMASI INTERVAL PARAMETER MODEL REGRESI SEMIPARAMETRIK *SPLINE TRUNCATED*

The Modelling Human Development Index in The Maluku Islands with Approaches Estimation Interval of Parameters Models Semiparametric Spline Truncated Linear Regression

Samsul Bahri Loklomin*

Program Studi Matematika Fakultas Ilmu Alam dan Teknologi Rekayasa, Universitas Halmahera
Jalan Raya Wari Ino, Tobelo, Maluku Utara, Indonesia

e-mail: samsul.bahriloklomin@gmail.com
Corresponding Author*

Abstrak

Indeks Pembangunan Manusia merupakan indeks komposit yang dihitung dari indeks kesehatan, pendidikan dan standar hidup layak. Penelitian ini bertujuan memodelkan faktor-faktor yang berpengaruh terhadap Indeks Pembangunan Manusia di Kepulauan Maluku. Pemodelan menggunakan pendekatan Regresi Semiparametrik Spline *Truncated* Linier untuk melihat faktor-faktor yang signifikan terhadap Indeks Pembangunan Manusia dengan metode interval konfidensi. Hasil penelitian ini menunjukkan bahwa semua variabel prediktor berpengaruh signifikan terhadap Indeks Pembangunan Manusia. Model yang terbaik dalam penelitian ini adalah dengan tiga titik knot dengan R-Square 99,97% dan nilai GCV minimum 0,346.

Kata Kunci: Regresi semiparametrik, spline truncated, estimasi interval, indeks pembangunan manusia

Abstract

The Human Development Index is a composite index that is calculated from the index of health, education and decent living standards. This research aims to model the factors that affect the Human Development Index of the Maluku Islands. This Modelling uses the Semiparametric Spline Truncated Linear Regression approach to see significant factors on the Human Development Index using the confidence interval method. The results of this study indicate that all predictor variables have a significant influence on the Human Development Index. The best model in this study is with three knots with R-Square 99,97% and a minimum GCV value of 0.346.

Keywords: Semiparametric regression, spline truncated, estimation of interval, human development index

Diterima : 13 Maret 2019

Direvisi: 20 Juni 2019

Disetujui: 25 Juni 2019

Copyright © 2019 : Author, Published by Barekeng: Jurnal ilmu matematika dan terapan

1. PENDAHULUAN

Tujuan utama pembangunan adalah menciptakan lingkungan yang memungkinkan rakyat menikmati umur panjang, kesehatan yang memadai, dan menjalankan kehidupan yang produktif. Indeks Pembangunan Manusia (IPM) adalah indikator yang dapat digunakan untuk mengukur hasil pembangunan. Paradigma pembangunan yang sedang berkembang saat ini adalah pertumbuhan ekonomi yang diukur dengan pembangunan manusia yang dilihat dengan tingkat kualitas hidup manusia di setiap negara [2]. Di Indonesia, dari 34 provinsi terdapat 2 provinsi di kepulauan Maluku yang masih tergolong dalam kategori IPM yang sedang jika dibandingkan dengan provinsi-provinsi lainnya di Indonesia, provinsi Maluku menempati posisi ke-24 [3] dan provinsi Maluku Utara menempati posisi ke-27 secara nasional [4]. Jika dilihat dari sejarah terbentuknya kedua provinsi tersebut sangat bertolak belakang dengan kehidupan pembangunan terutama permasalahan IPM, sebab kedua provinsi tersebut merupakan provinsi-provinsi yang sudah lama terbentuk jika dibandingkan dengan beberapa provinsi baru yang mempunyai IPM lebih baik.

Permasalahan IPM merupakan salah satu permasalahan sosial kependudukan yang dapat dimodelkan dengan regresi. Pada regresi tidak semua variabel penjelas dapat didekati dengan pendekatan parametrik. Oleh karena tidak adanya informasi tentang bentuk hubungan variabel penjelas tersebut dengan variabel responnya, sehingga harus digunakan pendekatan nonparametrik. Dengan menggabungkan dua pendekatan tersebut dalam suatu pendekatan regresi akan didapatkan suatu model semiparametrik. Model regresi semiparametrik digunakan apabila variabel respon memiliki pola hubungan tertentu dengan salah satu atau beberapa variabel prediktor, akan tetapi dengan variabel prediktor yang lain tidak diketahui bentuk pola hubungan antara variabel respon dan variabel prediktor.

