

Analisis Regresi Nonparametrik *Spline Truncated* untuk Menganalisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Provinsi Sulawesi Selatan

Devi Carolin Wongkar, Ruliana*, Muhammad Fahmuddin S.

Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Negeri Makassar, Indonesia

Keywords: Regresi Nonparametrik *Spline Truncated*, GCV, Tingkat Pengangguran Terbuka

Abstract:

The nonparametric regression analysis is a regression model used to determine the relationship between response variable and independent variables with unknown regression curve shapes. In the nonparametric approach, one of the frequently used estimators is the spline truncated. Spline truncated model is a segmented polynomial truncation model. The advantage of this model is that it is flexible because it has knot points that can show changes in data patterns. The unemployment rate in South Sulawesi Province in 2021 reached 5.72% and became the province with the second highest unemployment rate on Sulawesi Island. Therefore, spline truncated nonparametric regression modelling will be carried out in the case of unemployment rate with each of the factors that are thought to be influential because the regression curve is found not to form a certain pattern. Based on the analysis results, the best truncated spline nonparametric regression model was obtained using three knot points and obtained the minimum GCV value of 0.38 with a coefficient of determination (R^2) value of 89%. Factors that have a significant effect on the unemployment rate in South Sulawesi are mean years of schooling (x_1) and labour force participation rate (x_2).

1. Pendahuluan

Analisis regresi adalah salah satu metode statistika yang menggunakan variabel bebas sebagai penduga (penjelas) untuk variabel respon (Tiro, 2010). Terdapat tiga pendekatan yang digunakan dalam analisis regresi yaitu regresi parametrik, nonparametrik, dan semiparametrik (Dani dkk., 2021). Apabila kurva diasumsikan tidak diketahui polanya, maka penggunaan analisis regresi nonparametrik lebih disarankan untuk digunakan.

Pendekatan dengan regresi nonparametrik memiliki beberapa estimator yang telah dikembangkan oleh para peneliti sebelumnya diantaranya yaitu Kernel (Rifai, 2019), *spline truncated* (Wahba, 1990), dan deret Fourier (Prahutama dkk., 2018). Dari beberapa estimator tersebut, estimator yang paling sering digunakan ialah *spline truncated* (Dani & Adrianingsih, 2021). *Spline truncated* merupakan model potongan polinomial yang bersifat tersegmen atau terbagi. Kelebihan dari model ini yaitu bersifat fleksibel dikarenakan *spline truncated* mempunyai titik knot yang menunjukkan terjadinya perubahan pola data (Tripena & Prabowo, 2021).

Indikator utama yang digunakan untuk mengukur angka pengangguran dalam angkatan kerja adalah Tingkat Pengangguran Terbuka (TPT). Pada Agustus 2021, TPT di Sulawesi Selatan mencapai 5,72% dan mengalami penurunan sebesar 0,59% poin dibandingkan dengan TPT Agustus 2020 yang mencapai 6,31%. Meskipun mengalami penurunan, Provinsi Sulawesi Selatan masih menjadi provinsi dengan angka TPT tertinggi kedua di Pulau Sulawesi

* Corresponding author.

E-mail address: ruliana.t@unm.ac.id



(Badan Pusat Statistik, 2021).

Tingginya pengangguran dapat menyebabkan penurunan kesejahteraan masyarakat karena pendapatan yang diperoleh semakin menurun sehingga dapat menyebabkan terjadinya kemiskinan. Dalam perhitungan Indeks Pembangunan Manusia (IPM), salah satu indikator yang digunakan untuk mengukur kualitas pendidikan yaitu Rata-Rata Lama Sekolah (RLS) (Siskawati dkk., 2021). Selain itu, Umur Harapan Hidup (UHH) juga termasuk salah satu faktor yang dapat mempengaruhi pengangguran karena secara teori, permasalahan yang dialami pengangguran menunjukkan kegagalan dalam pembangunan sehingga hal tersebut berdampak pada kesehatan mental yang dialami oleh pekerja tersebut dan secara ekonomi tidak mendapatkan pemasukan yang akan berpengaruh terhadap masalah sosial (Chalifari dkk., 2020). Faktor lainnya yang mempengaruhi meningkatnya pengangguran yaitu Tingkat Partisipasi Angkatan Kerja (TPAK) sebab adanya ketidakseimbangan antara jumlah angkatan kerja dan lapangan kerja yang tersedia dapat menimbulkan masalah pengangguran (Salsabila dkk., 2022).

