

## **Pemodelan Faktor-Faktor Yang Mempengaruhi Tingkat Pengangguran Terbuka (Tpt) Di Provinsi Jawa Tengah Menggunakan Regresi *Spline Truncated* Multivariabel**

Zenitha Amalia Azhar<sup>1</sup>, Sri Sulistijowati Handajani<sup>2</sup>, Isnandar Slamet<sup>3</sup>

<sup>1,2</sup>Program Studi Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Sebelas Maret Kentingan Jl. Ir. Sutami No.36, Jebres, Kec. Jebres, Kota Surakarta, Jawa Tengah 57126, Indonesia

e-mail: zenithaazhar@student.uns.ac.id<sup>1</sup>, rr\_ssh@staff.uns.ac.id<sup>2</sup>, isnandarlamet@staff.uns.ac.id<sup>3</sup>

Received : December, 2023

Accepted : June, 2024

Published : June, 2024

### **Abstract**

Human life depends on work as it brings self-actualization to families, societies, and nations. Increasing the Open Unemployment Rate (OPR) is an employment problem. Statistically speaking, regression analysis is a tool for discovering how one or more variables (the predictors) affect another (the response variables). For this TPT case study in Central Java, researchers looked into the nonparametric regression model of spline reduced using the UBR and GCV approaches for knot selection. The results demonstrated that the GCV model produced MSE values of  $1.381e-01$  and  $R^2$  of 95.69%, while the UBR model generated MSE value of  $1.380e-01$ , and  $R^2$ .

**Keywords:** *Unemployment Index, Nonparametrik Spline Truncated Regression, Unbiased Risk (UBR), Generalized Cross Validation (GCV).*

### **Abstrak**

Sebagai sarana pemenuhan kebutuhan individu, komunal, dan nasional, tenaga kerja sangat penting bagi eksistensi manusia. Salah satu masalah yang berkaitan dengan tenaga kerja adalah meningkatnya Tingkat Pengangguran Terbuka (TPT). Menentukan pola hubungan dan pengaruh variabel prediktor terhadap variabel respon adalah tujuan dari analisis regresi, sebuah pendekatan statistik. Strategi optimal pemilihan titik knot pada studi kasus TPT di Jawa Tengah diselidiki dalam penelitian ini dengan menggunakan model regresi nonparametrik spline truncated dengan metode UBR dan GCV. Hasilnya menunjukkan bahwa model GCV menghasilkan nilai MSE sebesar  $1,381e-01$  dan  $R^2$  sebesar 95,69%, sedangkan model UBR menghasilkan nilai MSE sebesar  $1,381e-01$  dan  $R^2$ .

**Kata Kunci:** *TPT, Regresi Nonparametrik Spline Truncated, Unbiased Risk (UBR), Generalized Cross Validation (GCV)*

## 1. PENDAHULUAN

Kehidupan manusia bergantung pada pekerjaan. Untuk aktualisasi diri dan berkontribusi kepada keluarga, masyarakat, dan bangsanya, setiap orang membutuhkan pekerjaan. Sebenarnya, tidak semua warga negara memiliki hak tersebut. Peningkatan tingkat pengangguran merupakan kendala utama yang menghambat kemajuan dalam upaya untuk meningkatkan ketenagakerjaan. Kesenjangan yang semakin besar antara permintaan dan penawaran pekerjaan karena jumlah pekerjaan yang terbatas tidak dapat menampung seluruh angkatan kerja yang ada.

Salah satu cara untuk mengatasi masalah pengangguran adalah dengan mengidentifikasi variabel-variabel yang mempengaruhi tingkat pengangguran terbuka. Ada beberapa pendekatan yang dapat digunakan untuk tujuan ini. Menggunakan analisis regresi adalah salah satunya. Menemukan pola korelasi adalah tujuan dari analisis regresi, sebuah pendekatan statistik.

Model regresi nonparametrik sering kali menggunakan teknik-teknik seperti Fourier, Wavelet, Spline, Multivariate Adaptive Regression Splines (MARS), dan Kernel. Metode spline adalah yang paling terkenal di kalangan peneliti karena beberapa keunggulannya, termasuk fleksibilitas yang luar biasa, kemampuan untuk menunjukkan perubahan perilaku yang terjadi pada subdata. Dengan menggunakan estimator spline terpotong dan deret fourier, Dani (2021) menggunakan regresi nonparametrik untuk mensimulasikan Angka Kematian Ibu di Indonesia.

