

# Pemodelan Faktor yang Memengaruhi Angka Partisipasi Kasar SMA/ sederajat di Papua Menggunakan Regresi Nonparametrik *Spline Truncated*

Rizkiana Prima Rahmadina, Madu Ratna dan I Nyoman Budiantara  
Departemen Statistika, Institut Teknologi Sepuluh Nopember (ITS)  
*e-mail*: madu\_r@statistika.its.ac.id

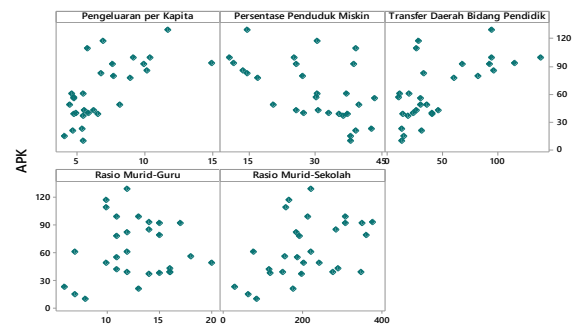
**Abstrak**—Angka Partisipasi Kasar (APK) adalah proporsi anak sekolah pada suatu jenjang terhadap penduduk pada kelompok usia tertentu. APK dapat digunakan sebagai tolok ukur keberhasilan program pembangunan pendidikan. Pada tahun 2018, APK SMA/ sederajat di Papua sebesar 63,26%, angka tersebut adalah yang paling rendah di Indonesia. Angka tersebut juga mengalami penurunan dari tahun sebelumnya. Adanya program wajib belajar 12 tahun, *quality education* dalam SDGs, dan penurunan nilai APK jenjang SMA/ sederajat di Provinsi Papua, maka perlu dilakukan upaya peningkatan dengan mengetahui faktor yang memengaruhi APK jenjang SMA/ sederajat. Penelitian ini menggunakan metode regresi nonparametrik *spline* karena lima variabel prediktor dengan unit observasi sebanyak 29 kabupaten/kota di Provinsi Papua tidak memiliki suatu pola tertentu. Didapatkan nilai GCV yang paling minimum sebesar 243,7483 dengan model menggunakan kombinasi titik knot (1,3,1,3,3). Pengujian asumsi residual menunjukkan semua asumsi terpenuhi dengan nilai koefisien determinasi atau  $R^2$  yang menunjukkan ukuran kebaikan model sebesar 95,86%. Variabel yang digunakan adalah pengeluaran per kapita, persentase penduduk miskin, transfer daerah bidang pendidikan, rasio murid-guru, dan rasio murid-sekolah.

**Kata Kunci**—Angka Partisipasi Kasar, GCV, Papua, Regresi Nonparametrik *Spline Truncated*, Titik Knot.

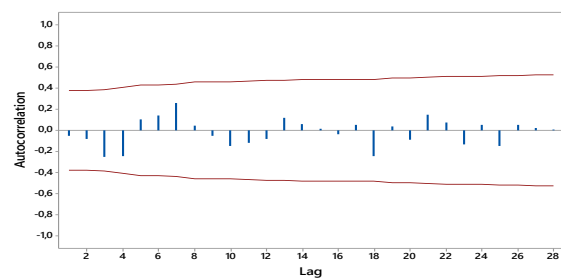
## I. PENDAHULUAN

MENURUT Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sistem Pendidikan Nasional (Sisdiknas) menyatakan bahwa wajib belajar adalah program pendidikan minimal yang harus diikuti oleh semua warga Negara Indonesia. Secara nasional program pemerintah terkait wajib belajar tertuang pada Program Indonesia Pintar (PIP) melalui penerbitan Kartu Indonesia Pintar (KIP) dengan Permen Dikbud No 19 tahun 2016 Pasal 2 huruf a yang berbunyi “PIP bertujuan untuk meningkatkan akses bagi anak usia 6 (enam) sampai dengan 21 (dua puluh satu) tahun untuk mendapatkan layanan pendidikan sampai tamat satuan pendidikan menengah dalam rangka mendukung pelaksanaan pendidikan menengah universal/rintisan wajib belajar 12 (dua belas) tahun”.

Pendidikan juga termasuk dalam *Sustainable Development Goals* (SDGs) yang keempat, yakni menjamin kualitas pendidikan yang inklusif dan merata serta meningkatkan kesempatan belajar sepanjang hayat untuk semua atau *quality education* [1]. *Quality education* merupakan salah satu tujuan dari negara-negara dunia untuk meningkatkan kualitas pendidikan dengan cara meningkatkan partisipasi penduduk



Gambar 1. Pola Hubungan Variabel Respon dan Prediktor.

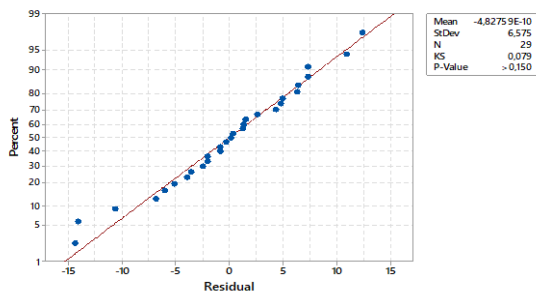


Gambar 2. Plot ACF.

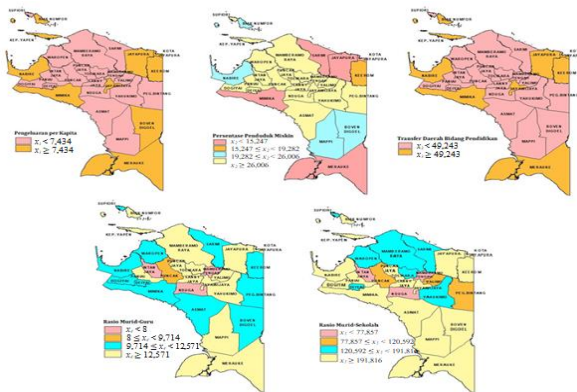
dalam pendidikan/sekolah tanpa memandang perbedaan jenis kelamin dan keterbatasan kemampuan penduduk.

Angka Partisipasi Kasar (APK) merupakan proporsi jumlah penduduk yang sedang bersekolah pada suatu jenjang pendidikan tertentu terhadap jumlah penduduk usia sekolah yang sesuai dengan jenjang pendidikan tersebut. Indikator ini mengindikasikan partisipasi sekolah penduduk pada jenjang pendidikan tertentu tanpa memperhatikan usia. Nilai APK bisa lebih dari 100 persen, karena populasi murid yang bersekolah pada suatu jenjang pendidikan tertentu mencakup anak di luar batas usia sekolah pada jenjang pendidikan tersebut. Hal ini disebabkan antara lain oleh adanya pendaftaran siswa usia dini, pendaftaran siswa yang telat bersekolah, atau pengulangan kelas.

