

Pemodelan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat Menggunakan Regresi Nonparametrik *Spline Truncated*

Novia Asri Kurniawati dan I Nyoman Budiantara¹⁾

Departemen Statistika, Fakultas Matematika, Komputasi, dan Sains Data

Institut Teknologi Sepuluh Nopember (ITS)

e-mail: i_nyoman_b@statistika.its.ac.id

Abstrak—Tingkat Pengangguran Terbuka (TPT) merupakan indikator yang digunakan untuk menilai kesejahteraan sosial masyarakat. Badan Pusat Statistika (BPS) tahun 2018 merilis data TPT di Indonesia menurut provinsi yang menunjukkan Provinsi Jawa Barat memiliki TPT tertinggi yaitu sebesar 8,22. Terdapat empat faktor dari indikator sosial kependudukan yang diduga berpengaruh terhadap tingkat pengangguran terbuka di Provinsi Jawa Barat tahun 2017 yaitu Tingkat Partisipasi Angkatan Kerja (TPAK), Dependency Ratio, Rata-rata Lama Sekolah dan Laju Pertumbuhan Ekonomi. Berdasarkan analisa menggunakan scatterplot data yang telah diperoleh dari Badan Pusat Statistika Provinsi Jawa Barat tahun 2017 menunjukkan tidak adanya pola hubungan antara Tingkat Pengangguran Terbuka (TPT) dengan faktor-faktor yang diduga berpengaruh sehingga pada penelitian ini digunakan metode regresi nonparametrik *spline truncated*. Hasil analisa menunjukkan bahwa variabel TPAK, Dependency Ratio, Rata-rata Lama Sekolah dan Laju Pertumbuhan Ekonomi berpengaruh secara signifikan terhadap Tingkat Pengangguran Terbuka (TPT) di Provinsi Jawa Barat tahun 2017 dengan nilai koefisien determinasi 95,23.

Kata Kunci—GCV, Koefisien Determinasi, Pengangguran Terbuka, Regresi Nonparametrik *Spline*.

I. PENDAHULUAN

TERCAPAINYA kesejahteraan masyarakat merupakan wujud dari pemerintah Indonesia dalam meningkatkan pembangunan nasional. Salah satu upaya pemerintah dalam meningkatkan kesejahteraan adalah meningkatkan stabilitas nasional, memacu pertumbuhan ekonomi, meningkatkan iklim investasi, dan menekan angka pengangguran. Pengangguran merupakan beban pekerjaan penting yang harus segera ditangani oleh pemerintah, khususnya Dinas Tenaga Kerja. Tingginya angka pengangguran di Indonesia disebabkan karena tidak ada kesesuaian antara penawaran tenaga kerja dengan kebutuhan di pasar tenaga kerja. Serta jumlah penduduk yang semakin meningkat yang tidak diimbangi dengan pertumbuhan lapangan usaha yang ada.

Jawa Barat merupakan provinsi dengan jumlah penduduk terbesar di Indonesia. Pada tahun 2017 penduduk Jawa Barat sudah mencapai 48,03 juta jiwa. Badan Pusat Statistika merilis jumlah pengangguran di Jawa Barat dari tahun 2014 hingga tahun 2017 mengalami fluktuasi. Tahun 2017 Provinsi Jawa Barat menunjukkan wilayah dengan tingkat pengangguran terbuka sebesar 8,22. Artinya dari 100 orang angkatan kerja Jawa Barat, ada sekitar 8-9 orang yang belum terserap di pasar kerja. Meskipun mengalami penurunan dari tahun sebelumnya, namun Jawa Barat merupakan wilayah dengan rata-rata tingkat pengangguran terbuka tertinggi di Indonesia. Hal ini ditunjukkan dengan

jumlah angkatan kerja sebesar 22,391 juta jiwa. Permasalahan ini tentunya menjadi perih yang perlu diperhatikan, karena nilai tersebut sangat jauh dari tingkat pengangguran terbuka nasional sendiri yang bernilai 5,34. Meningkatnya jumlah pengangguran merupakan salah satu masalah sosial yang harus dihadapi. Peningkatan tersebut disebabkan oleh faktor-faktor yang mempengaruhi. Menurut penelitian tentang pengangguran yang dilakukan Sari [1], Astuti [2], dan Wijaya [3] merupakan indikator yang berpengaruh yaitu indikator sosial, ekonomi, dan pendidikan.

Tujuan dari penelitian ini adalah untuk mengetahui variabel yang berpengaruh signifikan/nyata terhadap pengangguran terbuka di Jawa Barat dengan menggunakan pendekatan regresi *spline*. Regresi nonparametrik *spline* merupakan metode yang tepat untuk digunakan karena *spline* merupakan salah satu potongan polinomial yang memiliki sifat fleksibilitas yang lebih baik jika dibandingkan dengan polinomial biasa. Selain memiliki sifat fleksibilitas tersebut, *spline* memiliki kemampuan untuk menganalisa data pada sub-sub interval tertentu mengalami perubahan perilaku. Model terbaik dari regresi nonparametrik *spline* adalah model yang memiliki nilai *generalized cross validation* (GCV) minimum yang diperoleh dari nilai terkecil beberapa knot dan kombinasi knot.