*Spline* adalah segmen dengan kualitas tersegmentasi dan kontinu dalam polinomial. Kemampuan spline untuk beradaptasi dengan pola data yang berubah dan beradaptasi dengan ramalan data adalah salah satu dari banyak keunggulannya. Ini disebabkan oleh fakta bahwa spline menggunakan titik-titik knot sebagai acuan. Pola data berubah di titik knot ini. Dengan titik knot ini, spline lebih fleksibel daripada polinomial, yang memungkinkannya

mengikuti karakteristik lokal data atau fungsi dengan lebih baik. Dalam proses estimasi spline, masalah utama adalah menemukan parameter penghalus terbaik. Tantangan utama dalam memperoleh estimasi spline melalui optimasi Penalised Least Square (PLS) adalah menentukan titik-titik knot yang optimal. Semakin baik titik knot yang dipilih saat membuat model regresi, semakin baik model yang dihasilkan.

Saat menulis persamaan, angka Arab yang diapit oleh tanda kurung digunakan untuk menunjukkan nomor persamaan secara berurutan. Untuk membuat nomor persamaan menjadi rata kanan, nomor tersebut diberi tab, seperti pada contoh di bawah ini: Metode GCV dan UBR untuk memilih lokasi simpul terbaik akan dibandingkan dalam penelitian ini. Ada dua teknik untuk memilih titik-titik knot terbaik: UBR dan GCV. Masing-masing memiliki manfaat yang berbeda-beda. Pendekatan GCV memiliki beberapa keuntungan, seperti optimasi asimtotik, invarian terhadap transformasi data, kesederhanaan dan efisiensi dalam komputasi, dan tidak memerlukan informasi estimasi varians. Sedangkan pendekatan UBR lebih cocok untuk data dengan distribusi non-Gaussian atau anomali, metode pemilihan titik knot GCV lebih cocok untuk data dengan distribusi Gaussian atau normal (Wahba, 1990; Wang, 1998). Utami (2018) melakukan penelitian mengenai tingkat pengangguran terbuka dengan meneliti faktor-faktor yang dapat mempengaruhinya. Tujuan dari penelitian ini adalah untuk mengetahui model regresi spline truncated multivariabel yang mana, apakah model regresi spline truncated multivariabel dengan metode GCV atau metode UBR yang digunakan. Penelitian yang berjudul "Analisis regresi spasial dengan pembobot Queen Contiguity pada tingkat pengangguran terbuka di Jawa Tengah" (Kurnianto et al., 2021) dan "analisis faktor-faktor yang mempengaruhi tingkat pengangguran terbuka di Provinsi Jawa Timur dengan menggunakan regresi data panel" (Astuti et al., 2017).

Nasuha (2016) meneliti topik penerapan pendekatan Multivariate Adaptive Regression Splines (MARS) untuk memodelkan tingkat pengangguran terbuka di Jawa Tengah. Seperti yang telah dijelaskan sebelumnya, belum ada penelitian yang membandingkan teknik GCV dan UBR terkait hubungan TPT (Tingkat Pengangguran

Terbuka) di Jawa Tengah. Penelitian ini bertujuan untuk mengidentifikasi faktor-faktor yang mempengaruhi TPT di Jawa Tengah melalui analisis perbandingan pendekatan GCV dan UBR, serta identifikasi titik knot optimal untuk regresi nonparametrik spline truncated.

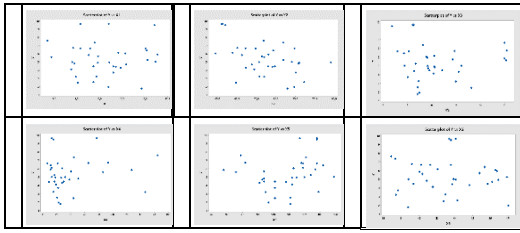
## **2. METODE PENELITIAN**

### **2.1 Data Penelitian**

Fokus utama penelitian ini adalah data TPT dari situs resmi BPS untuk kabupaten atau kota di Provinsi Jawa Tengah pada tahun 2022. Faktor-faktor respon, khususnya TPT di Jawa Tengah pada tahun 2022, dan variabel-variabel lain yang dianggap berdampak pada TPT di Jawa Tengah pada tahun 2022 menjadi variabel yang akan digunakan dalam studi ini.

### **2.2 Deskripsi Data**

Scatter plot digunakan untuk memvisualisasikan pola data pada variabel tingkat pengangguran terbuka dengan keenam variabel prediktor sebelum data tersebut dievaluasi. Metode regresi yang tepat dipilih berdasarkan pola hubungan antar variabel.