Beberapa penelitian sebelumnya telah membahas terkait regresi nonparametrik. Penelitian yang dilakukan oleh Dani dkk. (2021) meneliti tentang 'Pemodelan Regresi Nonparametrik *Spline Truncated* pada Data Longitudinal'. Selain itu, penelitian yang dilakukan oleh Kurniawati & Budiantara (2019) meneliti tentang 'Pemodelan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat Menggunakan Regresi Nonparametrik *Spline Truncated*'. Penelitian yang dilakukan oleh Adetya Arjun dkk. (2019) meneliti tentang 'Analisis Faktor-Faktor Yang Mempengaruhi Tingkat Pengangguran Terbuka di Kalimantan Menggunakan Regresi Nonparametrik *Spline Truncated*'.

2. Tinjauan Pustaka

2.1 Analisis Regresi

Analisis regresi adalah metode statistika yang memperhatikan hubungan antara satu variabel respon (*response variable*) dengan satu atau lebih variabel penjelas (*explanatory variable*) (Tiro, 2008). Secara umum bentuk model regresi linier dapat dilihat pada Persamaan (2.1) (Draper & Smith, 2014).

$$Y_i = \beta_0 + \sum_{k=1}^p \beta_k X_{ki} + \varepsilon_i \quad (2.1)$$

2.2 Analisis Regresi Nonparametrik *Spline Truncated*

Regresi nonparametrik *spline truncated* adalah metode regresi nonparametrik yang paling banyak digunakan (Izzah & Budiantara, 2020). Fungsi *spline truncated* didapatkan 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 akan membentuk kurva regresi yaitu $f(x_i)$, lebih rinci dapat dituliskan dalam persamaan berikut ini (Kurniawati & Budiantara, 2019).

$$y_i = \sum_{j=0}^p \beta_j x_i^j + \sum_{k=1}^r \beta_{p+r} (x_i - K_r)_+^p + \varepsilon_i, i = 1, 2, \dots, n \quad (2.2)$$

Fungsi *truncated* $(x_i - K_r)_+^p$ menghasilkan persamaan berikut (Kurniawati & Budiantara, 2019).

$$(x_i - K_r)_+^p = \begin{cases} (x_i - K_r)^p, & x_i \geq K_r \\ 0, & x_i < K_r \end{cases} \quad (2.3)$$

2.3 Estimasi Parameter

Estimasi parameter untuk model regresi nonparametrik *spline* dapat menggunakan metode *Ordinary Least Square* (OLS). Estimasi parameter ini bertujuan untuk meminimumkan jumlah kuadrat *error*. Dalam notasi matriks penulisan regresi nonparametrik *spline* dapat dinyatakan pada Persamaan (2.6) (Kurniawati & Budiantara, 2019).

$$y = \mathbf{X}\beta + \varepsilon \quad (2.4)$$

Sehingga, penaksir parameter regresi nonparametrik *spline* dengan metode OLS adalah sebagai berikut (Kurniawati & Budiantara, 2019).

$$\hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{y} \quad (2.5)$$

2.4 Pemilihan Titik Knot Optimal

Titik knot merupakan titik perpaduan dimana terdapat pada perubahan perilaku pola kurva dengan interval yang berlainan. Salah satu metode yang digunakan untuk memilih titik knot optimal adalah metode *Generalized Cross Validation* (GCV) (Kurniawati & Budiantara, 2019). Titik-titik knot optimal dapat diperoleh dari nilai GCV yang paling minimum. Metode GCV dapat dihitung dalam persamaan berikut (Pratiwi, 2020).

$$GCV(K) = \frac{MSE(K)}{[n^{-1}trace(\mathbf{I} - \mathbf{A})]^2} \quad (2.6)$$

2.5 Pengujian Parameter Model

Pengujian secara serentak merupakan pengujian yang dilakukan untuk mengetahui signifikansi parameter model secara bersama-sama. Hipotesis yang digunakan adalah:

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

$$H_1 : \text{Minimal ada satu } \beta_m \neq 0, m = 1, 2, \dots, p + r$$

Statistik uji yang digunakan adalah sebagai berikut (Syam dkk., 2019).

