

**PERBANDINGAN METODE *TRUNCATED SPLINE* DAN *B-SPLINE*
DALAM PEMODELAN TINGKAT PENGANGGURAN
TERBUKA DI WILAYAH SUMBAGSEL**

Skripsi

Oleh

**NAZWA FATONAH AZ-ZAHRA
2217031022**



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

ABSTRACT

COMPARISON OF TRUNCATED SPLINE AND B-SPLINE METHODS IN MODELING THE OPEN UNEMPLOYMENT RATE IN THE SUMBAGSEL REGION

By

NAZWA FATONAH AZ-ZAHRA

Nonparametric regression is a statistical method for modeling variables with nonlinear patterns where the functional form is not precisely known. This study compares the Truncated Spline and B-Spline methods to obtain an optimal model for handling highly fluctuated data through strategic knot selection. The modeling process involved simulations of various orders (linear, quadratic, cubic) and knot variations based on the minimum Generalized Cross Validation (GCV) criterion. The methods were applied to 2024 Open Unemployment Rate (TPT) data in the Sumbagsel region, using predictor variables such as population density (X_1), HDI (X_2), male labor force (X_3), female labor force (X_4), job seekers (X_5), and expected years of schooling (X_6). The results indicate that the linear Truncated Spline with one knot is the superior model, yielding a GCV of 0.5121, an MSE of 0.3270, and an R^2 of 66.74%. This demonstrates that the Truncated Spline approach is more effective and accurate than the B-Spline method in mapping labor dynamics within the region.

Keywords: Open Unemployment Rate, Nonparametric Regression, Truncated Spline, B-Spline, GCV, Sumbagsel.

ABSTRAK

PERBANDINGAN METODE *TRUNCATED SPLINE* DAN *B-SPLINE* DALAM PEMODELAN TINGKAT PENGANGGURAN TERBUKA DI WILAYAH SUMBAGSEL

Oleh

NAZWA FATONAH AZ-ZAHRA

Regresi nonparametrik merupakan metode statistika untuk memodelkan hubungan variabel yang polanya bersifat nonlinier dan tidak diketahui bentuk fungsinya secara pasti. Penelitian ini membandingkan metode *Truncated Spline* dan *B-Spline* guna memperoleh model terbaik dalam menangani data berfluktuasi tajam melalui pemilihan titik *knot* optimal. Pemodelan dilakukan melalui simulasi orde linear, kuadratik, dan kubik serta variasi jumlah titik *knot* berdasarkan kriteria *Generalized Cross Validation* (GCV) minimum. Metode tersebut diterapkan pada data Tingkat Pengangguran Terbuka wilayah Sumbagsel tahun 2024 dengan variabel prediktor meliputi kepadatan penduduk (X_1), Indeks Pembangunan Manusia (X_2), angkatan kerja laki-laki (X_3), angkatan kerja perempuan (X_4), pencari kerja (X_5), dan harapan lama sekolah (X_6). Hasil penelitian menunjukkan bahwa model *Truncated Spline* linear dengan satu titik *knot* merupakan model terbaik dengan nilai GCV sebesar 0,5121, MSE sebesar 0,3270, dan R^2 sebesar 66,74% yang menunjukkan bahwa pendekatan *Truncated Spline* lebih efektif dan akurat dibandingkan metode *B-Spline* dalam memetakan dinamika ketenagakerjaan di wilayah tersebut.

Kata-kata kunci: Tingkat Pengangguran Terbuka, Regresi Nonparametrik, *Truncated Spline*, *B-Spline*, GCV, Sumbagsel.

**PERBANDINGAN METODE *TRUNCATED SPLINE* DAN *B-SPLINE*
DALAM PEMODELAN TINGKAT PENGANGGURAN
TERBUKA DI WILAYAH SUMBAGSEL**

NAZWA FATONAH AZ-ZAHRA

Skripsi

Sebagai Salah Satu Syarat untuk Memperoleh Gelar
SARJANA MATEMATIKA

Pada

Jurusan Matematika

Fakultas Matematika dan Ilmu Pengetahuan Alam



**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS LAMPUNG
BANDAR LAMPUNG**

2026

Judul Skripsi : **PERBANDINGAN METODE *TRUNCATED SPLINE* DAN *B-SPLINE* DALAM PEMODELAN TINGKAT PENGANGGURAN TERBUKA DI WILAYAH SUMBAGSEL**

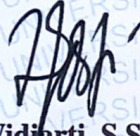
Nama Mahasiswa : **Nazwa Fatonah Az-Zahra**


Nomor Pokok Mahasiswa : **2217031022**

Jurusan : **Matematika**

Fakultas : **Matematika dan Ilmu Pengetahuan Alam**




Widiarti, S.Si., M.Si.
NIP 198005022005012003


Dr. Edwin Russel, S.E., M.Sc.
NIP 198406192024061001

2. Ketua Jurusan Matematika Matematika


Dr. Aang Nuryaman, S.Si., M.Si.
NIP. 197403162005011001

MENGESAHKAN

1. Tim Penguji

Ketua : Widiarti, S.Si., M.Si.

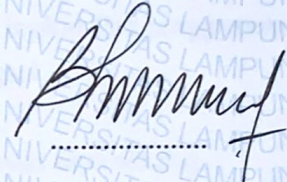


Sekretaris : Dr. Edwin Russel, S.E., M.Sc.



Penguji

**Bukan Pembimbing : Dr. Bernadhita Herindri
Samodera Utami, S.Si., M.Sc.**



2. Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam

Dr. Eng. Heri Satria, S.Si., M.Si.

NIP 197410012005011002



Tanggal Lulus Ujian Skripsi: 10 Juni 2026

PERNYATAAN SKRIPSI MAHASISWA

Yang bertanda tangan di bawah ini:

Nama : **Nazwa Fatonah Az-Zahra**
Nomor Pokok Mahasiswa : **2217031022**
Jurusan : **Matematika**
Judul Skripsi : **Perbandingan Metode *Truncated Spline* dan *B-Spline* dalam Pemodelan Tingkat Pengangguran Terbuka di Wilayah Sumbagsel**

Dengan ini menyatakan bahwa skripsi ini adalah hasil pekerjaan saya sendiri. Apabila kemudian hari terbukti bahwa skripsi ini merupakan hasil salinan atau dibuat oleh orang lain, maka saya bersedia menerima sanksi sesuai dengan ketentuan akademik yang berlaku.

Bandar Lampung, 10 Juni 2026

Penulis,



Nazwa Fatonah Az-Zahra

RIWAYAT HIDUP

Penulis bernama lengkap Nazwa Fatonah Az-Zahra, lahir di Bandar Lampung pada tanggal 16 April 2005. Penulis merupakan anak pertama dari pasangan Bapak Suprayitno dan Ibu Siti Maunah.

Penulis mengawali pendidikan di Taman Kanak-Kanak (TK) Tunas Karya pada tahun 2009 dan menempuh pendidikan dasar di SD Negeri 1 Perumnas Way Halim pada tahun 2010–2016. Kemudian, penulis melanjutkan jenjang pendidikan di MTs Negeri 2 Bandar Lampung pada tahun 2016–2019 dan Sekolah Menengah Atas (SMA) di SMA Negeri 15 Bandar Lampung.

Penulis diterima sebagai mahasiswi Program Studi S1 Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung melalui jalur SNMPTN pada tahun 2022. Selama menjadi mahasiswa, penulis aktif berpartisipasi dalam kegiatan tahunan Dies Natalis Jurusan Matematika, dengan menjabat sebagai Anggota Divisi Kesekretariatan pada tahun 2023, Anggota Divisi Kreatif pada tahun 2024, serta Anggota Divisi Kreatif Dies Natalis FMIPA Universitas Lampung pada tahun 2024.

Pada bulan Januari–Februari 2025, penulis melaksanakan Praktik Kerja Lapangan (PKL) di Perum Bulog Kantor Wilayah Lampung. Selanjutnya, pada bulan Juli–Agustus 2025, penulis melaksanakan Kuliah Kerja Nyata (KKN) di Kelurahan Gedong Pakuon, Kecamatan Teluk Betung Selatan, Kota Bandar Lampung, Provinsi Lampung. Kemudian, pada bulan Agustus–Oktober 2025, penulis melaksanakan magang di BRI Insurance Lampung.

KATA INSPIRASI

"And Allah is the best of planners."

(QS. Ali 'Imran: 54)

"Indeed, with hardship [will be] ease."

(QS. Al-Insyirah: 5)

"Dan sebaik-baiknya manusia adalah orang yang bermanfaat bagi orang yang lain"

(HR. Ath-Thabrani)

"No one can save you but yourself."

(Pamungkas ~ Be Okay Again Today)

PERSEMBAHAN

Segala puji bagi Allah SWT atas rahmat dan karunia-Nya, sehingga karya ini dapat diselesaikan. Dengan penuh rasa syukur, karya ini penulis persembahkan kepada:

Ayah dan Bunda tercinta

Sosok luar biasa yang selalu menjadi sumber kekuatan, yang tak pernah lelah memberikan kasih sayang, doa, dan dukungan. Semoga karya ini menjadi wujud kecil dari rasa terima kasih yang tak terhingga.

Dosen Pembimbing dan Penguji

Terima kasih kepada Bapak/Ibu dosen yang telah memberikan bimbingan, arahan, motivasi, serta ilmu yang sangat berharga selama proses penyusunan karya ini.