Pada tahun 2018 secara nasional Provinsi Papua menduduki peringkat terendah dan di bawah rata-rata dalam hal indikator partisipasi pendidikan. APK SD/MI/ sederajat sebesar 95,66% dan menurun pada setiap tingkatan pendidikan sehingga APK terendah ada pada tingkat Sekolah Menengah Atas (SMA)/ sederajat yaitu sebesar 63,26%, sedangkan rata-rata di Indonesia adalah sebesar 80,68%. Angka tersebut mengalami penurunan dari tahun sebelumnya, nilai APK jenjang SMA/ sederajat di Provinsi Papua pada tahun 2017 sebesar 65,49%.



Gambar 3. Plot Normalitas Residual.



Gambar 4. Peta Persebaran Kabupaten/Kota di Papua menurut Kelima Variabel Prediktor.

Oleh karena adanya program wajib belajar 12 tahun, *quality education* dalam SDGs, dan penurunan nilai APK jenjang SMA/ sederajat di Provinsi Papua, maka diperlukan suatu upaya peningkatan, yakni salah satunya dengan meneliti untuk mengetahui faktor-faktor yang memengaruhi Angka Partisipasi Kasar jenjang SMA/ sederajat di Provinsi Papua. Data Angka Partisipasi Kasar jenjang SMA/ sederajat dan faktor-faktor yang diduga memengaruhi dalam penelitian ini akan dimodelkan dengan menggunakan regresi nonparametrik *spline truncated*. Metode tersebut dipilih karena berdasarkan plot pada studi awal yang telah dilakukan, diketahui bahwa plot variabel-variabel yang diduga memengaruhi Angka Partisipasi Kasar jenjang SMA/ sederajat di Provinsi Papua tidak membentuk suatu pola data tertentu dan terjadi perubahan pola pada sub interval tertentu. Pemilihan metode regresi nonparametrik *spline truncated* didasari karena sifat fleksibilitas tinggi dalam menangani perubahan perilaku data, dapat memodelkan data yang memiliki pola berubah-ubah pada interval tertentu, metode ini lebih sederhana, interpretasi tidak rumit, dan perumusan matematis lebih sederhana dibanding metode regresi nonparametrik *spline* lainnya [2].

Untuk memilih model regresi *spline* yang terbaik, dilakukan pemilihan model dengan titik knot yang optimal. Salah satu metode yang dapat digunakan untuk menentukan titik knot optimal adalah *Generalized Cross Validation* (GCV). Pada penelitian ini diharapkan dapat memberikan pemodelan yang baik dan dapat mengetahui faktor-faktor yang memengaruhi APK SMA/ sederajat di Provinsi Papua sehingga dapat dijadikan salah satu acuan bagi pemerintah atau instansi terkait untuk meningkatkan nilai APK jenjang SMA/ sederajat di Provinsi Papua.

Tabel 1.

Analisis Ragam (ANOVA) Uji Parameter

Sumber Variasi	Derajat Bebas (df)	Jumlah Kuadrat (SS)	Rataan Kuadrat (MS)	F
Regresi	$p+K$	$\sum_{i=1}^n (\hat{y}_i - \bar{y})^2$	$\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{p+K}$	$\frac{MS_{regresi}}{MS_{error}}$
Error	$n-(p+K)-1$	$\sum_{i=1}^n (y_i - \hat{y}_i)^2$	$\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-(p+K)-1}$	-
Total	$n-1$	$\sum_{i=1}^n (y_i - \bar{y})^2$	-	-

Tabel 2.

Variabel Data Penelitian

Variabel	Keterangan	Skala
$y$	Angka Partisipasi Kasar Jenjang SMA/ sederajat	Rasio
$x_1$	Pengeluaran per Kapita	Rasio
$x_2$	Persentase Penduduk Miskin	Rasio
$x_3$	Transfer Daerah Bidang Pendidikan	Rasio
$x_4$	Rasio Murid-Guru	Rasio
$x_5$	Rasio Murid-Sekolah	Rasio

## II. LANDASAN TEORI

### A. Statistika Deskriptif

Statistika deskriptif merupakan metode-metode yang berkaitan dengan pengumpulan dan penyajian data sehingga memberikan informasi yang berguna [3]. Dalam statistika deskriptif terdapat dua macam penyajian, yaitu visualisasi data, seperti tabel, grafik, diagram, dan pictogram serta penyajian dalam bentuk ukuran-ukuran statistik, yakni ukuran pemusatan dan penyebaran. Ukuran pemusatan data meliputi rata-rata (*mean*), median dan modus, sedangkan ukuran penyebaran data berupa rentang data (*range*) dan varians.

### B. Regresi Nonparametrik

Regresi nonparametrik merupakan model regresi yang digunakan untuk mengetahui pola hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk kurva regresinya atau tidak terdapat informasi distribusi mengenai bentuk pola data [4]. Secara umum, Persamaan (1) merupakan model regresi nonparametric dengan  $y_i$  adalah variabel respon ke- $i$ ,  $f(x_i)$  adalah fungsi regresi yang tidak diketahui polanya,  $x_i$  merupakan variabel prediktor, dan  $\varepsilon_i$  adalah *error random* ke- $i$  dengan asumsi identik, independen, dan berdistribusi normal dengan *mean* nol dan variansi  $\sigma^2$ .

$$y_i = f(x_i) + \varepsilon_i, i = 1, 2, \dots, n \tag{1}$$

### C. Regresi Nonparametrik Spline Truncated

*Spline* merupakan potongan polinomial yang memiliki sifat tersegmen sehingga memiliki fleksibilitas yang tinggi dan memiliki kemampuan sangat baik untuk menangani data yang perilakunya berubah pada sub interval tertentu. Salah satu kelebihan pendekatan *spline* adalah model ini cenderung mencari sendiri estimasi data ke mana pola data tersebut bergerak. Kelebihan ini terjadi karena dalam *spline* terdapat titik-titik knot [5]. Masing-masing fungsi  $f$  pada regresi nonparametrik *spline truncated* multivariable berorde  $p$

Tabel 3.  
Karakteristik APK SMA/ sederajat dan Faktor-Faktor yang Diduga Berpengaruh