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2 / (p + r)}{\sum_{i=1}^n (y_i - \hat{y}_i)^2 / (n - (p + r) - 1)} \quad (2.7)$$

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

Kemudian, pengujian secara parsial berfungsi untuk mengidentifikasi apakah parameter secara individual memiliki pengaruh yang signifikan terhadap variabel respon. Hipotesis yang digunakan yaitu sebagai berikut.

$$H_0 : \beta_m = 0$$

$$H_1 : \beta_m \neq 0, m = 1, 2, \dots, p + r$$

Statistik uji yang digunakan dalam uji ini sebagai berikut (Kurniawati & Budiantara, 2019).

$$t_{hitung} = \frac{\hat{\beta}_j}{\sqrt{SE(\hat{\beta}_j)}} \quad (2.8)$$

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

2.6 Koefisien Determinasi

Koefisien determinasi merupakan kuantitas yang mampu menjelaskan sumbangan variabel prediktor terhadap variabel respon (Izzah & Budiantara, 2020). Berikut ini merupakan persamaan untuk menghitung nilai R^2 (Draper & Smith, 1992).

$$R^2 = \frac{SS_{regresi}}{SS_{Total}} = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.9)$$

2.7 Pengujian Asumsi Residual

Uji asumsi identik digunakan untuk mengetahui terjadinya homogenitas pada variansi residual. Uji ini menggunakan uji *Glejser* (Izzah & Budiantara, 2020). Hipotesis yang digunakan sebagai berikut.

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

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

Dengan statistik uji yang digunakan sebagai berikut (Izzah & Budiantara, 2020).

$$F_{hitung} = \frac{MS_{regresi}}{MS_{residual}} = \frac{\sum_{i=1}^n (|\hat{\varepsilon}_i| - |\bar{\varepsilon}|)^2 / (v-1)}{\sum_{i=1}^n (|\varepsilon_i| - |\hat{\varepsilon}_i|)^2 / (n-v)} \quad (2.10)$$

Tolak H_0 jika $F_{hitung} > F_{\alpha; (v-1, n-v)}$ atau $p\text{-value} < \alpha$.

Kemudian, uji asumsi independen digunakan untuk mengetahui ada atau tidak adanya korelasi antar residual dan uji ini dapat menggunakan beberapa cara salah satunya melakukan uji *Durbin-Watson* dengan hipotesis sebagai berikut.

$$H_0 : \rho = 0 \text{ (residual independen)}$$

$$H_1 : \rho \neq 0 \text{ (residual dependen)}$$

Statistik uji yang digunakan dalam uji ini sebagai berikut (Izzah & Budiantara, 2020).

$$d_{hitung} = \frac{\sum_{t=2}^n (\varepsilon_t - \varepsilon_{t-1})^2}{\sum_{t=1}^n \varepsilon_t^2} \quad (2.11)$$

Adapun kaidah keputusan dalam uji ini sebagai berikut, apabila (Izzah & Budiantara, 2020):

1. $0 < d_{hitung} < d_L$ atau $(4 - d_U) < d_{hitung} < 4$, maka tolak H_0 .
2. $d_U < d_{hitung} < (4 - d_U)$, maka gagal tolak H_0 .
3. $d_L \leq d_{hitung} \leq d_U$ atau $(4 - d_U) \leq d < (4 - d_L)$, maka tidak ada keputusan tolak H_0 ataupun gagal tolak H_0 .

Selanjutnya, pengujian asumsi distribusi normal merupakan uji asumsi yang digunakan untuk mendeteksi apakah residual data telah mengikuti distribusi normal. Menurut Daniel (1989) dalam Kurniawati & Budiantara (2019), uji ini dapat menggunakan uji *Kolmogorov-Smirnov*. Hipotesis yang digunakan sebagai berikut.

$$H_0 : F_0(x) = F(x) \text{ (residual mengikuti distribusi normal)}$$

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

Statistik uji yang digunakan sebagai berikut (Kurniawati & Budiantara, 2019).

$$D = \sup_{\varepsilon} |F_n(\varepsilon) - F_0(\varepsilon)| \quad (2.12)$$

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

2.8 Tingkat Pengangguran Terbuka

Tingkat Pengangguran Terbuka (TPT) yaitu indikator yang digunakan untuk mengukur tenaga kerja yang diserap oleh lapangan usaha di pasar kerja dan menggambarkan kurang termanfaatkannya pasokan tenaga kerja. Tingkat pengangguran terbuka dapat dihitung dalam persamaan berikut (Gatiningsih & Sutrisno, 2017).

$$\text{Tingkat Pengangguran} = \frac{\text{Jumlah Pengangguran}}{\text{Angkatan Kerja}} \times 100\% \quad (2.13)$$

3. Metode Penelitian

3.1 Sumber Data

Data yang akan digunakan dalam penelitian ini adalah data sekunder yang dipublikasikan oleh Badan Pusat Statistik Provinsi Sulawesi Selatan tahun 2021 (<https://sulsel.bps.go.id/>). Data yang digunakan adalah data Tingkat

Pengangguran Terbuka (TPT) dan faktor-faktor yang mempengaruhinya dimana pada penelitian ini terdapat unit observasi yaitu 21 kabupaten dan 3 kota.