Pada beberapa penelitian, variabel respon dapat memiliki hubungan linier dengan sebagian variabel prediktor, dan memiliki pola hubungan yang tidak diketahui bentuk kurva regresinya dengan sebagian variabel prediktor yang lain. Pada kasus seperti ini, lebih disarankan untuk menggunakan pendekatan regresi semiparametrik[5]. Model pendekatan regresi semiparametrik mempunyai latar belakang tersendiri dalam mengestimasi. Beberapa metode estimasi yang digunakan dalam regresi semiparametrik diantaranya adalah Kernel, Fourier, dan Spline. Salah satu model regresi dengan pendekatan regresi semiparametrik yang sering digunakan untuk melakukan estimasi terhadap kurva regresi adalah Spline *Truncated*[1]. Spline *Truncated* merupakan model yang mempunyai itererpretasi statistik dan interpretasi visual khusus. Kelebihan dari model Spline *Truncated* adalah jika data yang memiliki pola berubah-ubah pada sub-sub interval tertentu maka dapat dimodelkan dengan Spline *Truncated*[10]. Ada beberapa metode yang dapat digunakan untuk memilih titik knot optimal dalam estimator Spline *Truncated*, antara lain metode *Cross Validation* (CV), *Unbiased Risk* (UBR), *Generalized Maximum Likelihood* (GML), dan *Generalized Cross Validation* (GCV). Salah satu metode yang kerap digunakan dalam pemilihan titik knot optimal adalah GCV. Metode GCV memberikan beberapa kelebihan dibandingkan metode lainnya, diantaranya memiliki sifat optimal asimptotik, invarian terhadap transformasi dan dalam perhitungannya varians populasi tidak perlu diketahui. Berdasarkan uraian tersebut, permasalahan dalam penelitian ini bagaimana memodelkan IPM dengan pendekatan estimasi interval parameter model regresi semiparametrik Spline *Truncated*. Batasan masalah pada penelitian ini adalah pemilihan titik knot optimal menggunakan GCV satu knot, dua knot, tiga knot, dan kombinasi knot[9].

2. METODE PENELITIAN

2.1 Sumber Data

Penelitian ini menggunakan data sekunder tahun 2015 dari publikasi Badan Pusat Statistik Provinsi Maluku dan Maluku Utara dengan unit observasi meliputi 21 Kabupaten/Kota di Kepulauan Maluku.

2.2 Variabel Penelitian

Variabel penelitian yang digunakan dalam penelitian ini berupa data kategorik yaitu IPM menurut Kabupaten/Kota di Kepulauan Maluku tahun 2015 dan variabel-variabel yang diduga mempengaruhi IPM.

2.4 Langkah-langkah Penelitian

Secara umum langkah-langkah yang akan dilakukan dalam penelitian ini yaitu :

1. Membuat *scatter plot* antara variabel respon dengan variabel prediktor
2. Menentukan variabel komponen parametrik berdasarkan pola data antara variabel respon dan variabel prediktor
3. Menentukan variabel komponen nonparametrik berdasarkan pola data antara variabel respon dan variabel prediktor
4. Memodelkan data menggunakan regresi semiparametrik spline *truncated* dimana spline yang digunakan adalah satu knot, dua knot, tiga knot, dan kombinasi knot.
5. Memilih titik knot optimal dengan metode GCV
6. Menetapkan model terbaik dari nilai GCV terkecil
7. Memeriksa asumsi residual yang diisyaratkan yaitu sumsi identik, independen dan normalitas.
8. Menghitung MSE dan R^2 sebagai bagian dari kriteria kebaikan model
9. Menghitung interval konfidensi terpendek untuk parameter model regresi semiparametrik spline *truncated*
10. Mengambil kesimpulan

3. HASIL DAN PEMBAHASAN

3.1. Spline Truncate dalam Regresi Semiparametrik

Diberikan data berpasangan pengamatan independen (x_i, z_i, y_i) , $i = 1, 2, \dots, n$, dimana y_i adalah variabel respon, sedangkan x_i adalah variabel prediktor yang mengikuti pola parametrik dan z_i mengikuti pola nonparametrik. Pola antara (x_i, z_i, y_i) dapat dinyatakan dalam persamaan model regresi[4]:

$$y_i = \sum_{j=1}^p f_j(x_{ij}) + \sum_{k=1}^q g_k(z_{ik}) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (1)$$

dengan $f_j(x_{ij})$ adalah komponen parametrik dan $g_k(z_{ik})$ adalah komponen nonparametrik. Kurva komponen parametrik $f_j(x_{ij})$ dimodelkan dengan fungsi linier. Kurva komponen nonparametrik $g_k(z_{ik})$ dimodelkan dengan fungsi Spline linier dengan knot K_1, K_2, \dots, K_r . Dari persamaan (1) komponen parametrik dinyatakan dalam bentuk:

$$\sum_{j=1}^p f_j(x_{ij}) = \sum_{j=1}^p \lambda_j x_{ij} \quad (2)$$

Berdasarkan persamaan (1) komponen nonparametrik dapat disajikan dalam bentuk persamaan sebagai berikut:

$$\sum_{k=1}^q g_k(z_{ik}) = \sum_{k=1}^q \left(\beta_k z_{ik} + \sum_{k=1}^r \beta_{(1+j)k} (z_{ik} - K_{kr})_+ \right) \quad (3)$$

dengan demikian, dari (2) dan (3) model regresi semiparametrik tersebut dapat dituliskan sebagai berikut:

$$y_i = \sum_{j=1}^p \lambda_j x_{ij} + \sum_{k=1}^q \left(\beta_k z_{ik} + \sum_{k=1}^r \beta_{(1+j)k} (z_{ik} - K_{kr})_+ \right) + \varepsilon_i \quad (4)$$

dimana fungsi *truncated* diberikan oleh:

$$(z_{ik} - K_{kr})_+ = \begin{cases} (z_{ik} - K_{kr}), & z_{ik} \geq K_{kr} \\ 0, & z_{ik} < K_{kr} \end{cases}$$

3.2. Estimasi Interval Parameter dan Pemilihan Titik Knot Optimal

Diberikan model regresi linier berganda:

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_p x_{pi} + \varepsilon_i, \quad i = 1, 2, \dots, n$$

untuk mendapatkan estimasi interval suatu parameter dalam regresi parametrik untuk koefisien regresi β_j ,

diasumsikan bahwa *error* ε_i berdistribusi normal independen dengan mean nol dan varians σ^2 . Oleh karena itu, observasi y_i berdistribusi normal dengan mean $\beta_0 + \sum_{j=1}^p \beta_j x_j$ dan varians σ^2 . Karena estimator *least-square* $\hat{\beta}$ berdistribusi normal dengan vektor mean β dan matriks covarians $\sigma^2(\mathbf{X}'\mathbf{X})^{-1}$. Akibatnya bahwa distribusi $\hat{\beta}_j$ adalah normal dengan mean β_j dan varians $\sigma^2 C_{jj}$, dimana C_{jj} merupakan elemen diagonal ke- j dari matriks $(\mathbf{X}'\mathbf{X})^{-1}$. Akibatnya masing-masing statistik[1]:

$$T_j = \frac{\hat{\beta}_j - \beta_j}{\sqrt{\sigma^2 C_{jj}}}, j = 0, 1, \dots, p \tag{5}$$

Dari Persamaan (5) berdistribusi *t* dengan derajat bebas $n-p$, dimana σ^2 diestimasi dengan varians *error*. Berdasarkan Persamaan (5), maka diperoleh interval konfidensi $(1-\alpha) \times 100\%$ untuk koefisien regresi $\beta_j, j = 0, 1, \dots, p$ adalah sebagai berikut:

$$\hat{\beta}_j - t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \leq \beta_j \leq \hat{\beta}_j + t_{\alpha/2, n-p} \sqrt{\hat{\sigma}^2 C_{jj}} \tag{6}$$

Salah satu metode untuk mencari titik knot optimal yang sering digunakan adalah dengan metode GCV. Titik optimal diperoleh dari nilai GCV minimum. Metode GCV secara umum adalah sebagai berikut:

$$GCV(K) = \frac{MSE(K)}{[n^{-1} \text{trace}(I - A(K))]^2}$$

dengan $MSE(K) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2$, I adalah matriks identitas, n adalah jumlah pengamatan, $A(K)$ adalah matriks $\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'$

3.3. Deskripsi Indeks Pembangunan Manusia di Kepulauan Maluku

Kepulauan Maluku terdiri dari 21 Kabupaten/Kota yakni 11 Kabupaten/Kota di Provinsi Maluku dan 10 Kabupaten/Kota di Provinsi Maluku Utara. IPM di Kepulauan Maluku secara umum termasuk dalam kategori IPM sedang pada tahun 2015.