Variabel	Mean	Varians	SD	Min	Maks
y	63,26	1045,11	32,33	9,40	129,86
x <sub>1</sub>	6,96	6,61	2,57	4,13	14,92
x <sub>2</sub>	29,43	100,55	10,03	10,54	43,49
x <sub>3</sub>	46,25	1285,78	35,86	9,49	139,35
x <sub>4</sub>	12,62	11,46	3,39	6,00	20,00
x <sub>5</sub>	208,72	8866,85	94,16	28,00	377,00

Tabel 4.  
Perbandingan Nilai GCV

Knot	GCV Minimum
Satu titik knot	672,6121
Dua titik knot	594,2564
Tiga titik knot	446,4743
<b>Kombinasi knot (1,3,1,3,3)</b>	<b>243,7483</b>

Tabel 5.  
Analisis Ragam (ANOVA)

Sumber Variasi	df	SS	MS	F <sub>hitung</sub>	P-value
Regresi	16	28052,79	1753,299	17,383	6,67661x10 <sup>-6</sup>
Error	12	1210,336	100,8614		
Total	28	29263,13			

dengan titik knot λ<sub>1</sub>, λ<sub>2</sub>, ..., λ<sub>K</sub> dapat dituliskan menjadi Persamaan (2).

$$f(x_{li}) = \sum_{j=1}^p \beta_{lj} x_{li}^j + \sum_{k=1}^K \beta_{l(p+k)} (x_{li} - \lambda_{lk})_+^p \quad (2)$$

Sehingga diperoleh persamaan model regresi nonparametrik *spline truncated* sebagai berikut,

$$y_i = \beta_0 + \sum_{l=1}^q \sum_{j=1}^p \beta_{lj} x_{li}^j + \sum_{l=1}^q \sum_{k=1}^K \beta_{l(p+k)} (x_{li} - \lambda_{lk})_+^p + \varepsilon_i, \quad i = 1, 2, L, n \quad (3)$$

fungsi *truncated* (x<sub>i</sub> - λ<sub>k</sub>)<sub>+</sub><sup>p</sup> akan menghasilkan Persamaan (4),

$$(x_{li} - \lambda_{lk})_+^p = \begin{cases} (x_{li} - \lambda_{lk})^p, & x_{li} \geq \lambda_{lk} \\ 0, & x_{li} < \lambda_{lk} \end{cases} \quad (4)$$

dimana,

β<sub>j</sub> : parameter model polinomial, j = 1, 2, ..., p

x<sub>i</sub> : variabel prediktor i = 1, 2, ..., n

β<sub>p+k</sub> : parameter pada komponen *truncated* k = 1, 2, ..., K

K : banyaknya knot

λ<sub>k</sub> : titik-titik knot

#### D. Pemilihan Titik Knot Optimal

Titik knot merupakan adalah titik perpaduan bersama dimana terdapat perubahan perilaku pola data pada interval yang berlainan. Model regresi *spline* terbaik tergantung pada titik knot yang optimal. Salah satu metode yang dapat digunakan untuk menentukan titik knot optimal adalah *Generalized Cross Validation* (GCV). Pemilihan metode GCV didasarkan pada kelebihan yang dimilikinya antara lain: sederhana dan efisien dalam perhitungan, optimal secara asimtotik, tidak memerlukan informasi terhadap varians populasi (σ<sup>2</sup>), invarian terhadap transformasi [6]. Fungsi GCV dapat dituliskan seperti pada Persamaan (5).

$$CGV(\lambda_1, \lambda_2, \dots, \lambda_K) = \frac{MSE(\lambda_1, \lambda_2, \dots, \lambda_K)}{(n^{-1} \text{tr}[I - A(\lambda_1, \lambda_2, \dots, \lambda_K)])^2} \quad (5)$$

Tabel 6.  
Hasil Pengujian Signifikansi Parameter Secara Parsial

Variabel	Parameter	Estimator	t	P-value
Konstan	β <sub>0</sub>	354,767	2,237	0,045
x <sub>1</sub>	β <sub>1</sub>	13,771	2,2255	0,046
	β <sub>2</sub>	-12,6725	-2,183	0,0496
	β <sub>3</sub>	13,338	2,3598	0,036
x <sub>2</sub>	β <sub>4</sub>	-42,568	-4,836	0,0004
	β <sub>5</sub>	41,763	6,3199	3,832x10 <sup>-5</sup>
	β <sub>6</sub>	-12,326	-4,691	0,0005
x <sub>3</sub>	β <sub>7</sub>	-0,745	-1,938	0,077
	β <sub>8</sub>	1,624	2,8625	0,014
	β <sub>9</sub>	-111,09	-	4,1035
x <sub>4</sub>	β <sub>10</sub>	54,664	1,0677	0,3066
	β <sub>11</sub>	30,676	0,453	0,6586
	β <sub>12</sub>	30,596	5,2826	0,0002
x <sub>5</sub>	β <sub>13</sub>	3,496	4,5116	0,0007
	β <sub>14</sub>	2,0685	0,689	0,5039
	β <sub>15</sub>	-5,605	-1,599	0,1356
	β <sub>16</sub>	0,046	0,186	0,8556

Tabel 7.  
ANOVA Uji Glejser

Sumber	df	SS	MS	F <sub>hit</sub>	P-value
Regresi	16	125,6534	7,85334	0,26326	0,9927
Error	12	357,9756	29,8313		
Total	28	483,629			

Dengan **I** adalah matriks identitas, n adalah jumlah observasi, A(λ<sub>1</sub>, λ<sub>2</sub>, ..., λ<sub>k</sub>) merupakan X(X<sup>T</sup>X)<sup>-1</sup>X<sup>T</sup>, dan MSE(λ<sub>1</sub>, λ<sub>2</sub>, ..., λ<sub>k</sub>) = n<sup>-1</sup> ∑<sub>i=1</sub><sup>n</sup> (y<sub>i</sub> - ŷ<sub>i</sub>)<sup>2</sup>

#### E. Pengujian Parameter Model Regresi

Pengujian parameter model digunakan untuk mengetahui apakah variabel prediktor memiliki pengaruh signifikan terhadap variabel respon atau tidak. Tahapan dalam pengujian parameter adalah pengujian secara serentak dan pengujian secara parsial atau individu.

##### 1) Uji Serentak

Pengujian secara serentak dilakukan untuk mengetahui signifikansi parameter model regresi secara bersama-sama terhadap variabel respon.