Sahabat dan orang-orang terdekat

Terima kasih telah selalu hadir dalam setiap proses, terutama di saat sulit. Dukungan, doa, dan kebersamaan kalian sangat berarti dan menjadi kekuatan bagi penulis.

Almamater Tercinta,
Universitas Lampung

SANWACANA

Puji syukur kehadiran Allah Swt. atas rahmat, nikmat, dan karunia-Nya, sehingga penulis dapat menyelesaikan skripsi yang berjudul “Perbandingan Metode *Truncated Spline* dan *B-Spline* dalam Pemodelan Tingkat Pengangguran Terbuka di Wilayah Sumbagsel.”

Terselesainya skripsi ini tidak terlepas dari bantuan berbagai pihak, baik secara langsung maupun tidak langsung. Oleh karena itu, pada kesempatan ini penulis mengucapkan terima kasih kepada:

1. Ibu Widiarti, S.Si., M.Si., selaku pembimbing I yang telah dengan penuh kesabaran memberikan bimbingan, arahan, serta motivasi kepada penulis hingga skripsi ini dapat terselesaikan dengan baik.
2. Bapak Dr. Edwin Russel, S.E., M.Sc., selaku pembimbing II yang telah memberikan dukungan, arahan, serta masukan kepada penulis selama proses penyusunan skripsi ini.
3. Ibu Dr. Bernadhita Herindri Samodera Utami, S.Si., M.Sc., selaku dosen pembahas yang telah berkenan menguji serta dengan sabar memberikan masukan, kritik, dan saran kepada penulis.
4. Bapak Prof. Dr. La Zakaria, S.Si., M.Sc., selaku dosen pembimbing akademik.
5. Bapak Dr. Aang Nuryaman, S.Si., M.Si., selaku Ketua Jurusan Matematika.
6. Bapak Dr. Eng. Heri Satria., M.Si., selaku Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam.
7. Seluruh dosen, staf, dan karyawan Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam.

8. Kedua orang tua dan keluarga tercinta, Ayah, Bunda, dan Adek, yang senantiasa menghadirkan doa tanpa henti, kasih sayang yang tulus, serta dukungan yang menguatkan penulis dalam setiap langkah hingga skripsi ini dapat terselesaikan.
9. Sahabat kuliah tercinta, Ardila, Erisa, Nadhia, Kinasih, dan Hani, yang telah menemani penulis dalam setiap langkah, baik dalam suka maupun duka. Terima kasih atas kebersamaan, tawa, dan dukungan yang tidak pernah berhenti, terutama di saat penulis merasa lelah.
10. Sahabat-sahabat penulis, Melsha dan Titah yang senantiasa hadir untuk menemani, mendengarkan, serta memberikan dukungan di saat-saat sulit. Serta kepada Zora, Zikri, Daffa, Tara, Febi, Visca, Shafira, Sajda, Nisya, Dita, Areta, dan Hakim, terima kasih atas doa, dukungan, dan kebersamaan yang begitu hangat.
11. Semua pihak yang tidak dapat penulis sebutkan satu per satu, yang telah memberikan bantuan dan dukungan dalam proses penyelesaian skripsi ini.

Bandar Lampung, 10 Juni 2026

Nazwa Fatonah Az-Zahra

DAFTAR ISI

DAFTAR ISI	xiii
DAFTAR TABEL	xv
DAFTAR GAMBAR	xvi
I. PENDAHULUAN	1
1.1 Latar Belakang dan Masalah	1
1.2 Tujuan Penelitian	4
1.3 Manfaat Penelitian	4
II. TINJAUAN PUSTAKA	5
2.1 Statistika Deskriptif	5
2.2 Analisis Regresi	6
2.3 Regresi Nonparametrik	7
2.4 <i>Truncated Spline</i>	8
2.4.1 Menentukan Titik Knot yang Optimal	12
2.5 <i>B-Spline</i>	13
2.5.1 Penaksiran Parameter dalam Model <i>B-Spline</i>	16
2.5.2 Penentuan Model <i>B-Spline</i> yang Optimal	17
2.6 Evaluasi Model	18
2.6.1 Koefisien Determinasi	18
2.7 Tingkat Pengangguran Terbuka (TPT)	19
2.7.1 Kepadatan Penduduk (KP)	19
2.7.2 Indeks Pembangunan Manusia (IPM)	20
2.7.3 Angkatan Kerja Laki-laki dan Perempuan	20
2.7.4 Pencari Kerja	20
2.7.5 Harapan Lama Sekolah (HLS)	21
III. METODE PENELITIAN	22
3.1 Tempat dan Waktu Penelitian	22
3.2 Data Penelitian	22
3.3 Metode Penelitian	23
IV. HASIL DAN PEMBAHASAN	25

4.1	Eksplorasi Data Tingkat Pengangguran Terbuka	25
4.2	Grafik Hubungan Variabel Independen Terhadap Variabel Dependen	29
4.3	Pemodelan Regresi Nonparametrik <i>Truncated Spline</i>	33
4.3.1	Model <i>Truncated Spline</i> linear dengan 1 Titik Knot	33
4.3.2	Model <i>Truncated Spline</i> Linear dengan 2 Titik Knot	34
4.3.3	Model <i>Truncated Spline</i> Linear dengan 3 Titik Knot	36
4.3.4	Model <i>Truncated Spline</i> Kuadratik dengan 1 Titik Knot	37
4.3.5	Model <i>Truncated Spline</i> Kuadratik dengan 2 Titik Knot	38
4.3.6	Model <i>Truncated Spline</i> Kuadratik dengan 3 Titik Knot	39
4.3.7	Model <i>Truncated Spline</i> Kubik dengan 1 Titik Knot	41
4.3.8	Model <i>Truncated Spline</i> Kubik dengan 2 Titik Knot	42
4.3.9	Model <i>Truncated Spline</i> Kubik dengan 3 Titik knot	43
4.4	Pemodelan Regresi Nonparametrik <i>B-Spline</i>	45
4.4.1	Model <i>B-Spline</i> Linear dengan 1 Titik knot	46
4.4.2	Model <i>B-Spline</i> Linear dengan 2 Titik knot	47
4.4.3	Model <i>B-Spline</i> Linear dengan 3 Titik knot	48
4.4.4	Model <i>B-Spline</i> Kuadratik dengan 1 Titik knot	49
4.4.5	Model <i>B-Spline</i> Kuadratik dengan 2 Titik knot	50
4.4.6	Model <i>B-Spline</i> Kuadratik dengan 3 Titik knot	52
4.4.7	Model <i>B-Spline</i> Kubik dengan 1 titik knot	53
4.4.8	Model <i>B-Spline</i> Kubik dengan 2 Titik knot	54
4.4.9	Model <i>B-Spline</i> Kubik dengan 3 Titik knot	55
4.5	Perbandingan Performa Model <i>Truncated Spline</i> dan <i>B-Spline</i>	60
4.5.1	<i>Mean Squared Error</i> Model <i>Truncated Spline</i> dan <i>B-Spline</i>	60
4.5.2	Koefisien Determinasi (R^2) Model <i>Truncated Spline</i> dan <i>B-Spline</i>	61
4.6	Model Terbaik	61
4.7	Grafik Perbandingan Data Aktual dan Prediksi	63
	V. KESIMPULAN	65
	DAFTAR PUSTAKA	66
	LAMPIRAN	69

DAFTAR TABEL

1	Variabel Penelitian	22
2	Statistika Deskriptif Data Tingkat Pengangguran Terbuka dan Prediktornya	26
3	Statistika Deskriptif Data Tingkat Pengangguran Terbuka dan Prediktornya Setelah dilakukan Standardisasi	27
4	Nilai GCV <i>Truncated Spline</i> Linier Satu Titik Knot	34
5	Nilai GCV <i>Truncated Spline</i> Linear Dua Titik Knot	35
6	Nilai GCV <i>Truncated Spline</i> Linear Tiga Titik Knot	36
7	Nilai GCV <i>Truncated Spline</i> Kuadratik Satu Titik Knot	37
8	Nilai GCV <i>Truncated Spline</i> Kuadratik Dua Titik Knot	38
9	Nilai GCV <i>Truncated Spline</i> Kuadratik Tiga Titik Knot	40
10	Nilai GCV <i>Truncated Spline</i> Kubik Satu Titik Knot	41
11	Nilai GCV <i>Truncated Spline</i> Kubik Dua Titik knot	42
12	Nilai GCV <i>Truncated Spline</i> Kubik Tiga Titik knot	43
13	GCV minimum <i>Truncated Spline</i>	44
14	Nilai GCV <i>B-Spline</i> Linear Satu Titik knot	46
15	Nilai GCV <i>B-Spline</i> Linear Dua Titik knot	47
16	Nilai GCV <i>B-Spline</i> Linear Tiga Titik knot	48
17	Nilai GCV <i>B-Spline</i> Kuadratik Satu Titik knot	50
18	Nilai GCV <i>B-Spline</i> Kuadratik Dua Titik knot	51
19	Nilai GCV <i>B-Spline</i> Kuadratik Tiga Titik knot	52
20	Nilai GCV <i>B-Spline</i> Kubik Satu Titik knot	53
21	Nilai GCV <i>B-Spline</i> Kubik Dua Titik knot	54
22	Nilai GCV <i>B-Spline</i> Kubik Tiga Titik knot	56
23	GCV Minimum <i>B-Spline</i>	57
24	Mean Squared Error model <i>Truncated Spline</i> dan <i>B-Spline</i>	60
25	Koefisien Determinasi Model <i>Truncated Spline</i> dan <i>B-Spline</i>	61