Tabel 1. Deskriptif IPM Kepulauan Maluku

Variabel	Mean	Standar Deviasi	Maximum	Minimum
Y	64,05	5,56	79,30	58,26
X ₁	24,36	9,10	56,10	11,25
Z ₁	186,8	100,7	468,0	64,0
Z ₂	27,97	23,70	76,56	0,02
Z ₃	16,01	9,69	31,58	2,29
Z ₄	67,09	5,44	75,69	56,39

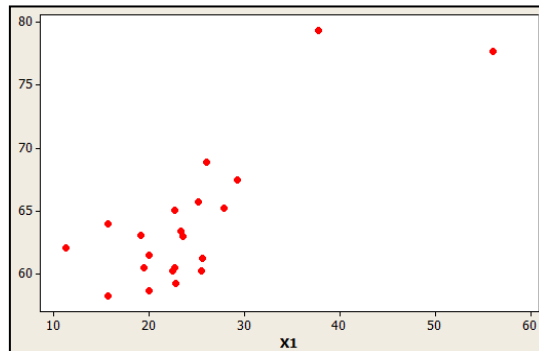
Berdasarkan Tabel 1, dapat dilihat karakteristik masing-masing variabel, baik variabel respon maupun variabel prediktor. Berikut adalah gambaran karakteristik untuk masing-masing variabel prediktor yaitu: persentase Kelompok Usia Sekolah di Perguruan Tinggi (X_1), Jumlah Sarana Kesehatan (Z_1), persentase Rumah Tangga dengan Akses Air Bersih (Z_2), persentase penduduk miskin (Z_3) dan persentase Tingkat Partisipasi Angkatan Kerja (TPAK) (Z_4).

- Rata-rata persentase kelompok usia sekolah di PT (X_1) di Kepulauan Maluku adalah 24,36 dengan standar deviasi sebesar 9,10. Kelompok usia sekolah di PT tertinggi pada Kota Ternate sebesar 56,10 persen dan terendah ada di Kabupaten Halmahera Tengah dengan angka 11,25 persen dengan *range* sebesar 44,85.
- Rata-rata jumlah sarana kesehatan (Z_1) di Kepulauan Maluku adalah 186,8 sarana dengan standar deviasi sebesar 100,7. Sarana kesehatan tertinggi pada Kabupaten Maluku Tengah sebesar 468 sarana dan terendah ada di Kabupaten Buru dengan angka 64 sarana dengan *range* sebesar 44,85 sarana.
- Rata-rata persentase RT dengan akses air bersih (Z_2) bersih di Kepulauan Maluku adalah 27,97 persen dengan standar deviasi sebesar 23,7 persen. RT dengan akses air bersih tertinggi pada Kota Ternate sebesar 76,56 persen dan terendah ada di Kabupaten Seram Bagian Timur dengan angka 0,02 persen dengan *range* sebesar 76,54 persen.

- d. Rata-rata persentase penduduk miskin (Z_3) di Kepulauan Maluku adalah 16,01 persen dengan standar deviasi sebesar 28,59 persen. Persentase penduduk miskin tertinggi pada Kabupaten Maluku Barat Daya sebesar 31,58 persen dan terendah ada di Kota Ternate dengan angka 2,99 persen.
- e. Rata-rata persentase TPAK (Z_4) di Kepulauan Maluku adalah 16,01 persen dengan standar deviasi sebesar 5,44 persen. PPM tertinggi pada Kabupaten Maluku Tenggara Barat sebesar 76,69 persen dan terendah ada di Kabupaten Pulau Taliabu dengan angka 56,39 persen dengan range sebesar 19,3 persen.

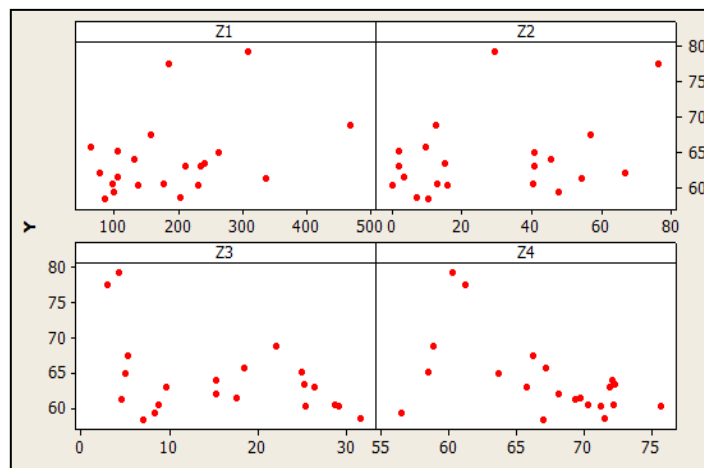
3.4. Penentuan Komponen Parametrik dan Variabel komponen Nonparametrik

Untuk melihat pola hubungan antara variabel respon dengan masing-masing variabel prediktor dapat dilihat dari *scatter plot*. Hasil *scatter plot* untuk masing-masing variabel respon dan variabel prediktor sebagai berikut:



Gambar 1. Scatter Plot Y dan X_1

Gambar 1, menunjukkan bahwa pola data cenderung membentuk pola hubungan tertentu yang dapat diregresikan secara parametrik, sehingga plot antara variabel IPM dengan Kelompok Usia Sekolah di Perguruan Tinggi (X_1) digunakan sebagai komponen parametrik.