Hipotesis :

$$H_0: \beta_1 = \beta_2 = \dots = \beta_{p+K} = 0$$

$$H_1: \text{Minimal ada satu } \beta_j \neq 0, \quad j = 1, 2, \dots, p + K$$

Statistik uji :

$$F_{hitung} = \frac{MS_{regresi}}{MS_{error}} \quad (6)$$

dimana p adalah banyaknya parameter polinomial tanpa β<sub>0</sub> dan K adalah banyaknya parameter untuk titik knot.

Perhitungan nilai statistik uji F didapatkan dari Analisis Ragam (ANOVA) seperti yang ditunjukkan Tabel 1.

Daerah penolakan H<sub>0</sub> adalah F > F<sub>(α,(p+K),(n-(p+K)-1))</sub> atau p-value < α. Jika H<sub>0</sub> ditolak maka dapat disimpulkan bahwa minimal terdapat satu parameter pada model regresi *spline* yang signifikan terhadap model [7].

##### 2) Uji Parsial

Pengujian parsial bertujuan untuk mengetahui parameter mana yang berpengaruh dan tidak berpengaruh signifikan terhadap model regresi [8].

Hipotesis :

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, p + K$$

Statistik uji :

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (7)$$

$SE(\hat{\beta}_j)$  adalah *standard error*  $\hat{\beta}_j$  dimana  $SE(\hat{\beta}_j) = \sqrt{Var(\hat{\beta}_j)}$  dengan  $Var(\hat{\beta}_j)$  merupakan elemen diagonal utama ke- $j$  dari matriks  $Var(\hat{\beta}_j)$ ,  $j = 1, 2, \dots, p + K$ . Daerah penolakan  $H_0$  adalah  $|t_{hitung}| > t_{(\alpha/2, n-(p+K)-1)}$  atau  $p\text{-value} < \alpha$  yang dapat disimpulkan bahwa parameter ke- $j$  berpengaruh signifikan terhadap variabel respon [7].

#### F. Koefisien Determinasi ( $R^2$ )

Salah satu kriteria untuk menentukan kebaikan model regresi adalah menggunakan koefisien determinasi. Semakin tinggi nilai  $R^2$  yang dihasilkan suatu model, maka semakin baik pula variabel-variabel prediktor dalam model dalam menjelaskan variabilitas variabel respon. Berikut merupakan rumus dari koefisien determinasi ( $R^2$ ). Nilai  $R^2$  dapat diperoleh dengan rumus pada Persamaan (8),

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

#### G. Pengujian Asumsi Residual

Uji asumsi residual adalah pengujian yang dilakukan untuk mengetahui apakah residual telah memenuhi asumsi. Residual dari model regresi *spline* harus memenuhi asumsi identik, independen, dan berdistribusi normal  $\sim$  IIDN  $(0, \sigma^2)$ . Terdapat tiga asumsi yang harus dipenuhi, yaitu identik, independen, dan berdistribusi normal.

##### 1) Asumsi Identik

Pengujian asumsi identik bertujuan untuk mengetahui homogenitas variansi residual. Variansi antar residual harus sama atau tidak terjadi heteroskedastisitas seperti yang dituliskan pada Persamaan (9),

$$var(y_i) = var(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n \quad (9)$$

Pengujian asumsi identik dilakukan dengan uji *Glejser*. Hipotesis yang digunakan dalam pengujian ini adalah sebagai berikut [9].

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

$$H_1 : \text{minimal terdapat satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, \dots, n$$

Persamaan (10) merupakan statistik uji uji *Glejser*.

$$F_{hitung} = \frac{[\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|)^2] / (p+K)}{[\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|)^2] / (n-(p+K)-1)} \quad (10)$$

Daerah penolakan  $H_0$  adalah  $F_{hitung} > F_{\alpha; (p+K), n-(p+K)-1}$  atau  $p\text{-value} < \alpha$  yang dapat disimpulkan bahwa terjadi kasus heteroskedastisitas, sehingga asumsi residual identik tidak terpenuhi.

##### 2) Asumsi Independen

Asumsi independen dilakukan untuk mengetahui apakah terdapat autokorelasi atau tidak pada residual. Untuk mengetahui apakah terjadi autokorelasi secara visual dilakukan dengan melihat plot dari *Autocorrelation Function* (ACF) dari residual. Untuk mendapatkan nilai ACF digunakan rumus Persamaan (11),

$$\hat{\rho}_k = \frac{\sum_{t=k+1}^n (e_{t-k} - \bar{e})(e_t - \bar{e})}{\sum_{t=1}^n (e_t - \bar{e})^2} \quad (11)$$

dimana  $\hat{\rho}_s$  merupakan korelasi antara  $e_i$  dan  $e_{i-s}$ . Asumsi independen dapat dideteksi dengan menggunakan interval konfidensi dengan Persamaan (12)

$$-t_{\alpha/2, n-1} SE(\hat{\rho}_s) < \rho_s < t_{\alpha/2, n-1} SE(\hat{\rho}_s) \quad (12)$$

Dimana

$$SE(\hat{\rho}_s) = \sqrt{\frac{1 + 2 \sum_{s=0}^{s-1} (\hat{\rho}_s)^2}{n}} \quad (13)$$

Apabila terdapat minimal satu nilai ACF yang keluar dari interval konfidensi, maka diindikasikan adanya autokorelasi antar residual. Sebaliknya, jika tidak terdapat nilai ACF yang keluar dari interval konfidensi, maka tidak terdapat kasus autokorelasi atau asumsi independen terpenuhi.

##### 3) Asumsi Distribusi Normal

Uji asumsi distribusi normal residual dilakukan untuk mengetahui residual mengikuti distribusi normal atau tidak. Pengujian normalitas dapat dilakukan dengan uji *Kolmogorov-Smirnov* dengan hipotesis sebagai berikut.

Hipotesis :

$$H_0 : F_n(x) = F_0(x) \text{ (residual berdistribusi normal)}$$

$$H_1 : F_n(x) \neq F_0(x) \text{ (residual tidak berdistribusi normal)}$$

Persamaan (14) merupakan nilai statistik uji *Kolmogorov-Smirnov*.

$$D = \sup |S(x) - F_0(x)| \quad (14)$$

$F_0(x)$  merupakan fungsi distribusi yang dihipotesiskan, sedangkan  $S(x)$  adalah fungsi peluang kumulatif yang dihitung dari data sampel. Daerah penolakan  $H_0$  adalah  $|D| > D_\alpha$  atau  $p\text{-value} < \alpha$ .  $D_\alpha$  adalah nilai kritis untuk uji *Kolmogorov Smirnov* satu sampel pada tabel *Kolmogorov-Smirnov* satu sampel. Jika tolak  $H_0$ , maka dapat disimpulkan bahwa residual tidak berdistribusi normal [9].