DAFTAR GAMBAR

1	Visualisasi Perubahan Pola Data	8
2	Diagram Alur Penelitian	24
3	<i>Scatterplot</i> Kepadatan Penduduk terhadap Tingkat Pengangguran Terbuka	29
4	<i>Scatterplot</i> Indeks Pembangunan Manusia terhadap Tingkat Pengangguran Terbuka	30
5	<i>Scatterplot</i> Angkatan Kerja Laki-laki terhadap Tingkat Pengangguran Terbuka	30
6	<i>Scatterplot</i> Angkatan Kerja Perempuan terhadap Tingkat Pengangguran Terbuka	31
7	<i>Scatterplot</i> Pencari Kerja terhadap Tingkat Pengangguran Terbuka .	32
8	<i>Scatterplot</i> Harapan Lama Sekolah terhadap Tingkat Pengangguran Terbuka	32
9	Grafik Perbandingan Data Aktual dan Data Prediksi	63

I. PENDAHULUAN

1.1 Latar Belakang dan Masalah

Menurut Draper dan Smith (1998), analisis regresi adalah metode statistik yang bertujuan untuk menjelaskan hubungan antara dua atau lebih variabel. Variabel yang dianalisis terdiri atas variabel dependen sebagai respons dan variabel independen sebagai prediktor. Berdasarkan pendekatan dalam mengestimasi kurva regresi, metode regresi dapat dibedakan menjadi regresi parametrik, regresi nonparametrik, dan regresi semiparametrik.

Pendekatan parametrik digunakan apabila bentuk fungsi hubungan antara peubah terikat dan bebas telah diketahui. Namun dalam praktiknya, tidak semua data memiliki bentuk pola hubungan yang diketahui secara jelas. Oleh karena itu, pendekatan yang sesuai untuk digunakan adalah regresi nonparametrik. Menurut Eubank (1999), regresi nonparametrik merupakan metode statistika yang digunakan untuk menganalisis hubungan antara variabel dependen dan variabel independen ketika bentuk kurva regresi tidak diketahui atau pola hubungannya tidak tampak secara pasti. Kurva regresi hanya diasumsikan sebagai suatu fungsi yang halus, yaitu fungsi yang dapat merepresentasikan bentuk hubungan tertentu. Oleh karena itu, regresi nonparametrik bersifat sangat fleksibel dalam memodelkan data. Dalam regresi nonparametrik, terdapat beberapa bentuk fungsi yang dapat digunakan, seperti *spline*, polinomial lokal, deret *Fourier*, dan *wavelet*. Di antara berbagai pendekatan tersebut, fungsi *spline* merupakan metode yang sering digunakan dalam pemodelan regresi nonparametrik.

Spline adalah fungsi yang dibentuk dari gabungan beberapa polinomial terpotong (*truncated*) yang disusun secara bertahap dan tetap mempertahankan kekontinuannya. Keunggulan *spline* terletak pada keberadaan titik-titik knot yang berfungsi sebagai

penghubung antarpotongan ketika terjadi perubahan pola pada data. Titik-titik knot tersebut berada pada setiap segmen polinomial dan diasumsikan halus (*smooth*), sehingga pendekatan *spline* menjadi lebih fleksibel dibandingkan penggunaan polinomial biasa (Eubank, 1999). Dalam pengembangannya, terdapat beberapa bentuk basis *spline*, namun dua yang paling banyak digunakan adalah *truncated spline* dan *B-spline*.

Truncated spline memiliki struktur sederhana sehingga mudah diinterpretasikan, namun memiliki kelemahan karena sensitif terhadap jumlah serta letak titik knot sedangkan *B-spline* menawarkan kestabilan numerik yang lebih baik dan mampu menghasilkan kurva yang lebih halus karena menggunakan basis fungsi yang bersifat lokal (Lyche & Mørken, 2008). Dengan kata lain, pemilihan basis *spline* yang tepat sangat menentukan kualitas model yang dihasilkan.

Terdapat beberapa penelitian terdahulu terkait regresi nonparameterik *spline truncated* dan *B-spline*, diantaranya yaitu Anggreni, dkk. (2018) mengkaji tentang pemodelan *spline truncated* pada jumlah kasus tuberkulosis di Provinsi Bali. Hasil penelitiannya, model terbaik adalah regresi nonparametrik *spline truncated* linear (orde 2) dengan satu titik knot. Selanjutnya penelitian oleh Suwarno (2023) mengenai pemodelan inflasi di Indonesia dengan metode *B-Spline*. Penelitian tersebut menunjukkan bahwa model terbaik diperoleh dengan titik knot optimal. Penelitian lainnya yaitu oleh Handajani, dkk. (2023) mengenai perbandingan pemodelan data suhu di enam stasiun di Jawa menggunakan metode *truncated spline* dan *B-spline*. Hasil penelitian tersebut menunjukkan bahwa model regresi *B-Spline* linear mampu menghasilkan nilai *mean square error* yang relatif kecil dan juga memberikan nilai koefisien determinasi yang lebih tinggi.

Metode *spline* telah banyak digunakan dalam berbagai bidang seperti kesehatan, ekonomi, dan lingkungan, namun penerapannya dalam analisis ketenagakerjaan khususnya dalam memodelkan tingkat pengangguran masih relatif jarang ditemukan. Padahal, karakteristik data ketenagakerjaan yang cenderung fluktuatif dan memiliki pola hubungan nonlinear sangat sesuai untuk dimodelkan menggunakan pendekatan *spline*.

Pengangguran merupakan masalah ketenagakerjaan yang berdampak pada kondisi sosial dan ekonomi daerah. Indikator yang digunakan untuk mengukurnya adalah

Tingkat Pengangguran Terbuka (TPT), yaitu persentase penganggur terhadap total angkatan kerja (Badan Pusat Statistika, 2023). Tingkat Pengangguran Terbuka (TPT) dipengaruhi oleh berbagai faktor, baik dari aspek demografi, kualitas sumber daya manusia, maupun kondisi pasar tenaga kerja.

Kepadatan penduduk dapat memengaruhi TPT karena peningkatan jumlah penduduk tanpa diikuti ketersediaan lapangan kerja akan memperketat persaingan kerja sehingga berpotensi meningkatkan pengangguran (Todaro & Smith, 2015). Dari aspek kualitas sumber daya manusia, Indeks Pembangunan Manusia (IPM) dan Harapan Lama Sekolah (HLS) mencerminkan tingkat pendidikan dan kualitas tenaga kerja. Semakin tinggi IPM dan HLS, semakin baik kualitas tenaga kerja sehingga peluang kerja meningkat dan TPT cenderung menurun (Badan Pusat Statistika, 2023).

Jumlah pencari kerja mencerminkan besarnya tekanan di pasar tenaga kerja, dimana peningkatan jumlah pencari kerja dapat meningkatkan risiko pengangguran jika tidak diimbangi dengan lapangan kerja yang tersedia (Badan Pusat Statistika, 2023). Sementara itu, Angka Kesempatan Kerja (AKL) dan Angka Kesempatan Penempatan (AKP) menunjukkan kemampuan pasar kerja dalam menyerap tenaga kerja. Semakin tinggi nilai kedua indikator tersebut, semakin besar peluang tenaga kerja terserap sehingga dapat menurunkan TPT (Kementerian Ketenagakerjaan RI, 2022).

Berdasarkan permasalahan tersebut, penelitian tertarik untuk menerapkan metode *Truncated Spline* dan *B-spline* pada data ketenagakerjaan di wilayah Sumatera Bagian Selatan tahun 2024 dengan satu variabel dependen yaitu TPT dan 6 variabel independen yaitu KP, IPM, AKL, AKP, PK, dan HLS. Penelitian ini bertujuan untuk membandingkan metode *truncated spline* dan *B-spline* dalam memodelkan tingkat pengangguran terbuka di wilayah Sumbagsel dan berjudul “Perbandingan Metode *Truncated Spline* dan *B-Spline* dalam Pemodelan Tingkat Pengangguran Terbuka di wilayah Sumbagsel”.

1.2 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut.

1. Membentuk model regresi nonparametrik *truncated spline* dan *B-spline* dalam menganalisis pengaruh KP, IPM, AKL, AKP, PK, dan HLS terhadap TPT di wilayah Sumbagsel
2. Membandingkan hasil pemodelan *truncated spline* dan *B-spline* berdasarkan kriteria *Generalized Cross Validation* (GCV), *Mean Square Error* (MSE), dan koefisien determinasi (R^2).
3. Menentukan model terbaik yang dapat menggambarkan hubungan KP, IPM, AKL, AKP, PK, dan HLS terhadap TPT di wilayah Sumbagsel.

1.3 Manfaat Penelitian

Manfaat dari penelitian ini adalah sebagai berikut.

1. Memberikan gambaran mengenai metode *Truncated spline* dan *B-spline* yang sesuai dalam memodelkan data TPT dengan variabel KP, IPM, AKL, AKP, PK, dan HLS.
2. Menjadi referensi bagi peneliti selanjutnya yang ingin menggunakan metode *truncated spline* dan *B-spline* dalam kasus serupa.
3. Menyediakan informasi empiris yang dapat digunakan sebagai bahan pertimbangan dalam pengambilan keputusan terkait pengangguran di wilayah Sumbagsel.