Gambar 2. Scatter Plot antara Y dan Z_1, Z_2, Z_3, Z_4

Gambar 2, menunjukkan bahwa hubungan antara variabel respon dan keempat variabel prediktor polanya cenderung tidak membentuk pola hubungan tertentu atau pola nonparametrik sehingga digunakan sebagai komponen nonparametrik.

Tabel 2. Komponen Variabel Prediktor

Variabel Prediktor	Komponen
Kelompok Usia Sekolah di Perguruan Tinggi	Parametrik
Jumlah sarana Kesehatan	Nonparametrik
Persentase Air Minum Bersih	Nonparametrik
Persentase Penduduk Miskin	Nonparametrik
Tingkat Partisipasi Angkatan Kerja	Nonparametrik

Berdasarkan Tabel 2, diatas dapat dilihat bahwa hanya satu variabel prediktor (X) sebagai komponen parametrik dan keempat variabel lainnya sebagai komponen nonparametrik (Z).

3.5. Pemodelan IPM di Kepulauan Maluku Menggunakan Spline *Truncated* dengan GCV

IPM di Kepulauan Maluku diestimasi dengan Spline *Truncated* linier. Pemodelan IPM pada penelitian ini dengan satu variabel komponen parametrik dan empat variabel komponen nonparametrik menggunakan Spline *Truncated* dengan satu titik knot, dua titik knot, tiga titik knot dan kombinasi knot. Penentuan model terbaik diperoleh dari nilai GCV optimal. Berikut ini merupakan tabel perbandingan nilai GCV untuk masing-masing model.

Tabel 3. Perbandingan Nilai GCV untuk Pemilihan Model

Variabel Prediktor	Komponen
Kelompok Usia Sekolah di Perguruan Tinggi	Parametrik
Jumlah Sarana Kesehatan	Nonparametrik
Persentase Air Minum Bersih	Nonparametrik
Persentase Penduduk Miskin	Nonparametrik
Tingkat Partisipasi Angkatan Kerja	Nonparametrik

Berdasarkan Tabel 3, diatas diperoleh knot dengan GCV minimum 0,346 berada pada 3 titik knot maka knot tersebut adalah yang terbaik, dengan koefisien determinasi yang diperoleh 99,97%. Dari titik knot optimal terpilih maka model persamaan regresi semiparametrik Spline *Truncated* linier yang terbentuk adalah:

$$\hat{y} = 1.113x_1 + 0.123z_1 - 0.789(z_1 - 121.71)_+ + 0.71(z_1 - 138.2)_+ + 0.003(z_1 - 187.67)_+ + 0.049z_2 + 0.04(z_2 - 10.95)_+ + 1.671(z_2 - 14.08)_+ - 1.855(z_2 - 23,45)_+ + 9.126z_3 - 22,665(z_3 - 7.07)_+ + 17.138(z_3 - 8.24)_+ - 3.259(z_3 - 11.74)_+ - 0.761z_4 + 17.976(z_4 - 59.15)_+ - 16.321(z_4 - 59.93)_+ - 1.955(z_4 - 62.3)_+$$

3.6. Estimasi Interval Parameter Model Regresi Semiparametrik Spline *Truncated*

Dari model regresi semiparametrik Spline *Truncated* linier yang terpilih berdasarkan titik knot optimal maka model persamaan regresi semiparametrik Spline *Truncated* linier tersebut diestimasi dengan tingkat kepercayaan 95% dengan probabilitas interval sebagai berikut:

$$P \left[\hat{B}_{jk} - b \sqrt{\frac{\hat{y}'(I - S(S'S)^{-1}S')\hat{y}(S'S)^{-1}}{n - (p + q(r + 1))}} \leq \tilde{B}_{jk} \leq \hat{B}_{jk} + a \sqrt{\frac{\hat{y}'(I - S(S'S)^{-1}S')\hat{y}(S'S)^{-1}}{n - (p + q(r + 1))}} \right] = 1 - \alpha \tag{7}$$

Dengan menggunakan Persamaan (7) maka diperoleh estimasi interval terpendek dari masing-masing variabel dan parameter-parameternya [6]. Estimasi interval parameter model Spline *Truncated* linier diberikan pada Tabel 4.