##### 4) Angka Partisipasi Kasar

Badan Pusat Statistik memiliki beberapa indikator yang digunakan untuk memantau program pendidikan yang telah diberikan pemerintah. Indikator-indikator tersebut antara lain: Angka Partisipasi Sekolah (APS), Angka Partisipasi Murni (APM), Angka Partisipasi Kasar (APK), Pendidikan Tertinggi yang Ditamatkan, Rata-rata Lama Sekolah, Angka Melek Huruf (AMH), Angka Buta Huruf, dan Pendidikan Anak Usia Dini. Partisipasi sekolah menggambarkan efektivitas program pendidikan dalam menyerap potensi pendidikan yang ada di masyarakat, yang artinya semakin tinggi nilainya menunjukkan semakin efektifnya suatu program. Angka Partisipasi Kasar adalah proporsi anak sekolah pada suatu jenjang terhadap penduduk pada kelompok usia tertentu. Semakin tinggi Angka Partisipasi Kasar suatu jenjang, maka partisipasi penduduk dalam melakukan kegiatan bersekolah semakin besar.

### III. METODE PENELITIAN

#### A. Sumber Data

Data yang digunakan dalam penelitian ini merupakan data sekunder yang diperoleh dari Badan Pusat Statistik (BPS) dan Kementerian Pendidikan dan Kebudayaan. Unit observasi

yang digunakan dalam penelitian ini adalah 29 kabupaten/kota di Provinsi Papua.

### B. Variabel Penelitian

Variabel respon dan prediktor yang digunakan dalam penelitian ini disajikan pada Tabel 2.

### C. Langkah Analisis

Langkah-langkah analisis dalam penelitian ini adalah:

1. Mendeskripsikan APK jenjang SMA/ sederajat di Provinsi Papua pada tahun 2018 dan faktor-faktor yang diduga memengaruhinya.
2. Mengidentifikasi pola hubungan antara variabel respon dengan masing-masing variabel prediktor dengan menggunakan *scatterplot*.
3. Memodelkan variabel respon menggunakan model regresi nonparametrik *spline truncated* dengan satu, dua, tiga, dan kombinasi knot.
4. Menentukan titik knot optimal yang didapatkan dari nilai GCV yang paling minimum.
5. Mendapatkan model regresi nonparametrik *spline truncated* terbaik dengan titik knot optimal.
6. Melakukan uji signifikansi parameter regresi nonparametrik *spline truncated* secara serentak dan parsial.
7. Melakukan uji asumsi residual identik, independen, dan berdistribusi normal (IIDN) dari model regresi nonparametrik *spline truncated*.
8. Menghitung nilai koefisien determinasi  $R^2$ .
9. Membuat interpretasi model dan menarik kesimpulan.

## IV. ANALISIS DAN DISKUSI

### A. Karakteristik Angka Partisipasi Kasar Jenjang SMA/ sederajat di Papua dan Faktor-Faktor yang Diduga Memengaruhinya

Karakteristik APK SMA/ sederajat di Provinsi Papua beserta faktor-faktor yang diduga memengaruhinya ditunjukkan melalui statistika deskriptif pada Tabel 3.

Berdasarkan Tabel 3 dapat diketahui variabel respon ( $y$ ) yang merupakan APK SMA/ sederajat di Provinsi Papua memiliki nilai rata-rata sebesar 63,26%. Terdapat 17 kabupaten/kota dengan nilai APK SMA/ sederajat di bawah rata-rata. Kabupaten Mimika merupakan wilayah dengan nilai APK SMA/ sederajat yang tertinggi di Provinsi Papua yaitu sebesar 129,86%, sedangkan nilai APK SMA/ sederajat terendah di Provinsi Papua, yaitu Kabupaten Puncak sebesar 9,40%. Sebaran data (standar deviasi) seluruh variabel yang digunakan menunjukkan nilai yang di bawah masing-masing rata-rata variabel tersebut. Hal tersebut mengartikan data yang digunakan bersifat homogen atau dengan kata lain penyebaran nilainya merata dan berada pada tingkat ketimpangan data yang cukup rendah.

### B. Pola Hubungan antara APK SMA/ sederajat dengan Faktor-Faktor yang Diduga Memengaruhinya

Pola hubungan yang terbentuk antara APK SMA/ sederajat ( $y$ ) dengan pengeluaran per kapita, persentase penduduk miskin, transfer daerah bidang pendidikan, rasio murid-guru, dan rasio murid-sekolah diidentifikasi secara visual menggunakan *scatterplot* pada Gambar 1.

Berdasarkan Gambar 1 diperoleh informasi bahwa pola hubungan antara APK SMA/ sederajat dengan masing-masing variabel prediktor tidak membentuk suatu pola tertentu. Hal ini mengindikasikan bahwa seluruh variabel yang diduga berpengaruh terhadap APK SMA/ sederajat termasuk dalam komponen nonparametrik. Oleh karena itu, regresi nonparametrik *spline truncated* dapat digunakan untuk melakukan dalam pemodelan.

### C. Pemilihan Titik Knot Optimal

Model terbaik merupakan model yang memiliki nilai GCV paling kecil. Berikut merupakan perbandingan nilai GCV dari masing-masing banyaknya titik knot.

Dari Tabel 4 dapat disimpulkan bahwa nilai GCV paling kecil, yaitu pada kombinasi knot (1,3,1,3,3), sehingga kombinasi knot (1,3,1,3,3) digunakan sebagai model terbaik regresi nonparametrik *spline truncated* dengan jumlah parameter model sebanyak 16 termasuk konstanta  $\beta_0$ .