3.2 Teknik Analisis

Langkah-langkah teknik analisis data dalam penelitian ini adalah:

1. Melakukan analisis statistika deskriptif.
2. Membuat *scatterplot* untuk mengetahui bentuk pola data.
3. Memodelkan data menggunakan model regresi nonparametrik *spline* dengan satu, dua, dan tiga knot.
4. Menentukan titik knot paling optimal berdasarkan GCV minimum.
5. Mendapatkan model regresi nonparametrik *spline* dengan titik knot paling optimal.
6. Menguji signifikansi parameter regresi nonparametrik *spline* secara serentak dan parsial.
7. Melakukan uji asumsi residual IIDN (Identik, Independen, dan Distribusi Normal). Jika pengujian asumsi residual tidak terpenuhi maka dilakukan transformasi, setelah itu lakukan kembali langkah 2 sampai langkah 6.
8. Menghitung nilai koefisien determinasi (R^2).
9. Interpretasi dan menarik kesimpulan.

4. Hasil dan Pembahasan

4.1 Analisis Deskriptif

Berikut ini merupakan statistika deskriptif data TPT dan faktor-faktor yang diduga mempengaruhinya yang ditampilkan pada Tabel 4.1.

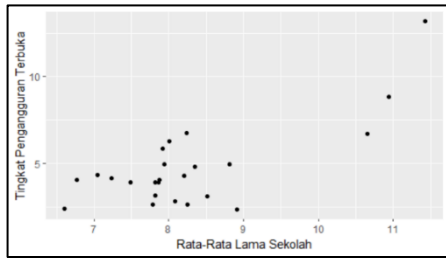
Tabel 1 Hasil Analisis Deskriptif

Variabel	Mean	Varians	Minimum	Maksimum
y	4,75	5,77	2,34	13,18
x_1	8,27	1,44	6,60	11,43
x_2	65,83	33,05	57,78	77,99
x_3	69,58	3,85	66,49	73,41
x_4	9,47	8,50	4,82	14,28

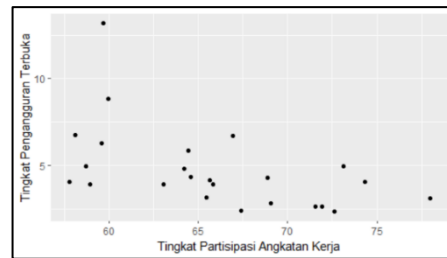
Berdasarkan Tabel 1, variabel y merupakan variabel TPT di Sulawesi Selatan tahun 2021 dengan rata-rata sebesar 4,75 persen dan nilai varians sebesar 5,77. Angka TPT terendah terjadi di Kabupaten Enrekang yaitu sebesar 2,34 persen, sedangkan angka TPT tertinggi terjadi di Kota Makassar yaitu sebesar 13,18 persen. Masing-masing setiap variabel independennya juga menggunakan nilai *mean*, varians, nilai minimum, dan nilai maksimum sebagai deskriptif data dengan penjelasan yang sama dengan variabel y .

4.2 Pola Hubungan Antara TPT dengan Variabel yang Diduga Berpengaruh

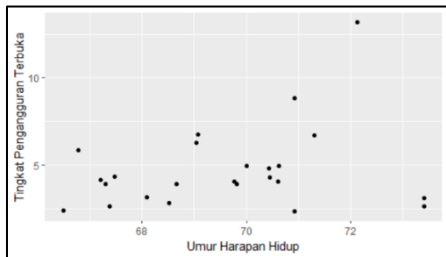
Berikut ini merupakan masing-masing *scatterplot* antara variabel respon dan masing-masing variabel independen.