II. TINJAUAN PUSTAKA

A. Statistika Deskriptif

Statistika Deskriptif merupakan metode pengumpulan dan penyajian suatu gugus data sehingga memberikan informasi yang berguna. Metode ini bertujuan untuk menguraikan tentang karakteristik dari suatu keadaan dan membuat deskripsi atau gambaran yang sistematis dan akurat mengenai fakta-fakta dan sifat-sifat dari fenomena yang diselidiki [4].

B. Regresi Nonparametrik *Spline Truncated*

Regresi nonparametrik merupakan salah satu model regresi yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor yang tidak diketahui bentuk kurva regresinya [5]. Model regresi nonparametrik dapat ditulis sebagai berikut.

$$y_i = f(x_i) + \varepsilon_i, i = 1, 2, 3, \dots, n \quad (1)$$

dimana y_i adalah variabel respon ke- i , $f(x_i)$ adalah fungsi regresi yang tidak diketahui bentuk kurva regresinya, dan residual random dinyatakan $\varepsilon_i \square IIDN(0, \sigma^2)$. Fungsi

spline truncated diperoleh dari hasil penjumlahan antara fungsi polinomial dengan fungsi truncated. Misal fungsi spline truncated berorde p dengan titik knot K_1, K_2, \dots, K_r , sehingga kurva regresi yang terbentuk yaitu $f(x_i)$, lebih rinci dapat dituliskan menjadi persamaan sebagai berikut:

$$f(x_i) = \sum_{j=0}^p \beta_j x_i^j + \sum_{k=1}^r \beta_{p+r} (x_i - K_r)_+^p \quad (2)$$

dimana

$$y_i = \sum_{j=0}^p \beta_j x_i^j + \sum_{k=1}^r \beta_{p+r} (x_i - K_r)_+^p + \varepsilon_i ; i = 1, 2, \dots, n$$

dengan,

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

Titik $x_i = K_r$ adalah titik knot yang menggambarkan pola perubahan fungsi pada sub interval yang berbeda dan nilai p adalah derajat polinomial.

Estimasi parameter model regresi nonparametrik spline dilakukan dengan menggunakan metode *Ordinary Least Square* (OLS) sebagai berikut.

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

C. Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan dimana terdapat perubahan perilaku pola kurva pada interval yang berlainan [6]. Salah satu metode yang sering digunakan untuk memilih titik knot optimal adalah metode *Generalized Cross Validation* (GCV). Model regresi spline terbaik diperoleh dari titik knot optimal dengan melihat nilai GCV terkecil. Metode GCV dapat dituliskan sebagai berikut.

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

Dimana \mathbf{I} merupakan matriks identitas, n adalah pengamatan, $K = (K_1, K_2, \dots, K_r)$ merupakan titik-titik knot,

$$MSE(K) = n^{-1} \sum_{i=1}^n (y_i - \hat{f}(x_i))^2 \quad (6)$$

Serta $\mathbf{A} = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T$.

D. Pengujian Parameter Model

1. Pengujian Secara Serentak

Pengujian secara serentak dilakukan untuk mengetahui signifikansi parameter model regresi secara bersama-sama. Hipotesis

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

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

Nilai $p+r$ merupakan banyak parameter dalam model regresi nonparametrik spline kecuali β_0 .

Statistik uji yang digunakan adalah sebagai berikut.

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} \quad (7)$$

Tolak H_0 jika $F_{hitung} > F_{\alpha, (p+r, n-(p+r)-1)}$ atau $p\text{-value} < \alpha$ [7].

2. Pengujian Secara Parsial

Pengujian secara individu berfungsi untuk mendeteksi apakah parameter secara individual mempunyai pengaruh yang signifikan terhadap variabel respon. Hipotesis pengujian secara parsial sebagai berikut.

$$H_0 : \beta_j = 0$$

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

Pengujian secara individu dilakukan dengan menggunakan uji t [7]. Statistik uji yang digunakan adalah sebagai berikut:

$$t_{hitung} = \frac{\hat{\beta}_j}{SE(\hat{\beta}_j)} \quad (8)$$

dengan

$$\text{Var}(\hat{\beta}_j) = \text{var} \left[(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \right] \quad (9)$$

Tolak H_0 jika $|t_{hitung}| > t_{\alpha/2; (n-(p+r)-1)}$ atau $p\text{-value} < \alpha$.

Sehingga menghasilkan kesimpulan bahwa parameter ke- j berpengaruh signifikan. R^2 digunakan sebagai indikator kebaikan model, yang diberikan sebagai berikut.

$$R^2 = \frac{SS_{Regresi}}{SS_{total}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (10)$$

E. Pengujian Asumsi Residual

1. Asumsi Identik

Asumsi identik terpenuhi apabila varians antar residual sama, atau tidak terjadi heteroskedastisitas [8].

$$\text{var}(y_i) = \text{var}(\varepsilon_i) = \sigma^2 ; i = 1, 2, \dots, n \quad (11)$$

Uji identik dapat menggunakan uji Glejser. Hipotesis yang digunakan adalah sebagai berikut:

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

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

Statistik uji yang digunakan adalah,

$$F_{hitung} = \frac{\left[\sum_{i=1}^n (|\varepsilon_i| - |\bar{\varepsilon}|^2) \right] / v-1}{\left[\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|^2) \right] / v-1} \quad (12)$$

Tolak H_0 jika $F_{hitung} > F_{\alpha; (v-1, n-v)}$ atau $p\text{-value} < \alpha$ dimana nilai v adalah banyaknya parameter model *Glejser*.