Gambar 1. Pola Hubungan Variabel Prediktor terhadap Variabel Respon

Tingkat pengguran terbuka dengan variabel presentase penduduk miskin ( $X_1$ ), persentase tingkat partisipasi angkatan kerja ( $X_2$ ), persentase rata lama sekolah ( $X_3$ ), PDRB per kapita atas dasar harga konstan ( $X_4$ ), rasio jenis kelamin ( $X_5$ ), angka ketergantungan ( $X_6$ ) menunjukkan variabilitas pada sub-interval tertentu dan tidak memiliki pola yang konsisten. Karena regresi parametrik tidak dapat diterima dengan pola data ini, regresi nonparametrik akan digunakan dalam investigasi ini.

### 2.3 Pemodelan dengan Metode GCV

Menemukan titik knot yang ideal adalah langkah pertama sebelum melakukan pemodelan. Hasil pemilihan titik knot terbaik menggunakan pendekatan GCV dengan satu, dua, dan tiga titik knot adalah sebagai berikut.

Tabel 1. Ringkasan Hasil Pemilihan Titik Knot Metode GCV

Titik Knot	Nilai GCV
Satu Titik Knot	3,369
Dua Titik Knot	2,747
Tiga Titik Knot	1,691

Dengan nilai GCV minimum sebesar 1,691, seperti yang ditunjukkan pada Tabel 1, penggunaan tiga titik knot merupakan pilihan terbaik dalam menggunakan metode GCV. Hasil estimasi parameter model regresi nonparametrik spline trimmed dengan menggunakan tiga titik knot dan pendekatan GCV ditunjukkan di bawah ini.

Tabel 2. Estimasi Parameter Model GCV

Parameter	Estimator
$\beta_0$	$\beta_0 = 113,2561$
$\beta_1$	$\beta_1 = -0,3456$
$\beta_2$	$\beta_2 = -7,2409$
$\beta_3$	$\beta_3 = 13,4394$
$\beta_4$	$\beta_4 = -6,1863$
$\beta_5$	$\beta_5 = -0,5099$
$\beta_6$	$\beta_6 = 2,1960$
$\beta_7$	$\beta_7 = -16,4572$
$\beta_8$	$\beta_8 = 30,2859$

$\beta_9$	$\beta_9 = -2,0992$
$\beta_{10}$	$\beta_{10} = -19,8605$
$\beta_{11}$	$\beta_{11} = 14,9944$
$\beta_{12}$	$\beta_{12} = 28,6881$
$\beta_{13}$	$\beta_{13} = 0,0601$
$\beta_{14}$	$\beta_{14} = -0,0082$
$\beta_{15}$	$\beta_{15} = -0,3511$
$\beta_{16}$	$\beta_{16} = 0,9017$
$\beta_{17}$	$\beta_{17} = -0,7121$
$\beta_{18}$	$\beta_{18} = 9,3824$
$\beta_{19}$	$\beta_{19} = -16,3310$
$\beta_{20}$	$\beta_{20} = 11,0573$
$\beta_{21}$	$\beta_{21} = 0,4049$
$\beta_{22}$	$\beta_{22} = 2,0029$
$\beta_{23}$	$\beta_{23} = -0,8966$
$\beta_{24}$	$\beta_{24} = -7,5991$

Model regresi nonparametrik terbaik yang menggunakan pendekatan GCV, dengan spline dan titik knot yang lebih pendek, adalah sebagai berikut, berdasarkan hasil estimasi parameter yang telah disebutkan sebelumnya.

$$\hat{y} = 113,2561 - 0,3456x_1 - 7,2409(x_1 - 12,439)_+ + 13,4394(x_1 - 13,432)_+ - 6,1863(x_1 - 14,921)_+ - 0,5099x_2 + 2,1960(x_2 - 74,731)_+ - 16,4572(x_2 - 75,941)_+ + 30,2859(x_2 - 77,755)_+ - 2,0992x_3 - 19,8605(x_3 - 9,541)_+ + 14,9944(x_3 - 9,927)_+ + 28,6881(x_3 - 10,506)_+ + 0,0601x_4 - 0,0082(x_4 - 66,403)_+ - 0,3511(x_4 - 72,845)_+ + 0,9017(x_4 - 82,508)_+ - 0,7121x_5 + 9,3824(x_5 - 101,582)_+ - 16,3310(x_5 - 102,161)_+ + 11,0573(x_5 - 103,031)_+ + 0,4049x_6 + 2,0029(x_6 - 44,875)_+ - 0,8966(x_6 - 45,414)_+ - 7,5991(x_6 - 46,222)_+$$

Untuk mengetahui apakah parameter pada variabel prediktor memiliki pengaruh yang signifikan atau tidak terhadap variabel respon, maka dalam pengujian model juga dilakukan uji signifikansi parameter.