II. TINJAUAN PUSTAKA

2.1 Statistika Deskriptif

Menurut Sudjana (1996), statistika deskriptif adalah bagian dari statistika yang bertugas menyajikan sekumpulan data secara ringkas dan teratur sehingga mampu memberikan informasi pokok dari data tersebut, baik melalui tabel, diagram, maupun ukuran-ukuran numerik. Tujuan utama statistika deskriptif adalah memberikan gambaran yang jelas, teratur, dan informatif mengenai karakteristik suatu data sehingga informasi yang terkandung di dalamnya mudah dipahami.

Hasan (2001) menegaskan bahwa statistika deskriptif dimanfaatkan untuk menyajikan karakteristik umum data melalui berbagai cara, seperti tabel dan grafik, sehingga pola maupun kecenderungan data dapat terlihat. Ditinjau dari jenisnya, statistika deskriptif umumnya mencakup:

1. Ukuran pemusatan data, seperti *mean*, median, dan modus,
2. Ukuran penyebaran data, seperti varians, standar deviasi, jangkauan (*range*), dan rentang antar kuartil,
3. Penyajian distribusi data melalui tabel distribusi frekuensi dan berbagai bentuk diagram.

2.2 Analisis Regresi

Menurut Draper dan Smith (1998), analisis regresi adalah metode statistik yang bertujuan untuk memodelkan hubungan antara dua atau lebih variabel. Dalam analisis tersebut, variabel dependen berperan sebagai variabel respons, sedangkan variabel independen berperan sebagai variabel penjelas atau prediktor. Tujuan utama dari analisis regresi adalah untuk menemukan atau memprediksi perkiraan bentuk kurva regresi yang tepat.

Dalam proses pengestimasi kurva regresi, dikenal tiga pendekatan utama, yakni regresi parametrik, regresi nonparametrik, dan regresi semiparametrik. Menurut Wang (2011), persamaan regresi dapat dituliskan sebagai berikut:

$$Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_i X_i + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.1)$$

dengan:

- Y_i : Peubah respon ke- i ,
- β_0 : Intersep,
- β_i : Koefisien regresi ke- i ,
- X_i : Peubah prediktor ke- i ,
- ε_i : Galat ke- i yang diasumsikan saling bebas dengan nilai tengah nol dan keragaman konstan
- n : Ukuran pengamatan.

Regresi parametrik diterapkan ketika model yang diasumsikan bersifat parametrik. Metode ini sangat efisien dalam pendugaan jika asumsi modelnya akurat, namun jika asumsi tersebut tidak terpenuhi, hal ini dapat berakibat pada interpretasi data yang tidak tepat. Sebaliknya, apabila tidak terdapat pengetahuan awal mengenai bentuk fungsi, maka pendekatan nonparametrik lebih disarankan (Hardle, 1990).

2.3 Regresi Nonparametrik

Menurut Eubank (1999), regresi nonparametrik merupakan teknik statistik yang digunakan untuk menganalisis hubungan antara variabel dependen dan independen ketika bentuk kurva regresinya tidak diketahui atau pola hubungan tersebut tidak jelas. Pada regresi nonparametrik, kurva regresi hanya diasumsikan sebagai suatu fungsi yang halus, yaitu fungsi yang dapat merepresentasikan bentuk hubungan tertentu. Oleh karena itu, regresi nonparametrik bersifat sangat fleksibel dalam memodelkan data. Dalam statistika nonparametrik, bentuk sebaran data tidak dibatasi, sehingga metode ini dapat digunakan baik pada data yang berdistribusi normal maupun tidak.

Secara umum, model regresi nonparametrik dapat direpresentasikan dengan persamaan berikut: Secara umum, model regresi nonparametrik dapat direpresentasikan dengan persamaan berikut:

$$y_i = f(x_i) + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (2.2)$$

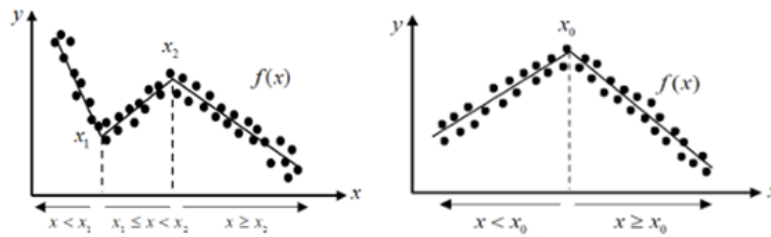
dengan:

- y_i : Variabel dependen pada data ke- i ,
- x_i : Variabel independen pada data ke- i ,
- $f(x_i)$: Fungsi regresi nonparametrik pada observasi ke- i yang bentuk fungsinya tidak diketahui,
- ε_i : Galat pada data ke- i yang diasumsikan saling independen, memiliki nilai rata-rata nol, dan variansi sebesar σ^2 .
- n : Ukuran pengamatan.

Regresi nonparametrik dipilih karena tidak menuntut asumsi bentuk kurva yang kaku, sehingga model yang terbentuk sepenuhnya ditentukan oleh data pengamatan (Hardle, 1990). Beberapa metode dalam regresi nonparametrik antara lain *spline*, polinomial lokal, deret *Fourier*, dan *wavelet*.

2.4 Truncated Spline

Menurut Budiantara (2019), metode *spline* merupakan bentuk potongan fungsi polinomial tersegmen (*truncated polynomial*) yang bersifat kontinu. Keunggulan metode ini terletak pada kemampuannya dalam menyesuaikan model terhadap data yang memiliki pola berubah-ubah di setiap sub-interval. Setiap segmen pada fungsi *spline* ditentukan oleh adanya titik knot, yaitu parameter pemulus yang berfungsi sebagai titik penghubung antarsegmen. Pada titik-titik knot tersebut biasanya terjadi perubahan pola atau perilaku data. Pola perubahan tersebut dapat bervariasi, tergantung pada karakteristik data yang diilustrasikan pada gambar berikut.



Gambar 1. Visualisasi Perubahan Pola Data

Apabila fungsi regresi f pada suatu persamaan didekati menggunakan fungsi *spline* berorde k dengan titik-titik knot K_1, K_2, \dots, K_r , maka bentuk umum dari fungsi *spline truncated* dapat dinyatakan sebagai berikut:

$$g(x_i) = \sum_{k=0}^m \beta_k x_i^k + \sum_{l=1}^r \beta_{(m+l)} (x_i - K_l)_+^m; \quad i = 1, 2, \dots, n \quad (2.3)$$

dengan

$$(x_i - K_l)_+^m = \begin{cases} (x_i - K_l)^m, & x_i \geq K_l \\ 0, & x_i < K_l \end{cases} \quad (2.4)$$

dimana:

- $g(x_i)$: Nilai estimasi fungsi *spline truncated* pada observasi ke- i ,
- x_i^k : Variabel independen ke- i , dengan orde k , untuk $k = 1, 2, \dots, m$,
- β_k : Parameter model polinomial untuk x_i^k ,
- $\beta_{(m+l)}$: Parameter pada komponen *truncated*, untuk $l = 1, 2, \dots, r$,
- r : Banyaknya titik *knot*.

Berdasarkan Persamaan (2.3), orde k merepresentasikan derajat dari polinomial, sedangkan K_l merupakan titik knot yang menandai adanya perubahan pola pada data. Dengan mensubstitusikan Persamaan (2.3) ke dalam persamaan (2.2), diperoleh bentuk model regresi nonparametrik *spline truncated* sebagai berikut:

$$y_i = \sum_{k=0}^m \beta_k x_i^k + \sum_{l=1}^r \beta_{(m+l)} (x_i - K_l)_+^m + \varepsilon_i \quad (2.5)$$

Misalkan terdapat variabel independen $x_{1i}, x_{2i}, \dots, x_{pi}$ maka regresi nonparametrik *spline truncated* multivariabel dapat dituliskan sebagai berikut (Budiantara, 2019):

$$y_i = \sum_{j=1}^p \sum_{k=0}^m \beta_{jk} x_{ji}^k + \sum_{j=1}^p \sum_{l=1}^r \beta_{j(m+l)} (x_{ji} - K_{jl})_+^m + \varepsilon_i \quad (2.6)$$

dengan:

$$(x_{ji} - K_{jl})_+^m = \begin{cases} (x_{ji} - K_{jl})^m, & x_{ji} \geq K_{jl} \\ 0, & x_{ji} < K_{jl} \end{cases} \quad (2.7)$$

dimana:

- β_{jk} : Parameter model polinomial, untuk $j = 1, \dots, p$,
dan $k = 0, 1, 2, \dots, m$,
- x_{ji} : Variabel independen j ke- i ,
- $\beta_{j(m+l)}$: Parameter model yang merepresentasikan komponen *truncated*,
untuk $l = 1, 2, \dots, r$,
- r : Banyaknya titik knot,
- K_{jl} : Titik-titik knot.