### D. Penaksiran Parameter

Penaksiran parameter menggunakan metode *Ordinary Least Square* (OLS). Berikut merupakan hasil estimasi parameter model regresi nonparametrik *spline truncated* dengan kombinasi titik knot (1,3,1,3,3) sebagai titik knot optimal. Satu titik knot pada variabel pengeluaran per kapita, tiga titik knot pada variabel persentase penduduk miskin, satu titik knot pada variabel transfer daerah bidang pendidikan, tiga titik knot pada variabel rasio murid-guru, dan satu titik knot pada variabel rasio murid-sekolah.

$$\hat{y} = 354,767 + 13,771x_1 - 12,6725(x_1 - 7,434)_+ + 13,338x_2 - 42,568(x_2 - 15,247)_+ + 41,763(x_2 - 19,282)_+ - 12,326(x_2 - 26,006)_+ - 0,745x_3 + 1,624(x_3 - 49,243)_+ - 111,09x_4 + 54,664(x_4 - 8)_+ + 30,676(x_4 - 9,714)_+ + 30,596(x_4 - 12,571)_+ + 3,496x_5 + 2,069(x_5 - 77,857)_+ - 5,605(x_5 - 120,592)_+ + 0,046(x_5 - 191,816)_+$$

### E. Pengujian Signifikansi Parameter

Terdapat dua tahapan dalam melakukan pengujian signifikansi parameter. Pertama adalah pengujian secara serentak, kemudian apabila kesimpulan dari uji serentak menunjukkan bahwa terdapat minimal satu parameter yang signifikan, maka dilanjutkan ke pengujian secara parsial/individu. Pengujian serentak dilakukan dengan menggunakan statistik uji  $F$  dengan hipotesis sebagai berikut.

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_{16} = 0$$

$$H_1 : \text{Minimal terdapat satu } \beta_j \neq 0, j = 1, 2, \dots, 16$$

Dengan menggunakan taraf kepercayaan 95% diperoleh nilai  $F_{(0,05;16;12)} = 2,5988$ . Berikut hasil pengujian parameter secara serentak yang ditampilkan pada Tabel 5.

Tabel 5 menunjukkan bahwa nilai statistik uji  $F_{hitung}$  sebesar 17,38 dengan  $p\text{-value} = 6,676614 \times 10^{-6}$ . Karena nilai statistik uji  $F_{hitung} > F_{(0,05;16;12)}$  dan nilai  $p\text{-value} < 0,05$  sehingga diperoleh keputusan tolak  $H_0$ . Artinya terdapat minimal satu parameter yang berpengaruh signifikan terhadap APK SMA/ sederajat di Papua. Berdasarkan hal tersebut, maka dapat dilanjutkan untuk dilakukan uji parameter secara parsial. Berikut hipotesis untuk melakukan uji parameter secara parsial.

$$H_0 : \beta_j = 0$$

$$H_1 : \beta_j \neq 0, j = 1, 2, \dots, 16$$

Dengan menggunakan taraf kepercayaan 95% diperoleh  $t_{(0,025;12)} = 2,179$ . Hasil pengujian parameter secara parsial ditampilkan pada Tabel 6.

Tabel 6 menunjukkan bahwa hasil pengujian individu dengan menggunakan nilai  $|t_{hitung}|$  dari ke-16 parameter, terdapat enam parameter yang tidak berpengaruh signifikan karena  $|t_{hitung}| < t_{(0,025;12)}$  dan  $p\text{-value} > 0,05$ , yaitu parameter  $\beta_7, \beta_{10}, \beta_{11}, \beta_{14}, \beta_{15}$ , dan  $\beta_{16}$ . Jika terdapat satu parameter yang signifikan dalam satu variabel prediktor maka variabel tersebut dikatakan berpengaruh signifikan terhadap variabel respon walaupun parameter lain tidak signifikan. Sehingga seluruh variabel prediktor berpengaruh signifikan terhadap APK SMA/ sederajat di Papua secara parsial.

#### F. Pengujian Asumsi Residual

Terdapat tiga asumsi residual yang harus dipenuhi sebuah model dalam analisis regresi nonparametrik *spline truncated*, yaitu asumsi residual identik, independen, dan berdistribusi normal.

##### 1) Asumsi Identik

Pengujian asumsi residual identik dilakukan untuk mengetahui apakah varians residual telah homogen atau tidak terjadi heterokedastisitas. Pengujian asumsi residual identik dilakukan menggunakan uji *Glejser* dengan hipotesis sebagai berikut.

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

$$H_1 : \text{minimal ada satu } \sigma_i^2 \neq \sigma^2 ; i = 1, 2, K, 29$$

Hasil pengujian asumsi identik dengan uji *Glejser* ditampilkan pada Tabel 7.

Berdasarkan Tabel 7 diperoleh nilai statistik uji  $F_{hitung} (0,26326) < F_{(0,05;16;12)} (2,5988)$  sehingga diperoleh keputusan gagal tolak  $H_0$ . Hal ini diperkuat dengan nilai  $p\text{-value}$  yang lebih besar dari signifikansi  $\alpha (0,05)$ , yaitu sebesar 0,9927. Sehingga dapat diartikan bahwa residual model regresi nonparametrik *spline* tidak terjadi heterokedastisitas dan asumsi residual identik terpenuhi.

##### 2) Asumsi Independen

Pengujian asumsi independen digunakan untuk mendeteksi terjadinya autokorelasi antar residual dari model. Pada penelitian ini, pengujian asumsi independen dilakukan dengan menggunakan plot *Autocorrelation Function (ACF)*.

Pada plot *Autocorrelation Function (ACF)* apabila terdapat minimal satu nilai ACF yang keluar dari interval konfidensi, maka artinya residual tidak independen. Diperoleh hasil plot ACF dari residual Gambar 2 berikut ini.

Berdasarkan plot ACF dari residual pada Gambar 2 diindikasikan bahwa tidak ada nilai autokorelasi pada *lag* berapapun yang keluar batas interval konfidensi atau dapat dikatakan tidak terdapat autokorelasi antar residual. Oleh karena itu, dapat disimpulkan bahwa residual telah memenuhi asumsi independen..

##### 3) Asumsi Distribusi Normal

Pengujian asumsi residual distribusi normal dilakukan dengan menggunakan uji *Kolmogorov-Smirnov*. Berikut merupakan hipotesis uji *Kolmogorov-Smirnov*.

$$H_0 : F_n(\epsilon) = F_0(\epsilon)$$

$$H_1 : F_n(\epsilon) \neq F_0(\epsilon)$$

Hasil uji *Kolmogorov-Smirnov* disajikan dalam bentuk plot dan  $p\text{-value}$  yang ditunjukkan pada Gambar 3. Titik plot residual yang ditunjukkan pada Gambar 3 menunjukkan penyebaran titik plot residual berada di sekitar garis normalitas, sehingga secara visual dapat dikatakan bahwa residual memenuhi asumsi normalitas. Namun, perlu dilakukan pengujian untuk mendukung hal tersebut. Pada Gambar 3 dapat diketahui nilai  $D$  sebesar 0,079 dan  $p\text{-value} > 0,150$ . Hal ini menghasilkan keputusan gagal tolak  $H_0$  karena  $|D| < D_\alpha (0,24)$  dan  $p\text{-value} > \alpha (0,05)$ . Sehingga dapat disimpulkan bahwa residual berdistribusi normal atau asumsi normalitas residual terpenuhi.