Tabel 3. Pengujian Parameter Model Regresi Secara Serentak

Sumber	Df	SS	MS	$F_{hitung}$	$p$ -value
Regresi	24	126,59	5,27	10,914	0,000209
Error	10	4,8327	0,48		
Total	34	131,4			

Tabel 3 menunjukkan bahwa nilai  $F_{hitung} > F_{0,05(10,24)}$  yaitu  $10,9142 > 2,25$  dan  $p$ -value  $< 0,05$  yaitu  $0,0002089 < 0,05$  maka  $H_0$  ditolak yang berarti bahwa secara bersama-sama variabel prediktor memiliki pengaruh yang signifikan terhadap variabel respon.

Tabel 4. Pengujian Secara Parsial

Var	Par	$t_{hitung}$	$p$ -value	Ket
	$\beta_0$	2,8816	0,016340	Signifikan
$x_1$	$\beta_1$	-2,6326	0,025050	Signifikan

	$\beta_2$	-3,0383	0,01249886	Signifikan
	$\beta_3$	3,1766	0,00987576	Signifikan
	$\beta_4$	-2,5521	0,02875918	Signifikan
$x_2$	$\beta_5$	-5,6611	0,00020921	Signifikan
	$\beta_6$	1,9805	0,07580948	Tidak
	$\beta_7$	-2,1736	0,05483985	Tidak
	$\beta_8$	2,3405	0,04130179	Signifikan
$x_3$	$\beta_9$	-6,3210	8,6694e-05	Signifikan
	$\beta_{10}$	-2,0134	0,07176719	Tidak
	$\beta_{11}$	0,3843	0,7087824	Tidak
	$\beta_{12}$	0,3919	0,7033686	Tidak
$x_4$	$\beta_{13}$	2,7351	0,02100699	Signifikan
	$\beta_{14}$	-0,0169	0,9868286	Tidak
	$\beta_{15}$	-0,4198	0,6834811	Tidak
	$\beta_{16}$	1,2758	0,2308745	Tidak
$x_5$	$\beta_{17}$	-2,2898	0,04503002	Signifikan
	$\beta_{18}$	5,7672	0,0001809	Signifikan
	$\beta_{19}$	-6,1909	0,00010267	Signifikan
	$\beta_{20}$	4,4368	0,00126073	Signifikan
$x_6$	$\beta_{21}$	1,8291	0,09731981	Tidak
	$\beta_{22}$	1,1616	0,2723652	Tidak
	$\beta_{23}$	-0,2770	0,7874236	Tidak
	$\beta_{24}$	2,2655	0,04693089	Signifikan

Hasil dari statistik uji pada Tabel 5 memperlihatkan hasil uji secara parsial. Variabel dapat dikatakan signifikan jika pada setiap variabel prediktor memiliki minimal satu  $H_0$  yang ditolak. Hal ini berarti bahwa variabel  $x_1, x_2, x_3, x_4, x_5,$  dan  $x_6$  memiliki parameter yang menolak  $H_0$  sehingga dapat dikatakan bahwa variabel-variabel tersebut memiliki pengaruh yang signifikan terhadap variabel respon. Meskipun 10 parameter yang tidak signifikan, namun semua variabel berpengaruh terhadap tingkat pengangguran terbuka.

tahap selanjutnya melakukan uji asumsi residual. Untuk hasil pengujian asumsi identic, diperlihatkan pada Tabel 5.

Tabel 5. Uji Gleiser Metode GCV

Sumber	Df	SS	MS	$F_{hitung}$	$p - value$
Regresi	24	1,801	0,075	0,8397	0,6556
Error	10	0,894	0,089		
Total	34	2,695			

Tabel 5 menunjukkan bahwa nilai  $F_{hitung} > F_{0,05(10,24)}$  yaitu  $0,8397 < 2,25$  atau  $p - value > \alpha$  yaitu  $0,6556 > 0,05$  maka  $H_0$  gagal ditolak yang berarti bahwa terdapat homoskedastisitas atau tidak terdapat heteroskedastisitas pada residu model regresi.