Untuk memperoleh hasil estimasi dari Persamaan (2.6), model tersebut dapat dilakukan transformasi bentuk, sehingga diperoleh Persamaan (2.8).

$$\hat{y} = \sum_{j=1}^p \sum_{k=0}^m \hat{\beta}_{jk} x_{ji}^k + \sum_{j=1}^p \sum_{l=1}^r \hat{\beta}_{j(m+l)} (x_{ji} - K_{jl})_+^m \quad (2.8)$$

dengan:

$$(x_{ji} - K_{jl})_+^m = \begin{cases} (x_{ji} - K_{jl})^m, & x_{ji} \geq K_{jl} \\ 0, & x_{ji} < K_{jl} \end{cases} \quad (2.9)$$

dimana:

- \hat{y} : Estimasi nilai y ,
- $\hat{\beta}_{jk}$: Estimasi parameter model β_{jk}
- x_{ji} : Variabel independen j ke- i
- $\hat{\beta}_{j(m+l)}$: Parameter estimasi model yang merepresentasikan komponen *truncated*, untuk $l = 1, 2, \dots, r$
- r : Banyaknya titik knot
- K_{jl} : Titik-titik knot

Persamaan (2.7) jika dijabarkan dalam bentuk matriks maka menjadi:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} 1 & x_1^1 & x_1^2 & \cdots & x_1^p & \vdots & (x_1 - K_1)_+^m & \cdots & (x_1 - K_r)_+^m \\ 1 & x_1^1 & x_2^2 & \cdots & x_2^p & \vdots & (x_2 - K_1)_+^m & \cdots & (x_2 - K_r)_+^m \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_1^1 & x_n^2 & \cdots & x_n^p & \vdots & (x_n - K_1)_+^m & \cdots & (x_n - K_r)_+^m \end{bmatrix} \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_{p+r} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dapat ditulis menjadi:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

dengan:

- \mathbf{y} : Vektor variabel dependen berukuran $n \times 1$
- \mathbf{X} : Matriks berukuran $n \times (m + r + 1)$
- $\boldsymbol{\beta}$: Vektor parameter yang akan diestimasi berukuran $(m + r + 1) \times 1$
- $\boldsymbol{\varepsilon}$: Vektor galat berukuran $n \times 1$

Parameter $\boldsymbol{\beta}$ ditaksir dengan menggunakan metode *Ordinary Least Squares* (OLS), di mana estimator $\boldsymbol{\beta}$ diperoleh melalui prosedur minimisasi pada metode OLS sebagai berikut:

$$\min\{Q(\boldsymbol{\beta})\} = \min\{(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta})\} \quad (2.10)$$

Jika persamaan (2.9) diuraikan, maka diperoleh bentuk sebagai berikut:

$$\begin{aligned} Q(\boldsymbol{\beta}) &= (\mathbf{y} - \mathbf{X}\boldsymbol{\beta})^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= (\mathbf{y} - \mathbf{X}^T\boldsymbol{\beta}^T)(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \\ &= \mathbf{y}^T\mathbf{y} - 2\boldsymbol{\beta}^T\mathbf{X}^T\mathbf{y} + \boldsymbol{\beta}^T\mathbf{X}^T\mathbf{X}\boldsymbol{\beta} \end{aligned} \quad (2.11)$$

Estimasi parameter diperoleh dengan menghitung turunan parsial $Q(\beta)$ terhadap β .

$$\frac{\partial Q(\beta)}{\partial \beta} = -2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \beta \quad (2.12)$$

Dengan menyetarakan hasil turunan pada Persamaan (2.11) dengan nol, diperoleh bentuk sebagai berikut:

$$-2\mathbf{X}^T \mathbf{y} + 2\mathbf{X}^T \mathbf{X} \beta = \mathbf{0} \quad (2.13)$$

$$\mathbf{X}^T \mathbf{X} \beta = \mathbf{X}^T \mathbf{y} \quad (2.14)$$

Dengan demikian, estimator β dapat dinyatakan sebagai berikut:

$$\hat{\beta} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y} \quad (2.15)$$

Model regresi nonparametrik *truncated spline* dengan derajat p dengan ($p = 1, 2, 3$) pada titik-titik knot K_1, K_2, K_3 membentuk model sebagai berikut:

a. Model *truncated spline* linear dengan satu titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ \quad (2.16)$$

b. Model *truncated spline* linear dengan dua titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ + \hat{\beta}_3 (x_1 - K_2)_+ \quad (2.17)$$

c. Model *truncated spline* linear dengan tiga titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 (x_1 - K_1)_+ + \hat{\beta}_3 (x_1 - K_2)_+ + \hat{\beta}_4 (x_1 - K_3)_+ \quad (2.18)$$

d. Model *truncated spline* kuadratik dengan satu titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_1^2 + \hat{\beta}_3 (x_1 - K_1)_+^2 \quad (2.18)$$

e. Model *truncated spline* kuadratik dengan dua titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_1^2 + \hat{\beta}_3 (x_1 - K_1)_+^2 + \hat{\beta}_4 (x_1 - K_2)_+^2 \quad (2.20)$$

f. Model *truncated spline* kuadratik dengan tiga titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_1^2 + \hat{\beta}_3 (x_1 - K_1)_+^2 + \hat{\beta}_4 (x_1 - K_2)_+^2 + \hat{\beta}_5 (x_1 - K_3)_+^2 \quad (2.21)$$

g. Model *truncated spline* kubik dengan satu titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_1^2 + \hat{\beta}_3 x_1^3 + \hat{\beta}_4 (x_1 - K_1)_+^3 \quad (2.22)$$

h. Model *truncated spline* kubik dengan dua titik knot:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_1^2 + \hat{\beta}_3 x_1^3 + \hat{\beta}_4 (x_1 - K_1)_+^3 + \hat{\beta}_5 (x_1 - K_2)_+^3 \quad (2.23)$$

i. Model *truncated spline* kubik dengan tiga titik knot:

$$\begin{aligned} \hat{y} = & \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_1^2 + \hat{\beta}_3 x_1^3 + \hat{\beta}_4 (x_1 - K_1)_+^3 + \\ & + \hat{\beta}_5 (x_1 - K_2)_+^3 + \hat{\beta}_6 (x_1 - K_3)_+^3 \end{aligned} \quad (2.24)$$

2.4.1 Menentukan Titik Knot yang Optimal

Titik knot adalah titik di mana terjadi perubahan karakteristik pola pada interval yang berbeda. Ketepatan dalam menentukan titik-titik knot yang optimal sangat memengaruhi kualitas penaksir pada regresi nonparametrik *spline truncated*. Oleh karena itu, penentuan lokasi titik knot yang tepat menjadi aspek krusial dalam penerapan regresi nonparametrik *spline truncated*. Pemilihan titik knot K_1, K_2, \dots, K_r yang optimal dapat dilakukan melalui beberapa pendekatan, di antaranya metode *Cross Validation* (CV) dan *Generalized Cross Validation* (GCV). Penentuan titik knot optimal berdasarkan kriteria GCV dilakukan dengan memilih nilai GCV yang paling minimum. Adapun persamaan GCV menurut Lai & Wang (2020) adalah sebagai berikut:

$$GCV(K_1, K_2, \dots, K_r) = \frac{MSE(K_1, K_2, \dots, K_r)}{\left[\frac{1}{n} \text{trace}(\mathbf{I} - \mathbf{A}(K_1, K_2, \dots, K_r))\right]^2} \quad (2.25)$$

dimana:

$$MSE(K_1, K_2, \dots, K_r) = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.26)$$

dengan

\mathbf{I}	:	Matriks identitas,
K_1, K_2, \dots, K_r	:	Titik-titik knot,
n	:	Total pengamatan,
y_i	:	Nilai aktual ke- i ,
\hat{y}_i	:	Nilai prediksi ke- i .

Matriks \mathbf{A} dinyatakan dengan persamaan:

$$\mathbf{A}(K_1, K_2, \dots, K_r) = \mathbf{X}(\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \quad (2.27)$$

Berdasarkan komponen-komponen matriks dan kriteria kebaikan model di atas, proses untuk mendapatkan nilai GCV yang paling minimum dilakukan secara sistematis. Dalam penelitian ini, pemilihan lokasi titik knot (K_1, K_2, \dots, K_r) yang meminimumkan nilai GCV tersebut dilakukan menggunakan pendekatan kuantil yang dikombinasikan dengan mekanisme *grid search*, dengan tahapan sebagai berikut:

1. Menentukan Kandidat Titik Knot

Menggunakan pendekatan kuantil dari sebaran data variabel independen. Hal ini bertujuan agar kandidat titik knot yang diuji tersebar secara representatif mengikuti distribusi data aktual.

2. *Grid Search*

Proses *grid search* dilakukan dengan menguji seluruh kombinasi kandidat titik *knot* yang telah ditentukan melalui iterasi komputasi.

3. Pemilihan Model Terbaik

Nilai GCV paling minimum pada Persamaan (2.25) menjadi dasar pemilihan kombinasi titik knot optimal di antara seluruh kombinasi yang telah diuji.

2.5 *B-Spline*

Basis fungsi *B-Spline* dipandang sebagai alternatif dari basis fungsi *truncated*, yang memiliki keterbatasan pada orde tinggi, jumlah knot yang banyak, maupun

penempatan knot yang terlalu berdekatan. Kondisi tersebut dapat menyulitkan penyelesaian persamaan, karena matriks yang terbentuk berpotensi mendekati singularitas (Budiantara, 2005). Dalam konstruksi model regresi *spline* dengan basis fungsi *B-Spline*, terdapat tiga kriteria utama yang harus diperhatikan, yaitu penentuan orde model, jumlah knot yang digunakan, serta lokasi penempatan knot (Rahmawati, dkk., 2017).