Tabel 4. Estimasi Interval Parameter Model Spline *Truncated* Linier

Variabel	Parameter	Batas Bawah	Batas Atas
X_1	λ	0,681	1,546
Z_1	β_{11}	0,023	0,222
	β_{21}	-1,388	-0,191
	β_{31}	0,146	1,275
	β_{41}	-0,059	0,065
Z_2	β_{12}	-0,193	0,291
	β_{22}	-1,072	1,153
	β_{32}	0,138	3,204
	β_{42}	-2,787	-0,923
	β_{13}	4,313	13,940

Z_3	β_{23}	-34,73	-10,59
	β_{33}	8,662	25,615
	β_{43}	-4,333	-2,185
Z_4	β_{14}	-1,670	0,148
	β_{24}	11,542	24,411
	β_{34}	-22,84	-9,798
	β_{44}	-5,026	1,153

Tabel 4, memperlihatkan bahwa dari 17 parameter pada model regresi semiparametrik Spline *Truncated* linier, terdapat 5 parameter yang tidak signifikan yakni parameter $\beta_{41}, \beta_{12}, \beta_{22}, \beta_{14}$ dan β_{44} . Secara keseluruhan semua variabel prediktor berpengaruh signifikan terhadap IPM.

3.7. Pengujian Asumsi Residual

Suatu model dengan kriteria terbaik adalah jika residual memenuhi asumsi identik, independen dan berdistribusi normal. Hasil uji identik disajikan pada Tabel 5 dengan hipotesis:

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

$$H_1: \text{Minimal ada satu } \sigma_i^2 \neq \sigma^2, i = 1, 2, \dots, n$$

Tabel 5. Anova Uji Glejser

Sumber Variasi	Df	SS	MS	F	P-Value
Regresi	17	0,532	0,031	3,458	0,167
Error	3	0,027	0,009		
Total	20	0,559			

Nilai *p-value* 0,167 lebih besar dari alpha 0,05, maka dapat disimpulkan bahwa tidak terdapat kasus heteroskedastisitas, sehingga varians residual memenuhi asumsi identik. Selanjutnya pengujian asumsi independen menggunakan *Run Test* dengan hipotesis:

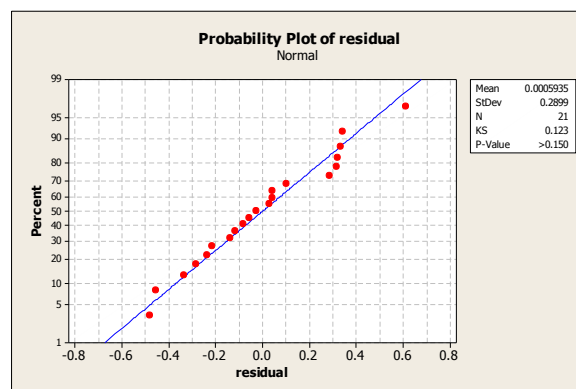
$$H_0 : \text{Residual random}$$

$$H_1 : \text{Residual tidak random}$$

Run Test terhadap residual menghasilkan nilai 1,00 lebih besar dari alpha 0,05 maka gagal tolak H_0 jadi disimpulkan residual bersifat random atau tidak ada korelasi antar residual, maka residual memenuhi asumsi independen. Langkah terakhir pengujian asumsi residual berdistribusi normal dengan hipotesis:

$$H_0 : F_0(x) = F(x) \text{ (Residual berdistribusi normal)}$$

$$H_1 : F_0(x) \neq F(x) \text{ (Residual tidak berdistribusi normal)}$$



Gambar 3. Residual berdistribusi normal

Dari Gambar 3, diatas diperoleh *p-value* >0,150 lebih besar dari 0,05 sehingga asumsi residual berdistribusi normal terpenuhi. Model tersebut adalah model terbaik karena sudah memenuhi uji klasik yakni, residual memenuhi asumsi identik, independen dan berdistribusi normal.