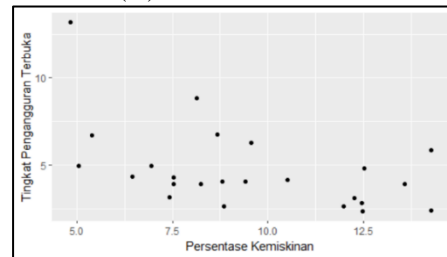
Gambar 1 Scatterplot antara TPT (y) dengan RLS (x₁)



Gambar 2 Scatterplot antara TPT (y) dengan TPAK (x₂)



Gambar 3 Scatterplot antara TPT (y) dengan UHH (x₃)



Gambar 4 Scatterplot antara TPT (y) dengan persentase kemiskinan (x₄)

Berdasarkan Gambar 1 sampai Gambar 4 menunjukkan bahwa *scatterplot* antara TPT dengan masing-masing variabel yang diduga mempengaruhinya menyebar tidak mengikuti suatu pola tertentu. Penggunaan metode regresi nonparametrik *spline truncated* teridentifikasi digunakan pada pola yang terjadi antara variabel respon dan variabel prediktornya.

4.3 Pemodelan Regresi Nonparametrik Spline Truncated

4.3.1 Pemodelan dengan Satu Titik Knot

Penggunaan titik-titik knot dengan nilai GCV yang paling optimum untuk model regresi nonparametrik *spline truncated* menggunakan satu titik knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 (x_{i1} - 10,54)_+ + \hat{\beta}_3 x_{i2} + \hat{\beta}_4 (x_{i2} - 74,28)_+ + \hat{\beta}_5 x_{i3} + \hat{\beta}_6 (x_{i3} - 72,14)_+ + \hat{\beta}_7 x_{i4} + \hat{\beta}_8 (x_{i4} - 12,54)_+$$

4.3.2 Pemodelan dengan Dua Titik Knot

Penggunaan titik-titik knot dengan nilai GCV yang paling optimum untuk model regresi nonparametrik *spline truncated* menggunakan dua titik knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 (x_{i1} - 6,70)_+ + \hat{\beta}_3 (x_{i1} - 7,29)_+ + \hat{\beta}_4 x_{i2} + \hat{\beta}_5 (x_{i2} - 58,19)_+ + \hat{\beta}_6 (x_{i2} - 60,67)_+ + \hat{\beta}_7 x_{i3} + \hat{\beta}_8 (x_{i3} - 66,63)_+ + \hat{\beta}_9 (x_{i3} - 67,48)_+ + \hat{\beta}_{10} x_{i4} + \hat{\beta}_{11} (x_{i4} - 5,01)_+ + \hat{\beta}_{12} (x_{i4} - 6,17)_+$$

4.3.2 Pemodelan dengan Tiga Titik Knot

Penggunaan titik-titik knot dengan nilai GCV yang paling optimum untuk model regresi nonparametrik *spline truncated* menggunakan tiga titik knot adalah sebagai berikut.

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \hat{\beta}_2 (x_{i1} - 8,77)_+ + \hat{\beta}_3 (x_{i1} - 9,16)_+ + \hat{\beta}_4 (x_{i1} - 9,85)_+ + \hat{\beta}_5 x_{i2} + \hat{\beta}_6 (x_{i2} - 66,85)_+ + \hat{\beta}_7 (x_{i2} - 68,50)_+ + \hat{\beta}_8 (x_{i2} - 71,39)_+ + \hat{\beta}_9 x_{i3} + \hat{\beta}_{10} (x_{i3} - 69,60)_+ + \hat{\beta}_{11} (x_{i3} - 70,16)_+ + \hat{\beta}_{12} (x_{i3} - 71,15)_+ + \hat{\beta}_{13} x_{i4} + \hat{\beta}_{14} (x_{i4} - 9,07)_+ + \hat{\beta}_{15} (x_{i4} - 9,84)_+ + \hat{\beta}_{16} (x_{i4} - 11,19)_+$$

4.4 Pemilihan Titik Knot Optimal

Berikut ini nilai GCV dari setiap hasil titik knot yang ditunjukkan pada tabel berikut.

Tabel 2 Nilai GCV Terkecil dari Tiap Pemodelan

Jumlah Knot	GCV
Satu	1,90
Dua	1,48
Tiga	0,38

Tabel 2 menunjukkan nilai GCV terkecil diperoleh model regresi nonparametrik *spline* dengan tiga titik knot. Jadi, dapat disimpulkan bahwa pemodelan regresi nonparametrik *spline truncated* dengan tiga titik knot merupakan model dengan titik knot yang paling optimal.