2. Asumsi Independen

Uji independent digunakan untuk mengetahui ada tidaknya korelasi antar residual. Salah satu cara untuk mendeteksi residual bersifat independen atau tidak dengan menggunakan plot *Autocorrelation Function* (ACF) [9].

Interval konfidensi untuk parameter ρ_k diberikan oleh.

$$-t_{n-1; \alpha/2} SE(\hat{\rho}_k) < \rho_k < t_{n-1; \alpha/2} SE(\hat{\rho}_k) \quad (13)$$

dimana $SE(\hat{\rho}_k) = \sqrt{\frac{1}{n} (1 + \rho_1^2 + \dots + 2\rho_{k-1}^2)}$. Apabila tidak

terdapat nilai autokorelasi yang keluar dari batas signifikansi menunjukkan asumsi independen terpenuhi.

3. Distribusi Normal

Pengujian asumsi berdistribusi normal dilakukan untuk mengetahui apakah residual suatu data telah mengikuti distribusi normal. Uji yang digunakan adalah uji *Kolmogorov-Smirnov* [10].

Hipotesis:

$H_0 : F_0(x) = F(x)$ (residual mengikuti distribusi normal)

$H_1 : F_0(x) \neq F(x)$ (residual tidak mengikuti distribusi normal)

Statistik uji yang digunakan sebagai berikut.

$$D = \sup_x |F_n(x) - F_0(x)| \quad (14)$$

Tolak H_0 apabila $D > D_\alpha$ atau $p\text{-value} < \alpha$.

F. Pengangguran Terbuka

Pengangguran merupakan keadaan seseorang belum mendapatkan pekerjaan dimana seseorang tersebut termasuk dalam angkatan kerja. Tingkat pengangguran terbuka adalah persentase jumlah pengangguran terhadap jumlah angkatan kerja. Tingkat pengangguran terbuka sebagai indikasi besarnya persentase angkatan kerja yang termasuk dalam pengangguran. Tingkat pengangguran yang tinggi menunjukkan bahwa terdapat banyak angkatan kerja yang tidak terserap pada pasar kerja [11].

III. METODOLOGI PENELITIAN

A. Sumber Data

Data yang digunakan pada penelitian ini adalah data sekunder yang diambil dari Badan Pusat Statistika Provinsi Jawa Barat tahun 2017. Unit penelitian yang digunakan merupakan 27 kabupaten dan kota di Provinsi Jawa Barat.

B. Variabel Penelitian

Variabel respon yang digunakan dalam penelitian ini adalah tingkat pengangguran terbuka Provinsi Jawa Barat, sedangkan variabel prediktor yang diduga berpengaruh adalah sebagai berikut.

Tabel 1. Variabel Penelitian

Variabel	Keterangan	Skala
Y	Tingkat Pengangguran Terbuka	Rasio
X ₁	Tingkat Partisipasi Angkatan Kerja	Rasio
X ₂	Dependency Ratio	Rasio
X ₃	Rata-rata Lama Sekolah	Rasio
X ₄	Laju Pertumbuhan PDRB	Rasio

C. Langkah Analisis

Langkah analisis yang digunakan pada penelitian ini adalah sebagai berikut.

1. Melakukan analisis statistika deskriptif terhadap variabel-variabel penelitian.
2. Membuat *scatterplot* antara Tingkat Pengangguran Terbuka di Provinsi Jawa Barat dengan masing-masing variabel yang diduga untuk mengetahui bentuk pola data.
3. Memodelkan data menggunakan regresi nonparametrik spline truncated linear dengan satu, dua, tiga, dan kombinasi knot.
4. Menentukan titik knot paling optimal berdasarkan nilai GCV paling minimum.
5. Mendapatkan model regresi spline dengan titik knot paling optimal.
6. Menguji signifikansi parameter regresi spline secara serentak
7. Melakukan uji parameter regresi spline secara parsial.
8. Melakukan uji asumsi residual IIDN dari model regresi spline.
9. Menginterpretasikan model dan menarik kesimpulan.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Tingkat Pengangguran Terbuka

Berikut merupakan karakteristik data dari empat faktor yang diduga berpengaruh terhadap tingkat pengangguran terbuka di Provinsi Jawa Barat.

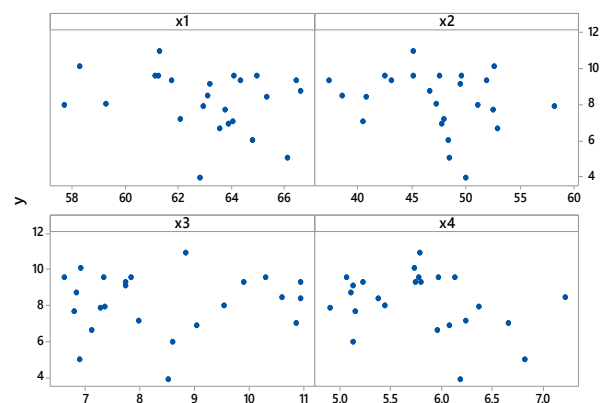
Tabel 2. Statistika Deskriptif

Variabel	Rata-rata	Variansi	Min	Maks
y	8,168	2,801	3,920	10,970
x ₁	63,021	5,624	57,690	66,600
x ₂	47,210	25,080	37,270	58,100
x ₃	8,432	2,189	6,610	10,930
x ₄	5,780	0,370	4,890	7,210