3.8. Interpretasi Model

Berdasarkan pengujian asumsi residual maka diperoleh model terbaik dengan 3 titik knot dengan koefisien determinasi 99,97 persen. Model regresi semiparametrik *Spline Truncated* tersebut dapat menjelaskan variabilitas IPM sebesar 99,97 persen dan terdapat 0,03 persen yang tidak dapat menjelaskan variabilitas IPM.

- a. Hubungan antara kelompok usia sekolah di PT (X_1) terhadap IPM (Y) dengan asumsi variabel lain konstan.

$$\hat{y} = 1,113x_1$$

jika terjadi kenaikan persentase kelompok usia sekolah di PT sebanyak satu persen maka IPM akan naik 1,113 persen.

- b. Hubungan antara jumlah sarana kesehatan (Z_1) terhadap IPM (Y) dengan asumsi variabel lain konstan yaitu

$$\hat{y} = 0.123z_1 - 0.789(z_1 - 121.71)_+ + 0.71(z_1 - 138.2)_+ + 0.003(z_1 - 187.67)_+$$

$$= \begin{cases} 0.123z_1 & , z_1 < 121.71 \\ -0.666z_1 + 96,029 & , 121.71 \leq z_1 < 138.2 \\ 0.044z_1 - 2.093 & , 138.2 \leq z_1 < 187.67 \end{cases}$$

Jika jumlah sarana kesehatan kurang dari 122 sarana, dan jumlah sarana kesehatan naik sebesar satu sarana, IPM akan naik sebesar 0,123 persen. Jika jumlah sarana kesehatan antara 122 dan 138 sarana, dan jumlah sarana kesehatan naik sebesar satu sarana, IPM akan turun sebesar 0,666 persen. Wilayah yang masuk dalam interval ini adalah Kabupaten Maluku Tenggara Barat, Buru, Buru Selatan, Kota Tual, Halmahera Tengah, Kepulauan Sula, Halmahera Timur, Pulau Morotai dan Pulau Taliabu. Jika jumlah sarana kesehatan antara 138 dan 187 sarana, dan jumlah sarana kesehatan naik sebesar satu sarana, IPM akan naik sebesar 0,044 persen.

- c. Hubungan sumber air bersih (Z_2) terhadap IPM (Y), dengan asumsi variabel lain konstan yaitu:

$$\hat{y} = 0.049z_2 + 0.04(z_2 - 10.95)_+ + 1.671(z_2 - 14.08)_+ - 1.855(z_2 - 23.45)_+$$

$$\hat{y} = \begin{cases} 1.711z_2 - 23.966 & , 14.08 \leq z_2 < 23.45 \\ -0.144z_2 + 19.534 & , z_2 \geq 23.45 \end{cases}$$

Jika persentase sumber air bersih terletak pada segmen 14,08 dan 23,45 dan jika persentase sumber air bersih naik sebesar satu persen, IPM akan naik sebesar 1,71 persen. Sedangkan pada segmen persentase sumber air bersih lebih besar atau sama dengan 23,45 persen, dan jika persentase sumber air bersih naik sebesar satu persen, IPM akan turun sebesar 0,14 persen. Wilayah yang masuk dalam interval ini adalah Kota Ambon, Halmahera Barat, Halmahera Tengah, Kepulauan Sula, Halmahera Selatan, Halmahera Utara, Halmahera Timur, Pulau Morotai, Kota Ternate dan Kota Tidore.

- d. Hubungan antara persentase penduduk miskin (Z_3) terhadap IPM (Y) dengan asumsi variabel lain konstan yaitu:

$$\hat{y} = 9.126z_3 - 22.665(z_3 - 7.07)_+ + 17.138(z_3 - 8.24)_+ - 3.259(z_3 - 11.74)_+$$

$$= \begin{cases} 9.126z_3 & , z_3 < 7.07 \\ -13.539z_3 + 160.242 & , 7.07 \leq z_3 < 8.24 \\ 3.599z_3 + 19.024 & , 8.24 \leq z_3 < 11.74 \\ 0.34z_3 + 57.285 & , z_3 \geq 11.74 \end{cases}$$

Model di atas menggambarkan persentase penduduk miskin terletak pada segmen 7,07 dan 8,24 persen, jika persentase penduduk miskin naik sebesar satu persen, IPM akan turun sebesar 13,54 persen. Selanjutnya jika persentase penduduk miskin terletak pada segmen 8,24 dan 11,74 persen, dan persentase penduduk miskin naik sebesar satu persen, IPM akan naik sebesar 3,59 persen. Wilayah yang masuk dalam interval ini adalah Kota Ambon, Halmahera Barat, Halmahera Utara, Pulau Morotai, Pulau Taliabu, Kota Ternate dan Kota Tidore. Sedangkan segmen persentase penduduk miskin lebih besar atau sama dengan 11,74, jika persentase penduduk miskin naik sebesar satu persen, IPM akan naik sebesar 0,34 persen. Wilayah yang masuk dalam interval ini adalah Kabupaten Maluku Tenggara Barat, Maluku Tenggara, Maluku Tengah, Buru, Aru, Seram Bagian Barat, Seram Bagian Timur, Maluku Barat Daya, Buru Selatan, Kota Tual, Halmahera Tengah dan Halmahera Timur.