Model regresi nonparametrik yang didekati dengan fungsi *B-Spline* berorde m dan memiliki k knot dapat diformulasikan dalam bentuk persamaan sebagai berikut (Ariesta, 2021):

$$h(x_i) = \sum_{J=1}^{m+k} \beta_J B_{J-m,m}(x_i) \quad i = 1, 2, 3, \dots, n \quad (2.28)$$

dengan:

- $B_{J-m,m}(x_i)$: Basis fungsi *B-Spline* dengan orde m , dimana $m = 2, 3, 4$, pada titik knot ke- J
- β_J : Parameter model regresi *B-Spline* untuk basis fungsi ke- J .

Prosedur pembentukan fungsi *B-Spline* berorde m dengan penentuan letak titik knot $\xi_1, \xi_2, \dots, \xi_k$, dimana $a < \xi_1 < \xi_2 < \dots < \xi_k < b$, langkah pertama adalah mendefinisikan sebanyak $2m$ titik knot tambahan, yaitu: $\xi_{-(m-1)} < \dots < \xi_{-1} < \xi_0 < \dots < \xi_{(m+k)}$ dimana $\xi_{-(m-1)} = \dots = \xi_0 = a$ dan $\xi_{k+1} = \dots = \xi_{(m+k)} = b$. Nilai a ditetapkan sebagai nilai minimum dari x , sedangkan nilai b ditetapkan sebagai nilai maksimum dari x .

Mengacu pada Devi et al. (2014), dengan menggabungkan Persamaan (2.1) dan (2.12), diperoleh model regresi nonparametrik yang didekati oleh fungsi *B-Spline* berorde m dan memiliki k knot sebagai berikut:

$$h(x_i) = \sum_{J=1}^{m+k} \beta_J B_{J-m,m}(x_i) \quad i = 1, 2, 3, \dots, n \quad (2.29)$$

Menurut Ariesta et al. (2021), fungsi *B-Spline* dengan orde m pada knot ke- J dirumuskan secara rekursif dalam bentuk berikut:

$$B_{J,m}(x) = \frac{x - \xi_J}{\xi_{J+m-1} - \xi_J} B_{J,m-1}(x) + \frac{\xi_{J+m} - x}{\xi_{J+m} - \xi_{J+1}} B_{J+1,m-1}(x) \quad (2.30)$$

dengan $J = -(m-1), \dots, k$, dan

$$B_{J,1}(x) = \begin{cases} 1, & \text{jika } \xi_J < x \leq \xi_{J+1} \\ 0, & \text{untuk yang lainnya} \end{cases} \quad (2.31)$$

Orde atau derajat dari *B-Spline* dilambangkan dengan m . Berdasarkan nilai orde m , basis fungsi *B-Spline* dikelompokkan ke dalam tiga kategori utama (Rahmawati, dkk., 2017), yaitu:

1. Basis fungsi *B-Spline* berorde $m = 2$ merupakan fungsi linear yang dirumuskan sebagai berikut:

$$B_{J,2}(x) = \frac{x - \xi_J}{\xi_{J+1} - \xi_J} B_{J,1}(x) + \frac{\xi_{J+2} - x}{\xi_{J+2} - \xi_{J+1}} B_{J+1,1}(x) \quad \text{dengan } J = -1, \dots, k \quad (2.32)$$

2. Basis fungsi *B-Spline* berorde $m = 3$ merupakan fungsi kuadratik yang dirumuskan sebagai berikut:

$$B_{J,3}(x) = \frac{x - \xi_J}{\xi_{J+2} - \xi_J} B_{J,2}(x) + \frac{\xi_{J+3} - x}{\xi_{J+3} - \xi_{J+1}} B_{J+1,2}(x) \quad \text{dengan } J = -2, \dots, k \quad (2.33)$$

3. Basis fungsi *B-Spline* berorde $m = 4$ merupakan fungsi kubik yang dirumuskan sebagai berikut:

$$B_{J,4}(x) = \frac{x - \xi_J}{\xi_{J+3} - \xi_J} B_{J,3}(x) + \frac{\xi_{J+4} - x}{\xi_{J+4} - \xi_{J+1}} B_{J+1,3}(x) \quad \text{dengan } J = -3, \dots, k \quad (2.34)$$

2.5.1 Penaksiran Parameter dalam Model *B-Spline*

Pendekatan terhadap kurva regresi pada Persamaan (2.28) dapat dilakukan dengan menerapkan fungsi *B-Spline* berorde m , yang disusun berdasarkan k titik knot, yaitu $\lambda = \{\xi_1, \dots, \xi_k\}$, dan dapat dinyatakan sebagai berikut:

$$h_\lambda(x_i) = \sum_{J=1}^{m+k} \beta_{\lambda J} B_{\lambda J-m,m}(x_i) \quad (2.35)$$

Sehingga model *B-spline* adalah:

$$h_i = \sum_{J=1}^{m+k} \beta_{\lambda J} B_{\lambda J-m,m}(x_i) + \varepsilon_i \quad (2.36)$$

Model regresi pada Persamaan (2.36) dapat dinyatakan sebagai berikut:

$$h_i = \beta_{\lambda 1} B_{\lambda 1-m,m}(x_i) + \beta_{\lambda 2} B_{\lambda 2-m,m}(x_i) + \dots + \beta_{\lambda(m+k)} B_{\lambda k,m}(x_i) + \varepsilon_i \quad (2.37)$$

Jika dinyatakan dalam bentuk matriks, maka diperoleh:

$$\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_n \end{bmatrix} = \begin{bmatrix} B_{\lambda 1-m,m}(x_1) & B_{\lambda 2-m,m}(x_1) & \dots & B_{\lambda k,m}(x_1) \\ B_{\lambda 1-m,m}(x_2) & B_{\lambda 2-m,m}(x_2) & \dots & B_{\lambda k,m}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ B_{\lambda 1-m,m}(x_n) & B_{\lambda 2-m,m}(x_n) & \dots & B_{\lambda k,m}(x_n) \end{bmatrix} \begin{bmatrix} \beta_{\lambda 1} \\ \beta_{\lambda 2} \\ \vdots \\ \beta_{\lambda(m+k)} \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

dapat dinyatakan sebagai berikut:

$$\mathbf{h} = \mathbf{B}_\lambda \boldsymbol{\beta}_\lambda + \boldsymbol{\varepsilon} \quad (2.38)$$

Parameter $\boldsymbol{\beta}_\lambda$ diestimasi melalui pendekatan *least squares spline*. Estimator $\hat{\boldsymbol{\beta}}_\lambda$ diperoleh dengan meminimalkan nilai jumlahan kuadrat *error*, sehingga diperoleh bentuk berikut:

$$\hat{\boldsymbol{\beta}}_\lambda = (\mathbf{B}_\lambda^T \mathbf{B}_\lambda)^{-1} \mathbf{B}_\lambda^T \mathbf{h} \quad (2.39)$$

dengan $\hat{\boldsymbol{\beta}}_\lambda = (\hat{\beta}_{\lambda 1}, \hat{\beta}_{\lambda 2}, \dots, \hat{\beta}_{\lambda(m+k)})^T$.

Model regresi nonparametrik yang diestimasi menggunakan fungsi *B-Spline* dinyatakan sebagai berikut:

$$\begin{aligned}
 \hat{\mathbf{h}} &= \mathbf{B}_\lambda \hat{\boldsymbol{\beta}}_\lambda \\
 &= \mathbf{B}_\lambda ((\mathbf{B}_\lambda^T \mathbf{B}_\lambda)^{-1} \mathbf{B}_\lambda^T \mathbf{h}) \\
 &= \mathbf{B}_\lambda (\mathbf{B}_\lambda^T \mathbf{B}_\lambda)^{-1} \mathbf{B}_\lambda^T \mathbf{h} \\
 &= \mathbf{S}_\lambda \mathbf{h}
 \end{aligned} \tag{2.40}$$

dengan matriks $\mathbf{S}_\lambda = \mathbf{B}_\lambda (\mathbf{B}_\lambda^T \mathbf{B}_\lambda)^{-1} \mathbf{B}_\lambda^T$ simetris dan definit positif (Eubank, 1999). Estimasi model *B-Spline* juga dapat dinyatakan sebagai berikut:

$$\hat{h} = \sum_{J=1}^{m+k} \hat{\beta}_{\lambda J} B_{\lambda J-m,m}(x) \tag{2.41}$$

Persamaan (2.41) dapat dinyatakan sebagai berikut:

$$\hat{h} = \hat{\beta}_{\lambda 1} B_{\lambda 1-m,m}(x) + \hat{\beta}_{\lambda 2} B_{\lambda 2-m,m}(x) + \cdots + \hat{\beta}_{\lambda(m+k)} B_{\lambda k,m}(x) \tag{2.42}$$

2.5.2 Penentuan Model *B-Spline* yang Optimal

Perolehan model *B-Spline* yang terbaik mensyaratkan penentuan jumlah serta posisi titik knot secara optimal. Salah satu pendekatan yang dapat diterapkan dalam proses tersebut adalah metode *Generalized Cross Validation* (GCV). Menurut Wood (2017), nilai GCV dihitung menggunakan persamaan sebagai berikut:

$$GCV(\lambda) = \frac{MSE(\lambda)}{\left[\frac{1}{n} \text{trace}(\mathbf{I} - \mathbf{S}_\lambda) \right]^2}$$

Dengan n menyatakan jumlah pengamatan dan \mathbf{I} merupakan matriks identitas berukuran $n \times n$, $MSE(\lambda) = n^{-1} \sum_{i=1}^n (h_i - \hat{f}_\lambda)^2$ dan $\mathbf{S}_\lambda = \mathbf{B}_\lambda (\mathbf{B}_\lambda^T \mathbf{B}_\lambda)^{-1} \mathbf{B}_\lambda^T$. Penentuan model terbaik dilakukan dengan membandingkan nilai GCV pada setiap orde dan titik knot, kemudian dipilih model dengan nilai GCV minimum.