4.5 Pengujian Parameter Model

4.5.1 Pengujian Secara Serentak

Tabel 3 Hasil Uji Serentak

Sumber Variasi	Df	R-Square	F-hitung	p-value
Regresi	16	0,89	3,65	0,045

Berdasarkan hasil uji serentak pada Tabel 3 didapatkan nilai $p\text{-value} < \alpha$ artinya minimal terdapat satu variabel prediktor yang memberikan pengaruh secara signifikan terhadap model.

4.5.2 Pengujian Secara Parsial

Tabel 4 Hasil Uji Parsial

Variabel	Parameter	Estimasi	p-value	Keputusan
x_1	β_0	4,93	0,01	Signifikan
	β_1	6,51	0,31	Tidak Signifikan
	β_2	-13,78	0,02	Signifikan
	β_3	28,74	< 0,001	Signifikan
	β_4	-7,71	0,02	Signifikan
x_2	β_5	8,64	0,03	Signifikan
	β_6	-9,09	0,12	Tidak Signifikan
	β_7	-3,00	0,44	Tidak Signifikan
	β_8	4,32	0,20	Tidak Signifikan
x_3	β_9	-3,44	0,20	Tidak Signifikan
	β_{10}	1,10	0,66	Tidak Signifikan
	β_{11}	-8,77	0,13	Tidak Signifikan
	β_{12}	3,53	0,41	Tidak Signifikan
x_4	β_{13}	-3,50	0,42	Tidak Signifikan
	β_{14}	1,57	0,74	Tidak Signifikan
	β_{15}	-10,38	0,26	Tidak Signifikan
	β_{16}	3,28	0,53	Tidak Signifikan

Tabel 4 menunjukkan hasil estimasi dari parameter yang diperoleh menggunakan metode OLS dengan model regresi nonparametrik *spline* dengan tiga titik knot dan berdasarkan hasil uji parsial didapatkan 5 parameter yang memberi pengaruh secara signifikan diantaranya $\beta_0, \beta_2, \beta_3, \beta_4,$ dan β_5 karena dari masing-masing parameter tersebut diperoleh nilai $p\text{-value} < \alpha$.

4.6 Pengujian Asumsi Residual

4.6.1 Asumsi Identik

Tabel 5 Hasil Uji Glejser

Sumber Variasi	df	F _{hitung}	p-value
Regresi	16	0,82	0,65

Tabel 5 menunjukkan hasil yang diperoleh yaitu nilai $p\text{-value} > \alpha$ artinya residual telah memenuhi asumsi identik.

4.6.2 Asumsi Independen

Tabel 6 Hasil Uji *Durbin-Watson*

d_{hitung}	$d_{L;0,05}$	$d_{U;0,05}$	$4-d_{L;0,05}$	$4-d_{U;0,05}$
2,02	1,03	1,78	2,99	2,23

Berdasarkan hasil yang diperoleh pada Tabel 6, nilai d_{hitung} lebih besar dari nilai $d_{U;0,05}$ dan nilai d_{hitung} lebih kecil dari nilai $4-d_{U;0,05}$ artinya residual telah memenuhi asumsi independen.

4.6.3 Asumsi Distribusi Normal

Tabel 7 Hasil Uji *Kolmogorov-Smirnov*

$p-value$	α
0,44	0,05

Berdasarkan Tabel 7 diperoleh bahwa nilai $p-value > \alpha$ artinya residual telah memenuhi asumsi distribusi normal.

4.7 Nilai Koefisien Determinasi

Berdasarkan Tabel 3, nilai koefisien determinasi yang didapatkan dari model regresi nonparametrik *spline truncated* yaitu sebesar 0,89 atau 89 persen. Nilai tersebut menjelaskan bahwa variabel prediktor secara signifikan memberikan pengaruh sebesar 89 persen terhadap tingkat pengangguran terbuka di Sulawesi Selatan tahun 2021, sedangkan sisanya yaitu 11 persen merupakan pengaruh dari variabel lain yang tidak terdapat dalam model.