Uji Durbin Watson menghasilkan nilai  $p - value > \alpha$  yaitu  $0,1 > 0,05$  maka  $H_0$  gagal

ditolak yang berarti bahwa tidak terdapat autokorelasi pada residu. Uji Kolmogorov-Smirnov menghasilkan nilai  $|D_{hitung}| > D_{35;0,05} = 0,22424$  dan nilai  $p - value > \alpha$  yaitu  $0,3662143 > 0,05$  maka  $H_0$  gagal ditolak yang berarti bahwa residu yang dihasilkan dari model regresi nonparametrik *spline truncated* berdistribusi normal.

## 2.4 Pemodelan dengan Metode UBR

Hasil dari pemilihan simpul ideal metode UBR dengan menggunakan satu, dua, dan tiga simpul adalah sebagai berikut.

Tabel 6 Ringkasan Hasil Pemilihan Titik Knot Metode UBR

Titik Knot	Nilai UBR
Satu Titik Knot	4,681e-12
Dua Titik Knot	3,787e-12
Tiga Titik Knot	2,356E-23

Berdasarkan Tabel 6, penggunaan tiga titik knot merupakan pilihan terbaik ketika menggunakan metode UBR, karena nilai UBR terkecil adalah 2,356E-23.

Hasil estimasi parameter model regresi nonparametrik *spline trimmed* dengan menggunakan tiga titik knot dan pendekatan UBR ditunjukkan di bawah ini.

Tabel 7 Estimasi Parameter Model UBR

Parameter	Estimator
$\beta_0$	$\beta_0 = 0,5854$
$\beta_1$	$\beta_1 = -0,2883$
$\beta_2$	$\beta_2 = 0,5548$
$\beta_3$	$\beta_3 = -1,4408$
$\beta_4$	$\beta_4 = 2,1359$
$\beta_5$	$\beta_5 = -0,4823$
$\beta_6$	$\beta_6 = 1,1581$
$\beta_7$	$\beta_7 = -0,9144$
$\beta_8$	$\beta_8 = 0,2767$
$\beta_9$	$\beta_9 = -1,0447$
$\beta_{10}$	$\beta_{10} = -2,4020$
$\beta_{11}$	$\beta_{11} = -0,6967$
$\beta_{12}$	$\beta_{12} = 8,9144$
$\beta_{13}$	$\beta_{13} = 0,0398$
$\beta_{14}$	$\beta_{14} = -0,2987$
$\beta_{15}$	$\beta_{15} = 0,3531$
$\beta_{16}$	$\beta_{16} = -0,0022$
$\beta_{17}$	$\beta_{17} = 0,3825$
$\beta_{18}$	$\beta_{18} = -2,8950$
$\beta_{19}$	$\beta_{19} = 6,2393$
$\beta_{20}$	$\beta_{20} = -5,8805$

$\beta_{21}$	$\beta_{21} = 0,2203$
$\beta_{22}$	$\beta_{22} = 1,3879$
$\beta_{23}$	$\beta_{23} = 1,1122$
$\beta_{24}$	$\beta_{24} = -5,4335$

Berdasarkan hasil estimasi parameter di atas, model regresi nonparametrik *spline truncated* dengan titik knot optimal menggunakan metode UBR yang terbentuk adalah sebagai berikut.

$$\hat{y} = 0,5854 - 0,2883x_1 + 0,5548(x_1 - 10,454)_+ - 1,4408(x_1 - 11,447)_+ + 2,1359(x_1 - 13,184)_+ - 0,4823x_2 + 1,1581(x_2 - 72,311)_+ - 0,9144(x_2 - 73,521)_+ + 0,2767(x_2 - 75,638)_+ - 1,0447x_3 - 2,4020(x_3 - 8,768)_+ - 0,6967(x_3 - 9,154)_+ + 8,9144(x_3 - 9,830)_+ + 0,0398x_4 - 0,2987(x_4 - 53,520)_+ + 0,3531(x_4 - 59,962)_+ - 0,0022(x_4 - 71,235)_+ + 0,3825x_5 - 2,8950(x_5 - 100,422)_+ + 6,2393(x_5 - 101,002)_+ - 5,8805(x_5 - 102,016)_+ + 0,2203x_6 + 1,3879(x_6 - 43,797)_+ + 1,1122(x_6 - 44,336)_+ - 5,4335(x_6 - 45,279)_+$$

Untuk mengetahui apakah parameter pada variabel prediktor berpengaruh secara signifikan terhadap variabel respon atau tidak, maka selanjutnya akan dilakukan pengujian terhadap model dengan menggunakan uji signifikansi parameter.