2.6 Evaluasi Model

Evaluasi model dilakukan untuk mengetahui seberapa baik model regresi *spline* yang dibangun dalam menggambarkan hubungan antara variabel independen dan variabel dependen. Beberapa ukuran yang umum digunakan untuk menilai ketepatan model adalah *Mean Squared Error* (MSE) dan Koefisien Determinasi (R^2).

2.6.1 Mean Squared Error

Mean Squared Error (MSE) merupakan salah satu kriteria yang paling sering digunakan untuk mengukur kinerja dan tingkat akurasi dari suatu model regresi. Secara historis, konsep pengukuran akurasi berbasis kuadrat galat ini berakar dari metode kuadrat terkecil (*least squares*) yang dikembangkan pertama kali oleh Adrien-Marie Legendre pada tahun 1805 dan Carl Friedrich Gauss pada tahun 1809 untuk meminimalkan jumlah kuadrat dari residual (Stigler, 1986). Nilai MSE diperoleh dari rata-rata penjumlahan kuadrat selisih antara nilai pengamatan (aktual) dengan nilai prediksi (*fitted value*). Secara matematis, rumus MSE didefinisikan sebagai berikut.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

dengan:

- n : Total pengamatan,
- y_i : Nilai aktual ke- i ,
- \hat{y}_i : Nilai prediksi ke- i .

2.6.1 Koefisien Determinasi

Menurut Ghozali (2021), koefisien determinasi (R^2) adalah indikator yang digunakan untuk menilai besarnya kontribusi variabel independen dalam menjelaskan variasi variabel dependen di dalam model. Nilai R^2 yang lebih tinggi menunjukkan bahwa model memiliki daya jelas yang semakin kuat terhadap data. Rumus koefisien

determinasi (R^2) dapat dinyatakan sebagai berikut:

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

dengan:

- n : Total pengamatan,
- y_i : Nilai aktual ke- i ,
- \hat{y}_i : Nilai prediksi ke- i ,
- \bar{y} : Rata-rata (*mean*) dari seluruh nilai aktual y .

2.7 Tingkat Pengangguran Terbuka (TPT)

Tingkat Pengangguran Terbuka (TPT) merupakan ukuran yang menunjukkan proporsi angkatan kerja yang belum bekerja dan masih aktif mencari pekerjaan terhadap keseluruhan angkatan kerja. TPT menggambarkan efektivitas suatu daerah dalam menyediakan lapangan pekerjaan. Menurut Sukirno (2019), tingkat pengangguran yang tinggi menunjukkan tidak seimbangnya antara pertumbuhan angkatan kerja dan kesempatan kerja yang tersedia. Dengan demikian, TPT menjadi ukuran penting dalam menilai keberhasilan pembangunan ekonomi suatu wilayah.

2.7.1 Kepadatan Penduduk (KP)

Kepadatan penduduk merupakan indikator kependudukan yang menunjukkan rata-rata jumlah penduduk yang menempati setiap kilometer persegi luas wilayah. Menurut BPS (2023), kepadatan penduduk mencerminkan tekanan jumlah penduduk terhadap ruang dan ketersediaan sumber daya di suatu daerah. Kepadatan penduduk yang tinggi mengindikasikan besarnya penawaran tenaga kerja (*labor supply*) yang apabila tidak diimbangi dengan ketersediaan lapangan kerja yang memadai, dapat memperketat persaingan kerja dan memicu peningkatan angka pengangguran (Todaro & Smith, 2015).

2.7.2 Indeks Pembangunan Manusia (IPM)

Indeks Pembangunan Manusia (IPM) adalah indikator komposit yang mengukur pencapaian rata-rata suatu wilayah dalam tiga dimensi dasar pembangunan manusia, yaitu umur panjang dan hidup sehat, pengetahuan, serta standar hidup layak (Badan Pusat Statistika, 2023). IPM mencerminkan kualitas modal manusia (*human capital*), di mana individu dengan tingkat kesehatan dan pendidikan yang lebih baik memiliki produktivitas serta daya saing yang lebih tinggi di pasar kerja. Secara teoritis, peningkatan kualitas pembangunan manusia diharapkan dapat memperluas peluang kerja dan menurunkan tingkat pengangguran terbuka (Mankiw, 2018).

2.7.3 Angkatan Kerja Laki-laki dan Perempuan

Angkatan kerja adalah penduduk usia kerja (15 tahun ke atas) yang bekerja atau sedang aktif mencari pekerjaan (Badan Pusat Statistika, 2023). Pemisahan angkatan kerja berdasarkan gender menjadi penting untuk melihat dinamika partisipasi tenaga kerja laki-laki dan perempuan yang seringkali memiliki karakteristik berbeda akibat struktur sosial dan ekonomi. Menurut Todaro & Smith (2015), meningkatnya jumlah angkatan kerja dari kedua kelompok gender tersebut dapat menambah tekanan pada pasar kerja apabila laju pertumbuhan kesempatan kerja tidak sebanding dengan penambahan jumlah penduduk yang siap masuk ke dunia kerja.

2.7.4 Pencari Kerja

Pencari kerja merupakan bagian dari angkatan kerja yang tidak memiliki pekerjaan dan sedang aktif melakukan upaya untuk mencari kerja, baik melalui pendaftaran di instansi terkait maupun jalur mandiri. Jumlah pencari kerja mencerminkan besarnya tekanan suplai tenaga kerja yang belum terserap oleh pasar kerja secara riil (Badan Pusat Statistika, 2023). Peningkatan jumlah pencari kerja yang signifikan tanpa diikuti oleh pembukaan lowongan pekerjaan yang setara akan secara langsung meningkatkan angka pengangguran terbuka di suatu wilayah.

2.7.5 Harapan Lama Sekolah (HLS)

Harapan Lama Sekolah (HLS) didefinisikan sebagai lamanya sekolah (dalam tahun) yang diharapkan akan dirasakan oleh anak pada umur tertentu di masa mendatang (Badan Pusat Statistika, 2023). HLS merupakan indikator investasi pendidikan jangka panjang yang berkontribusi pada peningkatan keterampilan dan kompetensi tenaga kerja. Secara teoritis, durasi pendidikan yang lebih lama meningkatkan keahlian individu sehingga memiliki daya tawar yang lebih baik dan menurunkan risiko pengangguran, meskipun keseimbangan antara kualifikasi pendidikan dengan kebutuhan industri tetap menjadi faktor penentu (Mankiw, 2018).

III. METODE PENELITIAN

3.1 Tempat dan Waktu Penelitian

Pelaksanaan penelitian ini dilakukan pada semester ganjil tahun ajaran 2025/2026 dan berlokasi di Jurusan Matematika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Lampung.

3.2 Data Penelitian

Penelitian ini menggunakan data sekunder berupa data tingkat pengangguran terbuka, kepadatan penduduk, indeks pembangunan manusia, Angkatan kerja laki-laki, Angkatan kerja Perempuan, pencari kerja, dan harapan lama sekolah di wilayah Sumbagsel tahun 2024 yang bersumber dari Badan Pusat Statistik (BPS), melalui situs resmi <https://www.bps.go.id/id>

Tabel 1. Variabel Penelitian

Variabel		Indikator
Dependen	Y	Tingkat Pengangguran Terbuka
Independen	X_1	Kepadatan Penduduk (KP)
	X_2	Indeks Pembangunan Manusia (IPM)
	X_3	Angkatan Kerja Laki-laki (AKL)
	X_4	Angkatan Kerja Perempuan (AKP)

Tabel 1. (Lanjutan)

Variabel		Indikator
Independen	X_5	Pencari Kerja (PK)
	X_6	Harapan Lama Sekolah (HLS)

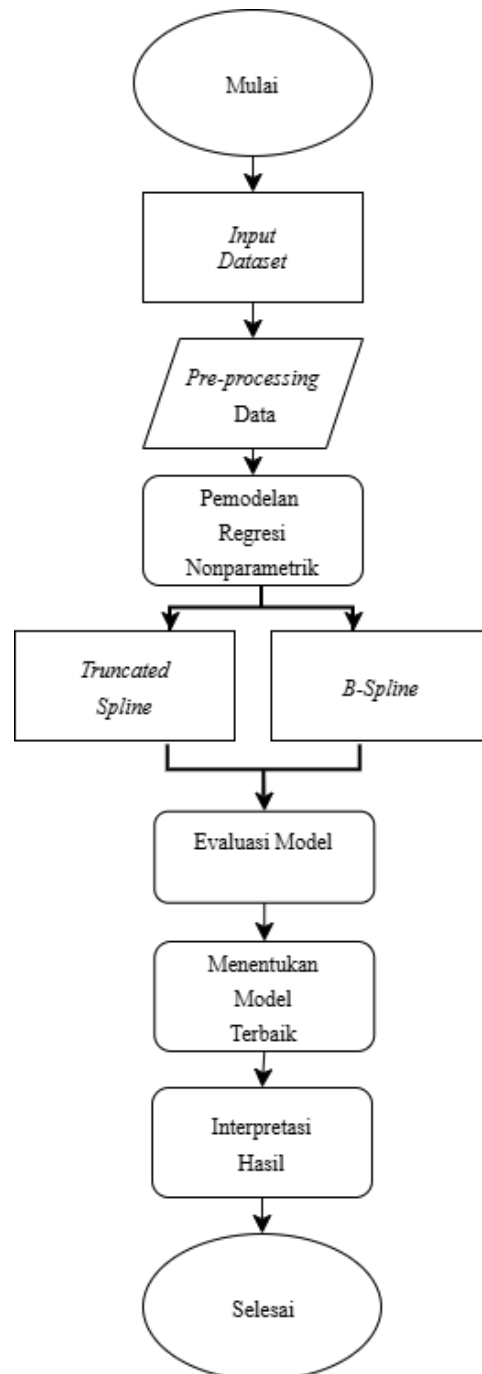
3.3 Metode Penelitian

Penelitian ini menggunakan metode regresi nonparametrik *spline* dengan data sekunder diperoleh dari Badan Pusat Statistik (BPS) dan diolah menggunakan RStudio dan *Microsoft Excel*.