4.8 Interpretasi Model Regresi Nonparametrik *Spline Truncated*

Model regresi nonparametrik *spline truncated* yang terbentuk menggunakan titik knot optimal yaitu menggunakan tiga titik knot ditunjukkan pada persamaan di bawah berikut.

$$\hat{y} = 4,93 + 6,51x_{i1} - 13,78(x_{i1} - 8,77)_+ + 28,74(x_{i1} - 9,16)_+ - 7,71(x_{i1} - 9,85)_+ + 8,64x_{i2} - 9,09(x_{i2} - 66,85)_+ - 3,00(x_{i2} - 68,50)_+ + 4,32(x_{i2} - 71,39)_+ - 3,44x_{i3} + 1,10(x_{i3} - 69,60)_+ - 8,77(x_{i3} - 70,16)_+ + 3,53(x_{i3} - 71,15)_+ - 3,50x_{i4} + 1,57(x_{i4} - 9,07)_+ - 10,38(x_{i4} - 9,84)_+ + 3,28(x_{i4} - 11,19)_+$$

Berikut ini adalah interpretasi dari setiap variabel yang berpengaruh secara signifikan terhadap TPT.

1. Dengan mengasumsikan variabel prediktor lain konstan selain x_1 , maka pengaruh dari variabel Rata-Rata Lama Sekolah (RLS) dan variabel Tingkat Pengangguran Terbuka (TPT) termasuk ke dalam komponen nonparametrik dapat dituliskan sebagai berikut.

$$\hat{y} = 8,64x_{i2} - 9,09(x_{i2} - 66,85)_+ - 3,00(x_{i2} - 68,50)_+ + 4,32(x_{i2} - 71,39)_+$$

$$= \begin{cases} 8,64x_{i2} & ; x_{i2} < 66,85 \\ -0,45 x_{i2} + 607,67 & ; 66,85 \leq x_{i2} < 68,50 \\ -3,45x_{i2} + 813,17 & ; 68,50 \leq x_{i2} < 71,39 \\ 0,87x_{i2} + 504,76 & ; x_{i2} \geq 71,39 \end{cases}$$

2. Dengan mengasumsikan variabel prediktor lain konstan selain x_2 , maka pengaruh dari variabel Tingkat Partisipasi Angkatan Kerja (TPAK) dan variabel Tingkat Pengangguran Terbuka (TPT) termasuk ke dalam komponen nonparametrik dapat dituliskan sebagai berikut.

$$\hat{y} = 8,64x_{i2} - 9,09(x_{i2} - 66,85)_+ - 3,00(x_{i2} - 68,50)_+ + 4,32(x_{i2} - 71,39)_+$$

$$= \begin{cases} 8,64x_{i2} & ; x_{i2} < 66,85 \\ -0,45 x_{i2} + 607,67 & ; 66,85 \leq x_{i2} < 68,50 \\ -3,45x_{i2} + 813,17 & ; 68,50 \leq x_{i2} < 71,39 \\ 0,87x_{i2} + 504,76 & ; x_{i2} \geq 71,39 \end{cases}$$

5. Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang telah dilakukan pada bab sebelumnya dapat ditarik kesimpulan sebagai berikut.

1. Model regresi nonparametrik *spline truncated* terbaik yang diperoleh yaitu dengan menggunakan tiga titik knot. Berikut ini model regresi yang dihasilkan.

$$\hat{y} = 4,93 + 6,51x_{i1} - 13,78(x_{i1} - 8,77)_+ + 28,74(x_{i1} - 9,16)_+ - 7,71(x_{i1} - 9,85)_+ + 8,64x_{i2} - 9,09(x_{i2} - 66,85)_+ - 3,00(x_{i2} - 68,50)_+ + 4,32(x_{i2} - 71,39)_+ - 3,44x_{i3} + 1,10(x_{i3} - 69,60)_+ - 8,77(x_{i3} - 70,16)_+ + 3,53(x_{i3} - 71,15)_+ - 3,50x_{i4} + 1,57(x_{i4} - 9,07)_+ - 10,38(x_{i4} - 9,84)_+ + 3,28(x_{i4} - 11,19)_+.$$

2. Variabel prediktor yang berpengaruh secara signifikan terhadap TPT di Sulawesi Selatan adalah variabel rata-rata lama sekolah (x_1) dan tingkat partisipasi angkatan kerja (x_2).