Berdasarkan Tabel 2, Tingkat Pengangguran Terbuka (Y) tertinggi pada tahun 2017 adalah Kabupaten Bekasi yaitu sebesar 10,97. Sedangkan kabupaten/kota yang memiliki TPT terendah di Kabupaten Pangandaran yaitu sebesar 3,340. Rata-rata TPT tiap kabupaten/kota di Provinsi Jawa Barat selama tahun 2017 adalah sebesar 8,168 dengan variansi sebesar 2,801. Tingkat Partisipasi Angkatan Kerja (X₁) terendah terdapat di Kabupaten Depok dan TPAK tertinggi terdapat pada Kabupaten Cianjur. *Dependency Ratio* (X₂) terendah sebesar 37,27 di Kota Subang dan tertinggi di Kabupaten Bogor. Rata-rata lama sekolah (X₃) Kabupaten Sukabumi mengenyam pendidikan lebih lama dibandingkan dengan Kabupaten Cirebon. Hal tersebut dapat dilihat dari tingginya angka rata-rata lama sekolah di Kota Kabupaten Sukabumi daripada di Kabupaten Cirebon. Laju pertumbuhan PDRB (X₄) terendah di Kabupaten Bogor dan laju pertumbuhan PDRB tertinggi terdapat di Kota Banjar.

B. Pola Data Antara Tingkat Pengangguran Terbuka dengan Variabel yang Diduga Mempengaruhi

Langkah pertama dalam melakukan analisis regresi adalah membuat *scatterplot* untuk mengetahui pola hubungan variabel prediktor terhadap variabel respon yaitu tingkat pengangguran terbuka. Gambar 1 menunjukkan bahwa pola hubungan antara tingkat pengangguran terbuka dengan seluruh variabel prediktor tidak membentuk pola tertentu, sehingga metode yang digunakan adalah regresi nonparametrik spline.



Gambar 1. *Scatterplot* Variabel Prediktor Terhadap Respon.

C. Pemodelan Tingkat Pengangguran Terbuka

Penelitian ini akan digunakan pemodelan tingkat pengangguran di Jawa Barat menggunakan pendekatan regresi nonparametrik *spline truncated linear*. Model *spline* yang digunakan adalah model dengan satu, dua, tiga, serta

kombinasi titik knot. Berikut merupakan model estimasi regresi nonparametrik *spline truncated* dengan satu titik knot menggunakan empat variabel prediktor.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+^1 + \hat{\beta}_{21}x_2 + \hat{\beta}_{22}(x_2 - K_2)_+^1 + \hat{\beta}_{31}x_3 + \hat{\beta}_{32}(x_3 - K_3)_+^1 + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_4)_+^1$$

Pemilihan titik knot optimal pada satu titik knot menghasilkan nilai GCV sebagai berikut

Tabel 3.
Nilai GCV Satu Titik Knot

GCV	Knot			
	x ₁	x ₂	x ₃	x ₄
2,723	62,418	48,323	8,902	6,121
2,778	57,872	37,695	6,698	4,937
2,780	62,236	47,898	8,814	6,074
2,823	62,600	48,748	8,990	6,168

Berdasarkan Tabel 3, diketahui bahwa nilai GCV paling minimum berdasarkan hasil iterasi untuk satu titik knot adalah sebesar 2,723 dengan rincian letak titik knot sebagai berikut.

$$K_1=62,418 ; K_2=48,323 ; K_3=8,902 ; K_4=6,121$$

Setelah mendapatkan knot optimum dari satu titik knot, langkah analisis selanjutnya dilakukan dengan pemilihan titik knot optimum menggunakan dua titik knot. Estimasi model regresi nonparametrik *spline truncated* dengan empat variabel prediktor menggunakan dua titik knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+^1 + \hat{\beta}_{13}(x_1 - K_2)_+^1 + \hat{\beta}_{21}x_2 + \hat{\beta}_{22}(x_2 - K_3)_+^1 + \hat{\beta}_{23}(x_2 - K_4)_+^1 + \hat{\beta}_{31}x_3 + \hat{\beta}_{32}(x_3 - K_5)_+^1 + \hat{\beta}_{33}(x_3 - K_6)_+^1 + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_7)_+^1 + \hat{\beta}_{43}(x_4 - K_8)_+^1$$

Berikut merupakan nilai GCV dengan menggunakan dua titik knot di setiap variabel prediktor.

Tabel 4.
Nilai GCV Dua Titik Knot

GCV	Knot			
	x ₁	x ₂	x ₃	x ₄
2,131	62,418	48,323	8,902	6,121
	62,600	48,748	8,990	6,168
2,381	57,690	37,270	6,610	4,890
	66,600	58,100	10,930	7,210
2,609	61,327	45,772	8,373	5,837
	61,509	46,197	8,461	5,884
2,659	62,236	47,898	8,814	6,074
	62,600	48,748	8,990	6,168

Tabel 4 menunjukkan bahwa nilai GCV minimum yang diperoleh adalah 2,131 dengan dua titik knot optimum untuk masing-masing variabel adalah sebagai berikut.

Tabel 5.