Proses analisis yang dilakukan dalam mengkaji studi kasus ini untuk menghasilkan estimasi parameter adalah sebagai berikut:

1. Input *dataset* ke dalam perangkat lunak RStudio.
2. Melakukan analisis deskriptif terhadap data tingkat pengangguran terbuka, KP, IPM, AKL, AKP, PK, dan HLS.
3. Melakukan standarisasi data menggunakan metode *Z-score*.
4. Membentuk model regresi nonparametrik *truncated spline* dengan 3 orde dan 3 titik knot.
5. Membentuk model regresi nonparametrik *B-spline* dengan 3 orde dan 3 titik knot.
6. Memilih model *truncated spline* dan *B-spline* yang optimal berdasarkan nilai GCV terkecil.
7. Membandingkan hasil pemodelan *truncated spline* dan *B-spline* untuk menentukan model terbaik dengan menggunakan kriteria evaluasi model *Mean Square Error* (MSE), dan koefisien determinasi (R^2).
8. Melakukan interpretasi hasil, yaitu melihat bagaimana pola hubungan variabel independen terhadap tingkat pengangguran terbuka serta menjelaskan metode mana yang lebih sesuai digunakan.

Alur penelitian dalam studi kasus ini ditampilkan pada Gambar 2.



Gambar 2. Diagram Alur Penelitian

V. KESIMPULAN

Berdasarkan hasil penelitian perbandingan metode *truncated spline* dan *B-spline* dalam pemodelan tingkat pengangguran terbuka di wilayah Sumbagsel dan pembahasan pada bab IV, maka berikut merupakan kesimpulan yang dapat diambil:

1. Penelitian ini telah berhasil membentuk model regresi nonparametrik *truncated spline* dan *B-spline* untuk menganalisis faktor-faktor yang memengaruhi TPT. Model ini melibatkan enam variabel independen melalui simulasi berbagai orde dan variasi titik knot untuk mendapatkan estimasi terbaik bagi data yang memiliki pola fluktuatif.
2. Berdasarkan kriteria kebaikan model, model *Truncated Spline* linear satu titik knot secara konsisten memberikan hasil yang lebih unggul dibandingkan *B-spline*. Model ini memiliki nilai GCV terkecil yaitu sebesar 0,5121, MSE terendah sebesar 0,3270, dan koefisien determinasi (R^2) tertinggi yaitu sebesar 0,6674 atau 66,74%.
3. Model terbaik yang paling tepat untuk menggambarkan hubungan variabel terhadap TPT di wilayah Sumbagsel adalah model *truncated spline* linear satu titik knot dengan persamaan sebagai berikut: Model terbaik yang paling tepat untuk menggambarkan hubungan variabel terhadap TPT di wilayah Sumbagsel adalah model *truncated spline* linear satu titik knot dengan persamaan sebagai berikut:

$$\hat{y}_i = 1,2565 + 4,2596x_{1i} - 0,6975x_{2i} - 0,6376x_{3i} + 4,4404x_{4i} + 0,0033x_{5i} + 0,0245x_{6i} + 4,1814(x_{1i} + 0,3169)_+ + 1,0995(x_{2i} + 0,7232)_+ + 0,4997(x_{3i} + 0,1300)_+ + 0,8102(x_{4i} + 0,8836)_+ - 0,1270(x_{5i} - 1,2684)_+ - 0,4261(x_{6i} - 0,2657)_+$$

DAFTAR PUSTAKA

- Anggreni, N. P. R., Suciptawati, N. L. P., & Srinadi, I. G. A. M. 2018. Model regresi nonparametrik spline truncated pada jumlah kasus tuberkulosis di Provinsi Bali tahun 2016. *E-Jurnal Matematika*. **7**(3):211-218.
- Ariesta, D., Gusriani, N., & Parmikanti, K. 2021. Estimasi Parameter Model Regresi Nonparametrik B-Spline Pada Angka Kematian Maternal. *Jurnal Matematika UNAND*, **10**(3): 342–354.
- Badan Pusat Statistik (BPS). 2023. Rata-Rata Lama Sekolah Penduduk Indonesia Tahun 2023. Jakarta.
- Budiantara, I. N. 2005. *Regresi Nonparametrik dan Spline: Teori dan Aplikasinya*. Jurusan Statistika, Institut Teknologi Sepuluh Nopember.
- Budiantara, I. N., Suryadi, F., Otok, B. W., & Guritno, S. 2006. Pemodelan B Spline dan Mars Pada Nilai Ujian Masuk Terhadap IPK Mahasiswa Jurusan Disain Komunikasi Visual UK. Petra Surabaya. *Jurnal Teknik Industri*.**8**(1):1-13.
- Budiantara, I. N. 2019. *Regresi Nonparametrik Spline Truncated*. ITS Press. Surabaya.
- Devi, A. R., Mukid, M. A., & Yasin, H. 2014. Analisis Inflasi Kota Semarang Menggunakan Metode Regresi Non Parametrik B-Spline. *Jurnal Gaussian*. **3**(2):193–202.
- Draper, N.R. & Smith, H. 1998. *Applied Regression Analysis*. 3th Edition. John Wiley and Sons, New York.

- Eubank, R.L. 1999. *Nonparametric Regression and Spline Smoothing*. 2nd Edition. Marcel Dekker, New York.
- Ghozali, I. 2021. *Aplikasi Analisis Multivariate dengan Program IBM SPSS 26*. Badan Penerbit Universitas Diponegoro.
- Hasan, M. I. 2001. *Pokok pokok Materi Statistika 1 (Statistika Deskriptif)*. Jakarta: Bumi Aksara
- Handajani, S. S., Pratiwi, H., Respatiwulan, Qona'ah, N., Ramadhania, M., Evitasari, N., & Apsari, N. E. 2023. Comparison of b-spline and truncated spline regression models for temperature forecast. *BAREKENG: Journal of Mathematics and Its Applications*. **17**(4):1969–1984.
- Hardle, W. 1990. *Applied Nonparametric Regression*. Cambridge University Press.
- Kementerian Ketenagakerjaan Republik Indonesia. 2022. *Profil Ketenagakerjaan Indonesia 2022*. Jakarta: Kemnaker RI
- Lai, H., & Wang, Z. 2020. Automatic knot selection in spline regression using generalized cross-validation. *Journal of Statistical Computation and Simulation*. **90**(12):2352-2368.
- Lyche, T., & Mørken, K. 2008. *Spline Methods*. In G. E. Fasshauer & L. L. Schumaker (Eds.), *Approximation Theory XIII: San Antonio 2007* (pp. 1–20). Springer.
- Mankiw, N. G. 2018. *Principles of Macroeconomics*. 8th Edition. Cengage Learning.
- Rahmawati, D., Sulianto, A., & Toharudin, T. 2017. Pemodelan regresi nonparametrik spline untuk data time series produksi padi di Jawa Barat. *Jurnal Statistika Universitas Muhammadiyah Semarang*. **5**(2):107–114.

Rahmawati, A. S., Ispriyanti, D., & Warsito, B. 2017. Pemodelan Kasus Kemiskinan Di Jawa Tengah Menggunakan Regresi Nonparametrik Metode B-Spline. *Jurnal Gaussian*. 6(1): 11–20.

Sudjana, S. 1996. *Metoda Statistika*. 6th edition. Bandung: Tarsito.

Sukirno, S. 2019. *Makroekonomi Teori Pengantar*. Jakarta: Rajawali Pers.

Stigler, S. M. 1986. *The history of statistics: The measurement of uncertainty before 1900*. Harvard University Press.

Suwarno, S. S. 2023. Regresi Nonparametrik B-Spline untuk Memodelkan Inflasi di Indonesia (Skripsi). Jurusan Matematika Fakultas Sains dan Teknologi UIN Maulana Malik Ibrahim, Malang.

Todaro, M. P., & Smith, S. C. 2015. *Economic Development*. 12th edition. Pearson Education.

Wang, Y. 2011. *Smoothing Splines: Methods and Applications*. CRC Press.

Wood, S. N. 2017. *Generalized Additive Models: An Introduction with R*. 2nd edition. CRC Press.