



**ESTIMASI REGRESI NONPARAMETRIK
PENALIZED SPLINE FILTER MENGGUNAKAN
METODE *PENALIZED LEAST SQUARE* PADA DATA
*TIME SERIES***

Skripsi
disusun sebagai salah satu syarat
untuk memperoleh gelar Sarjana Sains
Program Studi Matematika

oleh
Desca Nur Alizah
4111413015

**JURUSAN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS NEGERI SEMARANG
2017**



UNNES
UNIVERSITAS NEGERI SEMARANG

PERNYATAAN KEASLIAN TULISAN

Saya menyatakan bahwa skripsi ini bebas plagiat, kecuali yang secara tertulis dirujuk dalam skripsi ini dan disebutkan dalam daftar pustaka. Apabila dikemudian hari terbukti terdapat plagiat dalam skripsi ini, maka saya bersedia menerima sanksi sesuai ketentuan perundang-undangan.

Semarang,



Desca Nur Alizah

UNNES
UNIVERSITAS NEGERI SEMARANG

PENGESAHAN

Skripsi yang berjudul

Estimasi Regresi Nonparametrik Penalized Spline Filter menggunakan
Penalized Least Square pada *Data Time series*

disusun oleh

Desca Nur Alizah

4111413015

telah dipertahankan di hadapan sidang Panitia Ujian Skripsi FMIPA UNNES
pada 20 Oktober tanggal 2017.

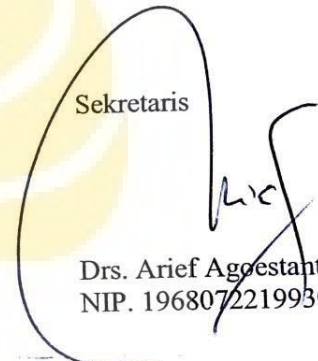
Panitia:

Ketua



Prof. Dr. Zaenuri, S.E., M.Si, Akt.
NIP. 196412231988031001

Sekretaris



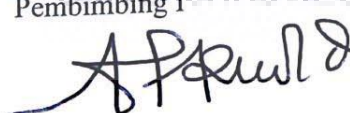
Drs. Arief Agoestanto, M.Si.
NIP. 196807221993031005

Ketua Penguji



Putriaji Hendikawati, S.Si., M.Pd., M.Sc.
NIP. 198208182006042001

Anggota Penguji/
Pembimbing I



Dr. Nur Karomah Dwidayati, M.Si.
NIP. 196605041990022001

Anggota Penguji/
Pembimbing II



Drs. Sugiman, M.Si.
NIP. 196401111989011001

MOTTO DAN PERSEMBAHAN

MOTTO

- ❖ Dan (ingatlah juga), tatkala Tuhanmu memaklumkan; "Sesungguhnya jika kamu bersyukur, pasti Kami akan menambah (nikmat) kepadamu, dan jika kamu mengingkari (nikmat-Ku), maka sesungguhnya azab-Ku sangat pedih". (QS. 'Ibrahim [14] : 7)
- ❖ Work hard in silence, let your succes be your noise. (Frank Ocean)
- ❖ Lillah maka tak lelah.

PERSEMBAHAN

- ❖ Untuk Ibu Mufriyah dan Bapak Machnurdin
- ❖ Untuk Ino dan Melan
- ❖ Untuk keluarga besar
- ❖ Untuk sahabat
- ❖ Untuk matematika Unnes
- ❖ Untuk Universitas Negeri Semarang

UNNES
UNIVERSITAS NEGERI SEMARANG

KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah memberikan nikmat dan karunia-Nya serta kemudahan sehingga penulis dapat menyelesaikan skripsi yang berjudul “Estimasi Regresi Nonparametrik Penalized Spline Filter menggunakan *Penalized Least Square* pada *Data Time series*”.

Penyusunan skripsi ini dapat diselesaikan berkat kerjasama, bantuan, dan dorongan dari berbagai pihak. Oleh karena itu penulis mengucapkan terima kasih kepada:

1. Prof. Dr. Fathur Rokhman, M.Hum., Rektor Universitas Negeri Semarang.
2. Prof. Dr. Zaenuri, S.E., M.Si., Akt., Dekan FMIPA Universitas Negeri Semarang.
3. Drs. Arief Agoestanto, M.Si., Ketua Jurusan Matematika FMIPA Universitas Negeri Semarang.
4. Drs. Mashuri, M.Si., Ketua Prodi Matematika FMIPA Universitas Negeri Semarang.
5. Dr. Nur Karomah Dwidayati, M.Si., selaku Dosen Pembimbing I yang telah memberikan bimbingan, pengarahan, nasehat, dan saran selama penyusunan skripsi ini.