#### G. Kebaikan Model ( $R^2$ )

Nilai koefisien determinasi menunjukkan kebaikan model yang digunakan. Nilai koefisien determinasi yang dihasilkan model terbaik sebesar 95,86%. Artinya, model regresi nonparametrik *spline truncated* dengan titik kombinasi (1,3,1,3,3) mampu menjelaskan variabilitas APK SMA/ sederajat di Papua sama dengan 95,86%. Atau, APK SMA/ sederajat di Papua dapat dijelaskan kelima variabel prediktor yang digunakan dalam penelitian ini sama dengan 95,86%, sedangkan sisanya dapat dijelaskan oleh variabel-variabel lain yang tidak terdapat dalam model.

#### H. Interpretasi Model

Model terbaik regresi nonparametrik *spline truncated* pada APK SMA/ sederajat di Provinsi Papua tahun 2018 merupakan model yang terdiri dari lima variabel prediktor dengan menggunakan titik kombinasi 1,3,1,3,3. Berikut adalah interpretasi model pada setiap variabel yang signifikan terhadap APK SMA/ sederajat di Provinsi Papua.

##### 1) Model Pertama

Dengan mengasumsikan semua variabel selain variabel  $x_1$  adalah konstan, maka pengaruh variabel pengeluaran per kapita ( $x_1$ ) terhadap APK SMA/ sederajat di Provinsi Papua adalah sebagai berikut.

$$\hat{y} = 354,767 + 23,771x_1 - 12,6725(x_1 - 7,434)_+$$

$$= \begin{cases} 354,767 + 13,771x_1 ; & x_1 < 7,434 \\ 448,974 + 1,0985x_1 ; & x_1 \geq 7,434 \end{cases}$$

Berdasarkan model yang diperoleh dapat dijelaskan bahwa apabila kabupaten/kota di Provinsi Papua dengan pengeluaran per kapita kurang dari Rp7,434 juta dan mengalami penambahan Rp1 juta dengan asumsi variabel lain tetap, maka APK SMA/ sederajat akan naik sebesar 13,771%. Terdapat 19 kabupaten yang termasuk pada interval ini. Selanjutnya, apabila kabupaten/kota di Papua memiliki pengeluaran per kapita lebih dari Rp7,434 juta dan mengalami kenaikan Rp1 juta dengan asumsi variabel lain tetap maka APK SMA/ sederajat akan naik 1,099%. Terdapat 10 kabupaten/kota yang masuk ke dalam interval kedua.

##### 2) Model Kedua

Semua variabel selain variabel  $x_2$  diasumsikan konstan, maka pengaruh persentase penduduk miskin ( $x_2$ ) terhadap APK SMA/ sederajat di Provinsi Papua adalah sebagai berikut

$$\hat{y} = 354,767 + 13,338x_2 - 42,568(x_2 - 15,247)_+ + 41,763(x_2 - 19,282)_+$$

$$= \begin{cases} 354,767 + 13,338x_2; & x_2 < 15,247 \\ 1103,8 - 29,32x_2; & 15,247 \leq x_2 < 19,282 \\ 198,527 - 12533x_2; & 19,282 \leq x_2 < 26,006 \\ 519,077 + 0,207x_2; & x_2 \geq 26,006 \end{cases}$$

Berdasarkan model tersebut dapat dijelaskan bahwa jika suatu kabupaten/kota memiliki persentase penduduk miskin kurang besar dari 15,25% persen dan mengalami pertambahan satu persen, maka persentase APK SMA/ sederajat cenderung naik 13,34%. Kabupaten/kota yang termasuk dalam interval ini antara lain: Kabupaten Jayapura, Merauke, Mimika, Sarmi, dan Kota Jayapura. Selanjutnya, pada interval kedua, hanya ada satu kabupaten yang menunjukkan apabila persentase penduduk miskin di suatu kabupaten/kota mengalami kenaikan sebesar 1% akan mengakibatkan APK SMA/ sederajat menurun sebesar 29,23%. Kabupaten tersebut adalah Kabupaten Keerom. Selanjutnya pada interval ketiga apabila persentase penduduk miskin di kabupaten/kota mengalami kenaikan sebesar 1%, maka APK SMA/ sederajat di kabupaten/kota tersebut cenderung naik 12,53%. Kabupaten Biak Numfor, Nabire, Boven Digoel, dan Mappi adalah kabupaten yang tergolong dalam interval ini. Kemudian pada interval keempat, apabila persentase penduduk miskin di kabupaten/kota lebih dari 26,006% dan mengalami pertambahan sebesar 1%, maka APK SMA/ sederajat di kabupaten/kota tersebut cenderung naik sebesar 0,207%. Terdapat 19 kabupaten/kota yang termasuk dalam interval ketiga.

3) Model Ketiga

Dengan mengasumsikan semua variabel selain variabel  $x_3$  adalah konstan, maka pengaruh variabel transfer daerah bidang pendidikan ( $x_3$ ) terhadap APK SMA/ sederajat di Provinsi Papua adalah sebagai berikut.

$$\hat{y} = 354,76 - 0,745x_3 + 1,624(x_3 - 43,243)_+ \\ = \begin{cases} 354,767 - 0,745x_3; & x_3 < 49,243 \\ 274,796 + 0,879x_3; & x_3 \geq 49,243 \end{cases}$$

Berdasarkan model tersebut, dapat diinterpretasikan apabila transfer daerah bidang pendidikan di kabupaten/kota lebih dari Rp49,243 miliar dan mengalami pertambahan sebesar Rp1 miliar, maka APK SMA/ sederajat di kabupaten/kota tersebut cenderung naik 0,879%. Kabupaten/kota yang tergolong dalam interval ini antara lain: Kabupaten Jayapura, Biak Numfor, Kep. Yapen, Merauke, Jayawijaya, Nabire, Mimika, Keerom, dan Kota Jayapura.