Letak Titik Knot Optimum dengan Dua Titik Knot

X ₁	X ₂	X ₃	X ₄
K ₁ =62,418	K ₃ =48,323	K ₅ =8,902	K ₇ =6,121
K ₂ =62,600	K ₄ =48,748	K ₆ =8,990	K ₈ =6,168

Langkah selanjutnya adalah pemilihan titik knot optimum menggunakan tiga titik knot. Estimasi model regresi nonparametrik *spline truncated* dengan tiga titik knot

menggunakan empat faktor variabel prediktor adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_{11}x_1 + \hat{\beta}_{12}(x_1 - K_1)_+^1 + \hat{\beta}_{13}(x_1 - K_2)_+^1 + \hat{\beta}_{14}(x_1 - K_3)_+^1 + \hat{\beta}_{21}x_2 + \hat{\beta}_{22}(x_2 - K_4)_+^1 + \hat{\beta}_{23}(x_2 - K_5)_+^1 + \hat{\beta}_{24}(x_2 - K_6)_+^1 + \hat{\beta}_{31}x_3 + \hat{\beta}_{32}(x_3 - K_7)_+^1 + \hat{\beta}_{33}(x_3 - K_8)_+^1 + \hat{\beta}_{34}(x_3 - K_9)_+^1 + \hat{\beta}_{41}x_4 + \hat{\beta}_{42}(x_4 - K_{10})_+^1 + \hat{\beta}_{43}(x_4 - K_{11})_+^1 + \hat{\beta}_{44}(x_4 - K_{12})_+^1$$

Nilai GCV untuk model regresi nonparametric *spline truncated* dengan tiga titik knot disajikan melalui Tabel 6.

Tabel 6.
Nilai GCV Tiga Titik Knot

GCV	Knot			
	x ₁	x ₂	x ₃	x ₄
1,251	62,600	48,748	8,990	6,168
	62,963	49,598	9,167	6,263
	63,145	50,023	9,255	6,310
	60,599	44,072	8,021	5,648
1,300	60,963	44,922	8,197	5,742
	63,872	51,723	9,608	6,500
	62,600	48,748	8,990	6,168
	62,781	49,173	9,079	6,216
1,324	63,327	50,448	9,343	6,358
	60,599	44,072	8,021	5,648
	60,963	44,922	8,197	5,742
	63,691	51,298	9,519	6,452

Tabel 6 menunjukkan bahwa nilai GCV yang dihasilkan melalui iterasi memiliki nilai paling minimum sebesar 1,251 dengan rincian letak knot sebagai berikut.

Tabel 7.

Letak Titik Knot Optimum dengan Tiga Titik Knot

X ₁	X ₂	X ₃	X ₄
K ₁ =62,600	K ₄ =48,748	K ₇ =8,990	K ₁₀ =6,168
K ₂ =62,969	K ₅ =49,598	K ₈ =9,167	K ₁₁ =6,263
K ₃ =63,145	K ₆ =50,023	K ₉ =9,255	K ₁₂ =6,310

Pemilihan titik knot dengan kombinasi titik knot perlu dilakukan karena terdapat kemungkinan jumlah titik knot optimal dari setiap variabel berbeda-beda. Pada Tabel 8 disajikan hasil nilai GCV untuk model regresi nonparametrik *spline truncated* dengan kombinasi knot.

Tabel 8.

Nilai GCV Kombinasi Titik Knot

GCV	Knot			
	x ₁	x ₂	x ₃	x ₄
0,737	62,600	48,748	8,902	6,121
	62,963	49,598	8,990	
	63,145	50,023		
	62,600	48,748	8,990	6,121
0,803	62,963	49,598	9,167	
	63,145	50,023	9,255	
	62,600	48,748	8,902	6,121
	62,963	49,598	8,990	6,168
0,859	63,145	50,023		
	62,600	48,748	8,990	6,121
	62,963	49,598	9,167	6,168
	63,145	50,023	9,255	

Informasi yang dapat diketahui berdasarkan Tabel 8 adalah nilai GCV paling minimum dihasilkan pada kombinasi titik knot (3,3,2,1) sebesar 0,737 dengan titik knot sebagai berikut.

Tabel 9.

Letak Titik Knot Optimum dengan Kombinasi Titik Knot

X ₁	X ₂	X ₃	X ₄
K ₁ =62,600	K ₄ =48,748	K ₇ =8,902	K ₉ =6,121
K ₂ =62,969	K ₅ =49,598	K ₈ =8,990	
K ₃ =63,145	K ₆ =50,023		

Setelah didapatkan hasil knot optimum yang diperoleh pada satu, dua, tiga serta kombinasi titik knot, dilakukan perbandingan sebagai dasar pemilihan model terbaik.

D. Pemilihan Titik Knot Optimal

Nilai GCV minimum pada pemilihan titik knot optimal dengan satu titik knot, dua titik knot, tiga titik knot, dan kombinasi titik knot ditampilkan sebagai berikut.

Tabel 10.
Nilai GCV Minimum Titik Knot Optimal

GCV	Jumlah Knot
2,723	1
2,131	2
1,251	3
0,737	Kombinasi (3,3,2,1)

Berdasarkan model dengan satu, dua, tiga maupun kombinasi knot, model terbaik didapatkan dari nilai GCV paling minimum yaitu pada kombinasi knot (3,3,2,1) dengan nilai GCV sebesar 0,737.