**Tabel 8.** Uji Parameter Model Regresi Secara Serentak

Sumber	Df	SS	MS	$F_{hitung}$	$p - value$
Regresi	24	113,286	4,72	2,5938	0,05930
Eror	10	18,1980	1,82		
Total	34	131,484			

Tabel 8 menunjukkan bahwa nilai  $F_{hitung} > F_{0,05(10;24)}$  yaitu  $2,5938 > 2,25$  dan  $0,059302 < 0,05$  yaitu  $0,0002089 < 0,05$  maka  $H_0$  ditolak yang berarti bahwa secara bersama-sama variabel prediktor memiliki pengaruh yang signifikan terhadap variabel respon.

**Tabel 9.** Pengujian Parameter Model Regresi Secara Parsial

Var	Par	$t_{hitung}$	$p - value$	Ket
	$\beta_0$	1,1239	0,2873224	Tidak
$X_1$	$\beta_1$	-0,7074	0,4954742	Tidak
	$\beta_2$	0,3534	0,731092	Tidak
	$\beta_3$	-0,7388	0,4769847	Tidak
	$\beta_4$	1,5119	0,161484	Tidak
$X_2$	$\beta_5$	-2,5556	0,02858561	Signifikan
	$\beta_6$	0,8392	0,4209566	Tidak
	$\beta_7$	-0,4605	0,6549699	Tidak
	$\beta_8$	0,2442	0,8120405	Tidak
$X_3$	$\beta_9$	-1,2966	0,2238898	Tidak
	$\beta_{10}$	-0,4780	0,6428962	Tidak
	$\beta_{11}$	-0,0813	0,93683	Tidak
	$\beta_{12}$	1,2336	0,2455565	Tidak
$X_4$	$\beta_{13}$	0,9736	0,3532287	Tidak
	$\beta_{14}$	-0,4329	0,6742638	Tidak

	$\beta_{15}$	0,3531	0,731349	Tidak
	$\beta_{16}$	-0,0054	0,9958253	Tidak
$X_5$	$\beta_{17}$	1,4596	0,1750751	Tidak
	$\beta_{18}$	-1,5371	0,1552825	Tidak
	$\beta_{19}$	2,2049	0,05199916	Tidak
	$\beta_{20}$	-2,6337	0,02500206	Signifikan
$X_6$	$\beta_{21}$	0,4603	0,6551287	Tidak
	$\beta_{22}$	0,2939	0,7748253	Tidak
	$\beta_{23}$	0,1560	0,8790726	Tidak
	$\beta_{24}$	-1,4571	0,1757624	Tidak

Hasil dari statistik uji pada Tabel 9 menunjukkan hasil pengujian secara parsial. Variabel dapat dikatakan signifikan jika pada setiap variabel prediktor memiliki minimal satu  $H_0$  yang ditolak. Hal ini berarti bahwa variabel  $X_2$  dan  $X_5$  memiliki parameter yang menolak  $H_0$  sehingga dapat dikatakan bahwa variabel-variabel tersebut memiliki pengaruh yang signifikan terhadap variabel respon.

Langkah berikutnya yaitu uji asumsi residual. yang dilakukan untuk melihat, apakah residual memenuhi asumsi identik, independen, dan berdistribusi normal (*IIDN*) atau tidak. Hasil uji asumsi identik disajikan dalam Tabel 10.

**Tabel 10.** Uji Glejser Metode UBR

Sumber	Df	SS	MS	$F_{hitung}$	$p - value$
Regresi	24	6,7733	0,2822	1,9330	0,1389
Residu	10	1,4599	0,1459		
Total	34	8,2333			

Tabel 10 menunjukkan bahwa nilai  $F_{hitung} > F_{0,05(10;24)} = 2,25$  atau  $p - value > \alpha$  yaitu  $0,1389 > 0,05$  maka  $H_0$  gagal ditolak yang berarti bahwa terdapat homoskedastisitas atau tidak terdapat heteroskedastisitas pada residu model regresi.

Uji Durbin Watson menghasilkan nilai  $p - value > \alpha$  yaitu  $0,98 > 0,05$  maka  $H_0$  gagal ditolak yang berarti bahwa terdapat tidak autokorelasi pada residu. Uji Kolmogorov-Smirnov menghasilkan nilai  $|D_{hitung}| = 0,09463417 < D_{35;0,05} = 0,22424$  dan  $p - value = 0,8834861 > \alpha = 0,05$  maka  $H_0$  gagal ditolak yang berarti bahwa residu berdistribusi normal.