6. Drs. Sugiman, M.Si., selaku Dosen Pembimbing II yang telah memberikan bimbingan, pengarahan, nasehat, dan saran selama penyusunan skripsi ini.
7. Putriaji Hendikawati, S.Si., M.Pd., M.Sc., selaku Dosen Penguji yang telah memberikan penilaian dan saran dalam perbaikan skripsi ini.
8. Dr. Rochmad, M.Si., selaku Dosen Wali yang telah memberikan bimbingan dan arahan.
9. Dosen-dosen Matematika Universitas Negeri Semarang yang telah membekali penulis dengan berbagai ilmu selama mengikuti perkuliahan sampai akhir penulisan skripsi ini.
10. Ibu Mufriyah dan Bapak Machnurdin tercinta, adik-adikku tersayang, Ino dan Melan serta seluruh keluarga yang senantiasa memberikan dukungan, semangat dan doa yang tiada putusya.
11. Bidikmisi Universitas Negeri Semarang yang telah memberikan dukungan secara materiil maupun non-materiil.
12. Ari, Lana, Rara sahabat yang tidak pernah bosan untuk selalu ada disegala suasana.
13. Teman-teman seperjuangan prodi Matematika FMIPA Unnes 2013 yang selalu memberikan semangat untuk bersama-sama berjuang dalam mendapat gelar S.Si ini.

14. Teman-teman HIMATIKA, MJC dan KKN Desa Ujung-Ujung yang telah memberikan banyak pengalaman organisasi, kemasyarakatan dan kekeluargaan.

15. Semua pihak yang tidak dapat disebutkan satu per satu yang telah memberikan bantuan dan semangat.

Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat banyak kekurangan. Oleh karena itu, penulis mengharapkan saran dan kritik yang membangun dari pembaca.

Semarang, Oktober 2017

Penulis



ABSTRAK

Alizah, Desca Nur. 2017. *Estimasi Regresi Nonparametrik Penalized Spline Filter Menggunakan Metode Penalized Least Square pada Data Time series.* Skripsi, Jurusan Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Negeri Semarang. Pembimbing Utama: Dr. Nur Karomah Dwidayati, M.Si. dan Pembimbing Pendamping: Drs. Sugiman, M.Si.

Kata Kunci: Regresi Nonparametrik, Spline, Penalized Spline Filter

Analisis regresi merupakan analisis statistika yang digunakan untuk menyelidiki pola hubungan fungsional antara variabel respon dan variabel prediktor. Pada umumnya untuk mengestimasi fungsi regresi digunakan pendekatan regresi parametrik. Dalam menerapkan regresi parametrik sering kali asumsi-asumsinya tidak terpenuhi. Solusi dari masalah tersebut adalah mengestimasi fungsi regresi menggunakan pendekatan regresi nonparametrik.

Penelitian ini difokuskan pada regresi nonparametrik dengan estimator Penalized Spline Filter (PSF). Penalized Spline Filter dengan keunggulannya memiliki fungsi penalti berupa fungsi filter dan sifat tersegmen kontinu yang memberikan fleksibilitas yang lebih baik sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik fungsi atau data yang diasumsikan erornya saling berkorelasi atau autokorelasi. Metode optimasi yang digunakan untuk estimasi parameter model PSF yaitu metode *Penalized Least Square* (PLS). Estimator parameter disimulasikan pada data nilai kurs rupiah terhadap dollar AS pada tanggal 1 November 2016 sampai 31 Januari 2017. Untuk mendapatkan model PSF terbaik, yang pertama ditentukan yaitu orde spline $p = 1$, selanjutnya dipilih banyaknya knot dengan algoritma *Full Search*. Knot optimal dan parameter penghalus λ optimal dipilih berdasarkan nilai GCV dan MSE terkecil yang dihitung menggunakan algoritma *Demmler-Reinsch Orthogonalization*.

Hasil penelitian didapati model PSF terbaik menggunakan banyak knot (K) = 1 dengan letak titik knot yaitu $K_1 = 13296$ dan parameter penghalus λ sebesar 100000 dengan nilai GCV terkecil sebesar 2544,012 dan MSE minimum sebesar 2385,051 dengan R^2 sebesar 84,3%. Hal ini menunjukkan bahwa model Penalized Spline Filter yang optimal dapat menjelaskan kurs rupiah terhadap dollar dalam setiap harinya pada tanggal 1 November 2016 sampai 31 Januari 2017, termasuk pola perubahan yang teridentifikasi dalam model tersebut.