E. Pengujian Signifikansi Parameter Model Regresi Nonparametrik Spline

1. Uji Serentak

Pengujian secara serentak ini dilakukan untuk menguji estimasi parameter model secara bersamaan (simultan). Nilai α yang digunakan sebesar 0,05. Berikut ini adalah hasil analisis ragam model regresi nonparametrik *spline*.

Tabel 11.
ANOVA Model Regresi Spline

Sumber variasi	df	Sum of Square (SS)	Mean Square (MS)	F _{hitung}
Regresi	13	61,35293	4,719457	
Residual	10	3,070061	0,3070061	15,37252
Total	23	64,423	-	

Berdasarkan hasil tabel ANOVA pada Tabel 11 diperoleh nilai *p-value* sebesar 6,78E-05. Karena *p-value* < α dengan $\alpha = 0,05$ maka dapat diambil keputusan tolak H_0 yang menunjukkan bahwa minimal ada satu parameter yang berpengaruh signifikan terhadap model. Selanjutnya dilakukan pengujian secara parsial untuk melihat parameter yang berpengaruh secara signifikan.

2. Uji Parsial

Pengujian secara parsial atau individu menggunakan uji *t*. Berikut merupakan hasil pengujian parameter secara individu dari model regresi spline yang ditampilkan pada Tabel 12.

F. Penaksiran Parameter Model Regresi Nonparametrik Spline Empat Variabel Prediktor

Model tersebut merupakan model dengan nilai GCV terkecil, didapatkan model terbaik yaitu model regresi nonparametrik spline dengan tiga titik knot dengan empat variabel prediktor.

$$\hat{y} = 73,119 + (-0,106)x_1 + (-22,565)(x_1 - 62,60)_+^1 + 54,148(x_1 - 62,96)_+^1 + (-31,117)(x_1 - 63,15)_+^1 + (-1,057)x_2 + 7,820(x_2 - 48,75)_+^1 + (-20,110)(x_2 - 49,60)_+^1 + 14,217(x_2 - 50,02)_+^1 + (-0,142)x_3 + 17,065(x_3 - 8,90)_+^1 + (-20,997)(x_3 - 8,99)_+^1 + (-1,205)x_4 + 1,005(x_4 - 6,12)_+^1$$

Tabel 12.
Hasil Pengujian Parameter Secara Serentak

Variabel	Parameter	Estimator	<i>p-value</i>	Keputusan
X ₁	β_0	73,11855	4,70E-05	Signifikan
	β_{11}	-0,10636	0,4899	Tidak Sig*
	β_{12}	-22,565	4,42E-05	Signifikan
	β_{13}	54,14841	5,19E-05	Signifikan
	β_{14}	-31,1174	0,0001	Signifikan
X ₂	β_{21}	-1,05667	2,99E-06	Signifikan
	β_{22}	7,819532	1,54E-05	Signifikan
	β_{23}	-20,1096	2,15E-05	Signifikan
	β_{24}	14,21741	3,18E-05	Signifikan
X ₃	β_{31}	-0,14216	0,6142	Tidak Sig*
	β_{32}	17,06528	0,0447	Signifikan
	β_{33}	-20,9973	0,0212	Signifikan
X ₄	β_{41}	-1,20542	0,0081	Signifikan
	β_{42}	1,00513	0,2605	Tidak Sig*

Sig* = Signifikan

G. Pengujian Asumsi Residual

1. Asumsi Identik

Berikut merupakan hasil pengujian asumsi residual identik dengan menggunakan uji *Glejser* pada Tabel 13.

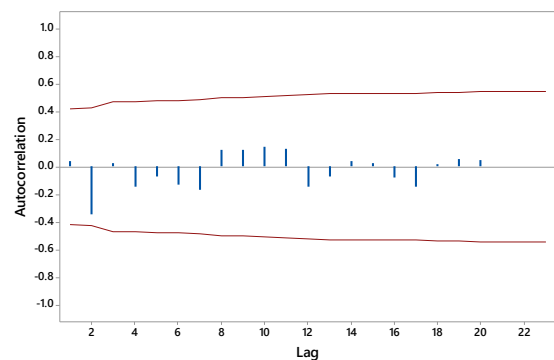
Tabel 13.
Hasil Pengujian Glejser

Sumber variasi	df	Sum of Square (SS)	Mean Square (MS)	F _{hitung}
Regresi	13	0,8728862	0,06714509	
Residual	10	0,7853663	0,07853663	0,8549525
Total	23	1,658252	-	

Berdasarkan hasil Tabel 13 diperoleh nilai statistik uji F sebesar 0,855 dengan *p-value* sebesar 0,612. Karena diperoleh *p-value* > α dengan $\alpha = 0,05$ maka dapat diambil keputusan gagal tolak H_0 yang menunjukkan bahwa tidak terdapat heteroskedastisitas. Sehingga dapat dikatakan bahwa asumsi identik pada residual terpenuhi.

2. Asumsi Independen

Berikut merupakan hasil uji independen menggunakan plot *Autocorrelation Function (ACF)*.