## 2.5 Perbandingan Metode GCV dan Metode UBR

Berikut ditampilkan tabel perbandingan metode UBR dan GCV dalam pemodelan regresi nonparametrik *spline truncated* pada kasus tingkat pengangguran terbuka di Jawa Tengah tahun 2022.

Tabel 11. Perbandingan Metode UBR dan GCV

	GCV	UBR
MSE	1,381e-01	5,199e-01
$R_{Adj}^2$	95,69%	83,77%

Berdasarkan Tabel 11 dapat dilihat bahwa metode GCV menghasilkan nilai MSE yang lebih kecil yaitu sebesar 1,381e-01 serta  $R_{Adj}^2$  juga yang lebih besar yaitu sebesar 95,69% dibandingkan dengan metode UBR. Berdasarkan hal tersebut, pada kasus tingkat pengangguran terbuka di Provinsi Jawa Tengah pada tahun 2022, pendekatan GCV merupakan metode yang lebih unggul dalam pemilihan titik knot optimal dibandingkan dengan metode UBR pada pemodelan regresi nonparametrik spline truncated.

### 3. KESIMPULAN

#### 3.1 Simpulan

Berdasarkan analisis yang telah dilakukan dapat ditarik kesimpulan sebagai berikut.

1. Model regresi nonparametrik spline trimmed untuk tingkat pengangguran terbuka, yang diperoleh dengan menggunakan pendekatan GCV minimal dengan tiga titik knot, terlihat seperti ini.

$$\hat{y} = 113,2561 - 0,3456x_1 - 7,2409(x_1 - 12,439)_+ + 13,4394(x_1 - 13,432)_+ - 6,1863(x_1 - 14,921)_+ - 0,5099x_2 + 2,1960(x_2 - 74,731)_+ - 16,4572(x_2 - 75,941)_+ + 30,2859(x_2 - 77,755)_+ - 2,0992x_3 - 19,8605(x_3 - 9,541)_+ + 14,9944(x_3 - 9,927)_+ + 28,6881(x_3 - 10,506)_+ + 0,0601x_4 - 0,0082(x_4 - 66,403)_+ - 0,3511(x_4 - 72,845)_+ + 0,9017(x_4 - 82,508)_+ - 0,7121x_5 + 9,3824(x_5 - 101,582)_+ - 16,3310(x_5 - 102,161)_+ + 11,0573(x_5 - 103,031)_+ + 0,4049x_6 + 2,0029(x_6 - 44,875)_+ - 0,8966(x_6 - 45,414)_+ - 7,5991(x_6 - 46,222)_+$$

2. Pendekatan UBR dengan tiga titik knot menghasilkan model regresi nonparametrik tereduksi spline minimum untuk tingkat pengangguran terbuka sebagai berikut.

$$\hat{y} = 0,5854 - 0,2883x_1 + 0,5548(x_1 - 10,454)_+ - 1,4408(x_1 - 11,447)_+ + 2,1359(x_1 - 13,184)_+ - 0,4823x_2 + 1,1581(x_2 - 72,311)_+ - 0,9144(x_2 - 73,521)_+ + 0,2767(x_2 - 75,638)_+ - 1,0447x_3 - 2,4020(x_3 - 8,768)_+ - 0,6967(x_3 - 9,154)_+ + 8,9144(x_3 - 9,830)_+ + 0,0398x_4 - 0,2987(x_4 - 53,520)_+ + 0,3531(x_4 - 59,962)_+ - 0,0022(x_4 - 71,235)_+ + 0,3825x_5 - 2,8950(x_5 - 100,422)_+ + 6,2393(x_5 - 101,002)_+ - 5,8805(x_5 - 102,016)_+ + 0,2203x_6 + 1,3879(x_6 - 43,797)_+ + 1,1122(x_6 - 44,336)_+ - 5,4335(x_6 - 45,279)_+$$

Hasil analisis data menunjukkan bahwa, dibandingkan dengan pendekatan UBR, metode GCV menghasilkan  $R^2$  yang lebih besar yaitu

95,69% dan nilai MSE yang lebih kecil yaitu 1,381e-01. Berdasarkan hal tersebut, pada kasus tingkat pengangguran terbuka di Provinsi Jawa Tengah tahun 2022, pendekatan GCV merupakan metode yang lebih unggul dalam pemilihan titik knot optimal dibandingkan dengan metode UBR pada pemodelan regresi nonparametrik spline truncated.