Hubungan antara TPAK (Z_3) terhadap IPM (Y), dengan asumsi variabel lain konstan yaitu:

$$\hat{y} = -0.761z_4 + 17.976(z_4 - 59.15)_+ - 16.321(z_4 - 59.93)_+ - 1.955(z_4 - 62.3)_+$$

$$= \begin{cases} 17.215z_4 - 196.837 & , 59.15 \leq z_4 < 59.93 \\ 30.894z_4 - 62.352 & , 59.93 \leq z_4 < 62.3 \end{cases}$$

Model tersebut menggambarkan persentase TPAK terletak pada segmen 59,15 hingga 59,93, dan segmen 59,93 hingga 62,3. Jika persentase TPAK naik sebesar satu persen, maka persentase IPM meningkat menjadi 17,21 persen di segmen 59,15 hingga 59,93. Wilayah yang masuk dalam interval ini adalah Kabupaten Maluku Tengah dan Kota Tual. Jika persentase TPAK naik sebesar satu persen, maka persentase IPM akan naik sebesar 30,89 persen di segmen 59,93 hingga 62,3. Wilayah yang masuk dalam interval ini adalah Kota Ambon dan Kota Ternate.

4. KESIMPULAN

Model terbaik yang diperoleh adalah dengan tiga titik knot. Model tersebut memiliki koefisien determinasi 99,97% dengan nilai GCV minimum 0,346. Pada pemodelan dengan pendekatan estimasi interval parameter model regresi semiparametrik Spline *Truncated* tersebut mampu menjelaskan semua variabel prediktor yakni variabel persentase kelompok usia sekolah di perguruan tinggi, jumlah sarana kesehatan, persentase rumah tangga dengan akses sumber air minum bersih, persentase penduduk miskin dan persentase TPAK berpengaruh terhadap variabel respon (IPM). Adapun ke depannya perlu dikembangkan pemilihan titik knot optimal dengan menggunakan *Unbiased Risk* (UBR) dan *Generalized Maximum Likelihood* (GML) dengan model Spline yang lain.

DAFTAR PUSTAKA

- [1] A. Rencher & G.B. Schaalje, "Linear Models in Statistics 2nd Edition", *John Wiley & Sonc Inc, New Jersey:2008*
- [2] Badan Pusat Statistika, "Laporan Tahunan BPS", *BPS RI, Jakarta, 2017*
- [3] Badan Pusat Statistika, "Maluku Dalam Angka Tahun 2016", *BPS, Ambon, 2017.*
- [4] Badan Pusat Statistika, "Maluku Utara Dalam Angka Tahun 2016", *BPS, Ternate, 2017.*
- [5] G. Wahba., & S. A. Wold, "Completely Automatic French Curve: Fitting Spline Functions by Cross-Validation", *Comm. Statist. 1975.*
- [6] I. N. Budiantara, "Metode U, GML, CV dan GCV Dalam Regresi Nonparametrik Spline", *Majalah Ilmiah Himpunan Matematika Indonesia (MIHMI), 6, 285-290, 2000*
- [7] I. N. Budiantara, "Model Spline dengan Knot Optimal", *Jurnal FMIPA Universitas Jember, Jember, 2006*
- [8] L. Annisah, "Estimasi Kurva Regresi Semiparametrik dengan Pendekatan Spline Linier Pada Data Kemiskinan di Provinsi Jawa Tengah Tahun 2012", *Tesis Jurusan Statistika, FMIPA ITS, Surabaya, 2016*
- [9] M. Nafi, "Estimasi Interval Spline Dalam Regresi Nonparametrik", *Tesis Jurusan Statistika, FMIPA ITS, Surabaya, 2010*
- [10] S. B. Loklomin, I. N. Budiantara, & I. Zain, "Estimation of Interval Parameters Semiparametric Regression Models Spline Truncated", *Internasional Conference on Mathematics Science and Education (ICoMSE), UNM Malang, 2017*