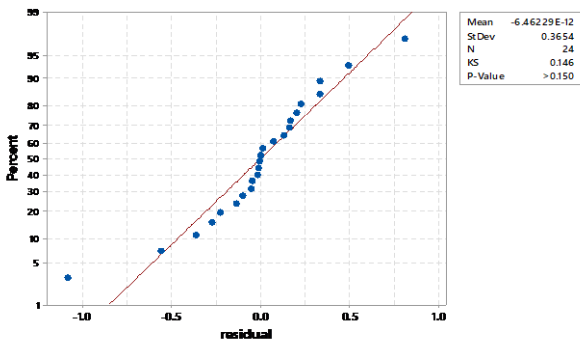


Gambar 2. Plot ACF Residual.

Berdasarkan Gambar 2 dapat dilihat bahwa tidak terdapat ρ_k yang keluar dari batas signifikansi. Hal ini menunjukkan bahwa tidak terdapat kasus autokorelasi pada residual. Sehingga dapat disimpulkan bahwa residual memenuhi asumsi independen.

3. Distribusi Normal

Uji normalitas dilakukan untuk mengetahui apakah residual mengikuti distribusi normal.



Gambar 3. Uji Kolmogorov-Smirnov.

Berdasarkan hasil uji Kolmogorov-Smirnov pada Gambar 3 diperoleh nilai *p-value* sebesar >0,150. Karena *p-value* > α dengan $\alpha = 0,05$ maka dapat diambil keputusan gagal tolak H_0 yang menunjukkan bahwa residual berdistribusi normal.

H. Nilai Koefisien Determinasi (R^2)

Model regresi nonparametrik spline linier dengan kombinasi knot menghasilkan nilai R^2 sebesar 95,23% yang artinya keempat variabel prediktor mampu menjelaskan sebesar 95,23% permasalahan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat.

I. Interpretasi Model Regresi Nonparametrik Spline

1. Hubungan antara tingkat partisipasi angkatan kerja (X_1) terhadap tingkat pengangguran terbuka (Y) dengan asumsi variabel lain dianggap konstan adalah sebagai berikut.

$$\hat{y} = (-0,106)x_1 + (-22,565)(x_1 - 62,60)_+^1 + 54,148(x_1 - 62,96)_+^1 + (-31,117)(x_1 - 63,15)_+^1$$

$$= \begin{cases} -0,106x_1 & ; & x_1 < 62,60 \\ -22,671x_1 + 1412,57 & ; & 62,60 \leq x_1 < 62,96 \\ 31,447x_1 - 1996,59 & ; & 62,96 \leq x_1 < 63,15 \\ 0,36x_1 - 31,55 & ; & x_1 \geq 63,15 \end{cases}$$

Apabila kabupaten/kota dengan nilai tingkat partisipasi angkatan kerja kurang dari 62,60 persen dan nilainya naik 0,1 persen dengan asumsi variabel prediktor lain tetap, maka tingkat pengangguran terbuka akan turun sebesar 0,106 persen. Kabupaten/kota yang termasuk pada interval ini adalah Kabupaten Cianjur, Kabupaten Kuningan, Kabupaten Cirebon, Kabupaten Sumedang, Kabupaten Bekasi, Kabupaten Bandung Barat, Kota Bogor, dan Kota Sukabumi.

2. Hubungan antara *dependency ratio* (X_2) terhadap tingkat pengangguran terbuka (Y) dengan asumsi variabel lain dianggap konstan adalah sebagai berikut.

$$\hat{y} = (-1,057)x_2 + 7,820(x_2 - 48,75)_+^1 + (-20,110)(x_2 - 49,60)_+^1 + 14,217(x_2 - 50,02)_+^1$$

$$= \begin{cases} -1,057x_2 & ; & x_2 < 48,75 \\ 6,763x_2 - 381,23 & ; & 48,75 \leq x_2 < 49,60 \\ -13,347x_2 + 616,23 & ; & 49,60 \leq x_2 < 50,02 \\ 0,97x_2 - 94,90 & ; & x_2 \geq 50,02 \end{cases}$$

Apabila kabupaten/kota dengan nilai *dependency ratio* diantara 49,60 persen hingga 50,02 persen dan apabila naik 0,1 persen, maka tingkat pengangguran terbuka yang termasuk dalam interval tersebut akan cenderung turun sebesar 13,347 persen. Kabupaten/kota yang termasuk pada segmen ini adalah Kabupaten Bandung.

3. Hubungan antara rata-rata lama sekolah (X_3) terhadap tingkat pengangguran terbuka (Y) dengan asumsi variabel lain dianggap konstan adalah sebagai berikut.

$$\hat{y} = (-0,142)x_3 + 17,065(x_3 - 8,90)_+^1 + (-20,997)(x_3 - 8,99)_+^1$$

$$= \begin{cases} -0,142x_3 & ; & x_3 < 8,90 \\ 16,923x_3 - 151,88 & ; & 8,90 \leq x_3 < 8,99 \\ -4,074x_3 + 36,88 & ; & x_3 \geq 8,99 \end{cases}$$

Apabila kabupaten/kota dengan nilai tingkat rata-rata lama sekolah lebih dari 8,99 persen dan naik 0,1 persen mengakibatkan penurunan tingkat pengangguran terbuka sebesar 4,074 persen. Kabupaten/kota yang termasuk pada segmen ini adalah Kota Bogor, Kota Sukabumi, Kota Bandung, Kota Cirebon, Kota Bekasi, Kota Depok, Kota Cimahi, dan Kota Tasikmalaya.