4) Model Keempat

Semua variabel selain variabel  $x_4$  diasumsikan konstan, maka pengaruh persentase rasio murid-guru ( $x_4$ ) terhadap APK SMA/ sederajat di Provinsi Papua adalah sebagai berikut.

$$\hat{y} = 354,767 - 111,09x_4 + 54,664(x_4 - 8)_+ + 30,676(x_4 - 9,714)_+ - 30,596(x_4 - 12,571)_+ \\ = \begin{cases} 354,767 - 111,09x_4; & x_4 < 8 \\ -82,545 - 56,426x_4; & 8 \leq x_4 < 9,714 \\ -380,532 - 25,75x_4; & 9,714 \leq x_4 < 12,571 \\ -765,154 + 4,846x_4; & x_4 \geq 12,571 \end{cases}$$

Pada interval pertama menunjukkan bahwa apabila rasio murid-guru kabupaten/kota kurang dari 8 maka setiap

kenaikan satu satuan akan menurunkan APK SMA/ sederajat sebesar 111,09%. Terdapat tiga kabupaten yang termasuk interval ini, yaitu Kabupaten Nduga, Mamberamo Tengah, dan Intan Jaya. Kemudian pada interval keempat, apabila rasio murid-guru di kabupaten/kota lebih dari 12,571 serta mengalami pertambahan sebesar satu satuan, maka APK SMA/ sederajat di kabupaten/kota tersebut cenderung naik sebesar 4,846%. Terdapat 14 kabupaten/kota yang berada pada interval keempat.

5) Model Kelima

Dengan mengasumsikan semua variabel selain variabel  $x_5$  adalah konstan, maka pengaruh variabel rasio murid-sekolah ( $x_5$ ) terhadap APK SMA/ sederajat di Provinsi Papua adalah sebagai berikut.

$$\hat{y} = 354,767 + 3,496x_5 + 2,069(x_5 - 77,257)_+ - 5,605(x_5 - 120,592)_+ - 0,046(x_5 - 191,816)_+ \\ = \begin{cases} 354,767 + 3,496x_5; & x_5 < 77,857 \\ 193,68 + 5,565x_5; & 77,857 \leq x_5 < 120,592 \\ 869,599 - 0,04x_5; & 120,592 \leq x_5 < 191,816 \\ 869,645x_5; & x_5 \geq 191,816 \end{cases}$$

Berdasarkan model yang diperoleh dapat dijelaskan bahwa apabila rasio murid-sekolah kabupaten/kota kurang dari 77,857 maka setiap kenaikan rasio murid-sekolah satu satuan akan meningkatkan APK SMA/ sederajat sebesar 3,496%. Kabupaten yang termasuk ke dalam interval pertama, yaitu Kabupaten Nduga, Mamberamo Tengah, dan Puncak.

Secara visual pengelompokkan wilayah untuk masing-masing variabel berdasarkan titik knot dapat dilihat melalui Gambar 4.

V. KESIMPULAN

APK SMA/ sederajat di Provinsi Papua memiliki nilai rata-rata sebesar 63,26%. Terdapat 17 kabupaten/kota dengan nilai APK SMA/ sederajat di bawah rata-rata, hal itu menunjukkan bahwa lebih dari setengah wilayah Provinsi Papua memiliki nilai APK SMA/ sederajat di bawah rata-rata. Sebaran data APK SMA/ sederajat menunjukkan nilai yang di bawah rata-rata, yaitu 32,33%. Hal tersebut mengartikan data APK SMA/ sederajat pada masing-masing wilayah kabupaten/kota di Provinsi Papua bersifat homogen atau dengan kata lain penyebaran nilainya merata dan berada pada tingkat ketimpangan data yang cukup rendah. Pada tahun 2018, di Kabupaten Mimika murid yang menempuh studi pada jenjang SMA/ sederajat melebihi jumlah anak usia 16-18 tahun yang bertempat tinggal di Kabupaten Mimika. Kabupaten Mimika merupakan wilayah dengan nilai APK SMA/ sederajat yang tertinggi di Provinsi Papua yaitu sebesar 129,86%, sedangkan nilai APK SMA/ sederajat terendah di Provinsi Papua, yaitu Kabupaten Puncak sebesar 9,40%.

Model terbaik yang digunakan untuk memodelkan APK SMA/ sederajat di Papua tahun 2018 ialah model regresi nonparametrik *spline truncated* dengan titik kombinasi (1,3,1,3,3). Semua variabel yang digunakan dalam penelitian ini berpengaruh secara signifikan terhadap APK SMA/ sederajat di Papua, yaitu variabel pengeluaran per kapita, persentase penduduk miskin, transfer daerah bidang pendidikan, rasio murid-guru, dan rasio murid-sekolah. Berikut merupakan model regresi *spline* yang diperoleh.

$$\hat{y} = 354,767 + 13,771x_1 - 12,6725(x_1 - 7,434)_+ + 13,338x_2 - 42,568(x_2 - 15,247)_+ + 41,763(x_2 - 19,282)_+ - 12,326(x_2 - 26,006)_+ - 0,745x_3 + 1,624(x_3 - 49,243)_+ - 111,09x_4 + 54,664(x_4 - 8)_+ + 30,676(x_4 - 9,714)_+ + 30,596(x_4 - 12,571)_+ + 3,496x_5 + 2,069(x_5 - 77,857)_+ - 5,605(x_5 - 120,592)_+ + 0,046(x_5 - 191,816)_+$$

Model telah memenuhi ketiga asumsi residual. Nilai koefisien determinasi dari model tersebut sama dengan 95,86%.

#### DAFTAR PUSTAKA

[1] Badan Pusat Statistik Provinsi Papua, *Indikator Pendidikan Provinsi*

*Papua Tahun 2018*. Jayapura: Badan Pusat Statistika Provinsi Papua, 2020.

- [2] R. L. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. Florida: CRC Press, 1999.
- [3] R. E. Walpole, *Pengantar Statistika*. Jakarta: Gramedia Pustaka Utama, 1995.
- [4] G. Wahba, *Spline Model for Observational Data*. Philadelphia: Society for Industrial and Applied Mathematics, 1990.
- [5] W. Härdle, *Applied Nonparametric Regression*. New York: Cambridge University Press, 1992.
- [6] I. N. Budiantara, "Model spline dengan knots optimal," *J. Ilmu Dasar*, vol. 7, no. 6, 2006.
- [7] N. Draper and H. Smith, *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama, 1992.
- [8] D. C. Porter and D. N. Gujarati, *Basic Econometrics*. New York: McGraw-Hill Education, 2008.
- [9] W. W. Daniel, *Statistik Nonparametrik*. Jakarta: Gramedia Pustaka Utama, 1989.