References

- Adetya Arjun, D., Sifriani, & Syaripuddin. (2019). Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka di Kalimantan Menggunakan Regresi Nonparametrik Spline Truncated. *Prosiding Seminar Nasional Matematika, Statistika, Dan Aplikasinya 2019*, 1, 115–121.
- Badan Pusat Statistik. (2021). *Keadaan Ketenagakerjaan Provinsi Sulawesi Selatan Agustus 2021*.
- Chalifari, Anwar, K., & Muhammad Abdy Yusuf, dan. (2020). Pengaruh Angka Harapan Hidup (AHH) dan Konsumsi Per Kapita Terhadap Pengangguran. *Jurnal Ekonomi Dan Pembangunan*, 11.
- Dani, A. T. R., & Adrianingsih, N. Y. (2021). Pemodelan Regresi Nonparametrik dengan Estimator Spline Truncated vs Deret Fourier. *Jambura Journal of Mathematics*, 3(1), 26–36.
- Dani, A. T. R., Ni'matuzzahroh, L., Ratnasari, V., & Budiantara, I. N. (2021). Pemodelan Regresi Nonparametrik Spline Truncated pada Data Longitudinal. *Inferensi*, 4(1), 47.
- Daniel, W. W. (1989). *Statistika Nonparametrik Terapan*. Gramedia.
- Draper, N. R., & Smith, H. (1992). *Analisis Regresi Terapan (Second Edition)*. PT. Gramedia Pustaka Utama.
- Draper, N. R., & Smith, H. (2014). *Applied Regression Analysis (Third Edition)*. John Wiley & Sons.
- Gatiningsih, & Sutrisno, E. (2017). *Kependudukan dan Ketenagakerjaan (Vol. 1)*. Fakultas Manajemen Pemerintahan IPDN.
- Izzah, N., & Budiantara, I. N. (2020). Pemodelan Faktor-faktor yang Mempengaruhi Tingkat Partisipasi Angkatan Kerja Perempuan di Jawa Barat Menggunakan Regresi Nonparametrik Spline Truncated. *Inferensi*, 3, 21–27.
- Kurniawati, N. A., & Budiantara, I. N. (2019). Pemodelan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat Menggunakan Regresi Nonparametrik Spline Truncated. *Jurnal Sains Dan Seni ITS*, 8, 334–340.
- Prahatama, A., Suparti, & Utami, T. W. (2018). Modelling Fourier Regression for Time Series Data- A Case Study: Modelling Inflation in Foods Sector in Indonesia. *Journal of Physics: Conference Series*, 1–9.
- Pratiwi, L. P. S. (2020). Pemilihan Titik Knot Optimal Model Spline Truncated dalam Regresi Nonparametrik Multivariabel dengan GCV. *Jurnal Matematika*, 10, 78–90.
- Rifai, N. A. K. (2019). Pendekatan Regresi Nonparametrik dengan Fungsi Kernel untuk Indeks Harga Saham Gabungan. *Jurnal Statistika*, 1, 23–61.
- Salsabila, N. A., Andriani, S., Mirisda, & Nohe, D. A. (2022). Analisis Pengaruh Tingkat Partisipasi Angkatan Kerja dan Indeks Pembangunan Manusia Terhadap Tingkat Pengangguran Terbuka Menggunakan Regresi Probit dan Logit. *Prosiding Seminar Nasional Matematika, Statistika, Dan Aplikasinya*, 2, 344–353.
- Siskawati, N., Surya, R. Z., & Sudeska, E. (2021). Pengaruh Harapan Lama Sekolah, Rata-Rata Lama Sekolah terhadap Tingkat Pengangguran Terbuka di Kabupaten/Kota Provinsi Riau. *Jurnal Selodang Mayang*, 7, 173–177.
- Syam, R., Sanusi, W., & Adawiyah, D. R. (2019). Model Regresi Nonparametrik dengan Pendekatan Spline (Studi Kasus: Berat Badan Lahir Rendah di Rumah Sakit Ibu dan Anak Siti Fatimah Makassar). *Journal of Mathematics, Computations and Statistics*, 2, 70–81.
- Tiro, M. A. (2008). *Analisis Regresi dengan Data Kategori (Edisi Ketiga)*. Andira Publisher.

Tiro, M. A. (2010). *Analisis Korelasi dan Regresi (Edisi Ketiga)*. Andira Publisher.

Tripena, A., & Prabowo, A. (2021). Estimated Spline in Nonparametric Regression with a Generalized Cross Validation and Unbiased Risk Approach. *Proceedings of the 11th Annual International Conference on Industrial Engineering and Operations Management Singapore*, 3788–3798.

Wahba, G. (1990). *Spline Models for Observational Data*. SIAM.