4. Hubungan antara laju pertumbuhan PDRB (X_4) terhadap tingkat pengangguran terbuka (Y) dengan asumsi variabel lain dianggap konstan adalah sebagai berikut.

$$\hat{y} = (-1,205)x_4 + 1,005(x_4 - 6,12)_+^1$$

$$= \begin{cases} -1,205x_4 & ; & x_4 < 6,12 \\ 0,20x_4 - 6,15 & ; & x_4 \geq 6,12 \end{cases}$$

Apabila kabupaten/kota dengan laju pertumbuhan PDRB kurang dari 6,12 persen dan naik 0,1 persen mengakibatkan kenaikan tingkat pengangguran terbuka sebesar 4,48 persen. Kabupaten/kota yang termasuk pada segmen ini adalah Kabupaten Bogor, Kabupaten Sukabumi, Kabupaten Cianjur, Kabupaten Garut, Kabupaten Tasikmalaya, Kabupaten Cirebon, Kabupaten Subang, Kabupaten Purwakarta, Kabupaten Karawang, Kabupaten Bekasi, Kabupaten Bandung Barat, Kota Sukabumi, Kota Cirebon, Kota Bekasi, Kota Cimahi, Kota Tasikmalaya, dan Kota Banjar.

V. KESIMPULAN DAN SARAN

Dari penelitian yang telah dilakukan, dapat disimpulkan bahwa tingkat pengangguran terbuka (TPT) di Provinsi Jawa Barat didapatkan kabupaten/kota yang memiliki TPT tertinggi pada tahun 2017 adalah Kabupaten Bekasi yaitu sebesar 10,97. Sedangkan kabupaten/kota yang memiliki TPT terendah di Kabupaten Pangandaran yaitu sebesar 3,340. Rata-rata TPT tiap kabupaten/kota di Provinsi Jawa Barat selama tahun 2017 adalah sebesar 8,168 dengan varians sebesar 2,801.

Pemodelan dengan regresi nonparametrik *spline truncated*, model terbaik dihasilkan menggunakan empat variabel. variabel yang berpengaruh terhadap TPT di Provinsi Jawa Barat adalah Tingkat Partisipasi Angkatan Kerja (TPAK) (x_1), *Dependency Ratio* (x_2), Rata-rata Lama Sekolah (x_3), dan Laju Pertumbuhan Ekonomi (PDRB) (x_4). Koefisien determinasi (R^2) yang dihasilkan oleh model sebesar 95,23% yang memiliki arti bahwa keempat variabel prediktor mampu menjelaskan sebesar 95,23% permasalahan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat sisanya dijelaskan oleh variabel lain yang tidak termasuk dalam model.

Saran yang dapat diberikan dalam penelitian ini adalah sebaiknya Pemerintah Provinsi Jawa Barat baiknya menekan tingkat pengangguran terbuka dengan upaya peningkatan partisipasi angkatan kerja terutama di Kota Bandung, menaikkan angka *dependency ratio*, meningkatkan pendidikan yang dilihat dari rata-rata lama sekolah, dan meningkatkan angka laju pertumbuhan PDRB karena faktor tersebut berpengaruh secara signifikan. Penelitian ini merupakan pemodelan regresi nonparametrik *spline* yang belum memperhitungkan pengaruh spasial, oleh karena itu pada penelitian selanjutnya perlu dikembangkan menjadi regresi nonparametrik spasial.

DAFTAR PUSTAKA

- [1] R. Sari and et al, "Pemodelan Pengangguran Terbuka di Jawa Timur Menggunakan Regresi Spline Multivariabel," *J. Sains dan Seni ITS*, vol. 1, 2012.
- [2] W. Astuti, "Analisis Faktor yang Berpengaruh Terhadap Tingkat Pengangguran Terbuka di Provinsi Jawa Timur Menggunakan Regresi Data Panel," *J. Sains dan Seni ITS*, vol. 6, 2017.
- [3] A. Wijaya and et al, "Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Provinsi Aceh Tahun 2015 dengan Regresi Nonparametrik Spline Truncated," Surabaya, 2018.
- [4] R. Walpole, *Pengantar Metode Statistika*, 3rd ed. Jakarta: Bumi Aksara, 1995.
- [5] R. Eubank, *Nonparametric Regression and Spline Smoothing*, 2nd ed. Texas: Department of Statistics Southern Methodist Dallas University, 1999.
- [6] Budiantara, "Metode U, GLM, CV, dan GCV dalam Regresi Non-Parametrik Spline," *Maj. Ilm. Himpun. Mat. Indones.*, vol. 6, pp. 41–45, 2000.
- [7] D. N.R and S. H., *Analisis Regresi Terapan*. Jakarta: Gramedia Pustaka Utama, 1992.
- [8] D. Gujarati, *Basic Econometrics*, 4th ed. New York: Mc Graw Hill Companies, 2004.
- [9] W. Wei, *Time Series Analysis: Univariate and Multivariate Methods*. Pearson Addison Wesley, 2006.
- [10] W. Daniel, *Statistika Nonparametrik Terapan*. Jakarta: Gramedia Pustaka Utama, 1989.
- [11] BPS Jawa Barat, "Laporan Eksekutif Keadaan Angkatan Kerja Provinsi Jawa Barat Agustus 2018," 2018.