#### 3.2 Saran

Penekanan penelitian yang telah dicapai dapat diterapkan pada penyelidikan lebih lanjut. Berikut ini adalah beberapa contoh pengembangan:

1. Untuk memastikan apakah ada dampak kedekatan lokasi antara lokasi-lokasi yang diteliti, penelitian selanjutnya dapat mempertimbangkan variabel-variabel tambahan yang tidak termasuk dalam penelitian ini dan mencoba menggunakan teknik-teknik alternatif seperti pendekatan spasial.
2. Untuk melakukan perbandingan yang lebih menyeluruh, teknik pemilihan titik knot lainnya dengan menggunakan data tingkat pengangguran terbuka, termasuk GML, dapat diteliti.
3. Dapat mengembangkan 4 atau 5 titik knot yang digunakan dalam penelitian.

#### PERNYATAAN PENGHARGAAN

Kepada pembimbing pertama, Ibu Dra. RR Sri Sulistijowati H, M.Si. dan pembimbing kedua, Bapak Drs. Isnandar Slamet, M.Sc., Ph.D., penulis mengucapkan terima kasih yang sebesar-besarnya atas waktu, tenaga, dan pikiran yang telah diberikan kepada penulis untuk membimbing penulis dalam menyelesaikan skripsi ini.

#### DAFTAR PUSTAKA

- Astuti, W. I., Ratnasari, V., & Wibowo, W. (2017). Analisis faktor yang berpengaruh terhadap tingkat pengangguran terbuka di Provinsi Jawa Timur menggunakan regresi data panel. *Jurnal Sains Dan Seni ITS*, 6(1), D144-D149.
- Dani, A. T. R., & Adrianingsih, N. Y. (2021). Pemodelan Regresi Nonparametrik

- engan Estimator Spline Truncated vs Deret Fourier. *Jambura Journal of Mathematics*, 3(1), 26-36.
- Eubank, R. L. (1999). *Nonparametric regression and spline smoothing*. CRC press.
- Hanani Wijaya, A. F. (2018). *Analisis Faktor-Faktor yang Mempengaruhi Tingkat Pengangguran Terbuka (TPT) di Provinsi Aceh dengan Regresi Nonparametrik Spline Truncated* (Skripsi Sarjana). Institut Teknologi Sepuluh November.
- Kurnianto, D., Arya, M. A. N., Kharisudin, I., & Fauzi, F. (2021, February). Analisis Regresi Spasial dengan Pembobot Queen Contiguity pada Tingkat Pengangguran Terbuka di Povinsi Jawa Tengah Tahun 2019. In *PRISMA, Prosiding Seminar Nasional Matematika (Vol. 4, pp. 595-601)*.
- Kurniawati, N. A., & Budiantara, I. N. (2020). Pemodelan Tingkat Pengangguran Terbuka di Provinsi Jawa Barat Menggunakan Regresi Nonparametrik Spline Truncated. *Jurnal Sains dan Seni ITS*, 8(2), D334-D340.
- Nasuha, F. (2016). *Pemodelan Tingkat Pengangguran Terbuka di Kabupaten/Kota di Provinsi Jawa Tengah dengan Pendekatan Multivariate Adaptive Regression Splines (MARS)*. (Skripsi, Institut Teknologi Sepuluh Nopember).
- Putri, P. S. (2023). *Pemodelan Faktor-Faktor yang Mempengaruhi Indeks Pembangunan Manusia Kabupaten/Kota di Provinsi Jawa Tengah Menggunakan Regresi Semiparametrik Spline Truncated* (Skripsi). Universitas Sebelas Maret, Program Studi Statistika.
- Sari, S. U. R. (2016). *Perbandingan Model Regresi Nonparametrik Spline Multivariabel dengan Menggunakan Metode GCV dan UBR Dalam Pemilihan Titik Knot Optimal (Studi Kasus Data Angka Kematian Maternal di Jawa Timur)* (Doctoral dissertation, Institut Teknologi Sepuluh Nopember).
- Utami, T. W. (2018). Pendekatan Regresi Semiparametrik Spline Truncated untuk Pemodelan Tingkat Pengangguran Terbuka di Jawa Tengah. *Jurnal Statistika Universitas Muhammadiyah Semarang*, 6(2).