DAFTAR ISI

Daftar Isi	Halaman
HALAMAN JUDUL.....	i
PERNYATAAN KEASLIAN TULISAN.....	Error! Bookmark not defined.
PENGESAHAN.....	Error! Bookmark not defined.
MOTTO DAN PERSEMBAHAN.....	v
KATA PENGANTAR.....	vi
ABSTRAK.....	ix
DAFTAR ISI.....	x
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN.....	xvi
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang.....	1
1.2 Rumusan Masalah.....	5
1.3 Batasan Masalah.....	5
1.4 Tujuan Penelitian.....	6
1.5 Manfaat Penelitian.....	6

1.6	Sistematika Penulisan.....	7
1.6.1	Bagian Awal	7
1.6.2	Bagian Isi	7
1.6.3	Bagian Akhir.....	8
BAB 2 TINJAUAN PUSTAKA.....		9
2.1	Landasan Teori.....	9
2.1.1	Regresi Parametrik.....	9
2.1.2	Regresi Nonparametrik	10
2.1.3	Pemodelan Data <i>Time series</i> dengan Regresi Nonparametrik.....	11
2.1.4	Pengujian Asumsi	13
2.1.5	Spline	16
2.1.6	Fungsi Spline Linier.....	18
2.1.7	Penalized Spline Filter.....	19
2.1.8	Pemilihan Parameter Penghalus λ Optimal	21
2.1.9	Pemilihan Knot Optimal	24
2.1.10	Kurs.....	26
2.2	Penelitian yang Relevan	28
2.3	Kerangka Berpikir	30

BAB 3 METODE PENELITIAN	35
3.1 Identifikasi Masalah	35
3.2 Fokus Penelitian	35
3.3 Jenis dan Sumber Data	36
3.4 Variabel Penelitian	36
3.5 Analisis dan Pemecahan Masalah	36
3.5.1 Estimasi regresi nonparametrik Penalized Spline Filter (PSF)....	36
3.5.2 Pemodelan data <i>time series</i> menggunakan model regresi nonparametrik Penalized Spline Filter.....	38
3.6 Penarikan Kesimpulan.....	39
BAB 4 HASIL DAN PEMBAHASAN	41
4.1 Hasil.....	41
4.1.1 Estimasi Regresi Nonparametrik Penalized Spline Filter (PSF) Menggunakan Metode Penalized Least Square (PLS).....	41
4.1.2 Pemodelan Data <i>Time series</i> Menggunakan Model Regresi Nonparametrik Penalized Spline Filter (PSF)	45
4.1.2.1 Deskripsi Data.....	45
4.1.3 Penerapan Program R 3.3.2 Pada Data Kurs Rupiah Terhadap Dollar Amerika Serikat	54

4.1.3.1Menentukan Titik Knot dan Parameter Penghalus λ Optimal.....	54
4.1.3.2	Model Penalized Spline Filter Optimal	56
4.1.3.3	Perbandingan Hasil Prediksi dengan Data Asli	59
4.2	Pembahasan.....	60
BAB 5 PENUTUP		63
5.1	Simpulan.....	63
5.2	Saran.....	64
DAFTAR PUSTAKA		65
LAMPIRAN		68



DAFTAR TABEL

Tabel	Halaman
Tabel 4.1. Statistik Deskriptif Data Kurs Tanggal 1 November 2016 sampai 31 Januari 2017.....	46
Tabel 4.2. Hasil <i>Output</i> Uji Normalitas Residual.....	48
Tabel 4.3. Hasil <i>Output</i> Uji Stasioneritas	49
Tabel 4.4. Hasil <i>Output</i> Uji Autokorelasi.....	52
Tabel 4.5. Banyak Knot, Parameter Penghalus λ Optimal, GCV dan MSE	55
Tabel 4.6 Koefisien parameter	56



DAFTAR GAMBAR

Gambar	Halaman
Gambar 2.1. Fungsi spline linier dengan 1 titik knot	19
Gambar 2.2. Kerangka Berpikir	34
Gambar 3.1. Diagram Alir Penelitian.....	40
Gambar 4.1. Grafik Data Kurs Beli Tanggal 1 November 2016 sampai 31 Januari 2017	45
Gambar 4.2. Plot Data Kurs Rupiah.....	46
Gambar 4.3. Plot PACF Data Kurs	47
Gambar 4.4. Hasil Output Correlogram Residual	51
Gambar 4.5. Kurva Estimasi Pola Hubungan Data Kurs ke $i - 1$ dan Data Kurs ke i	57
Gambar 4.6. Kurva Estimasi Kurs setelah dikembalikan terhadap Waktu (i)...	58
Gambar 4.7. Grafik Data Asli dan Hasil Prediksi	59

DAFTAR LAMPIRAN

Lampiran	Halaman
Lampiran 1. Data Kurs Rupiah terhadap Dolah Amerika Serikat.....	69
Lampiran 3. Hasil <i>Output</i> Uji Normalitas Residual	71
Lampiran 4. Hasil <i>Output</i> Uji Stationeritas	72
Lampiran 5. Hasil <i>Output</i> Uji Autokorelasi	73
Lampiran 6. Tabel <i>Durbin-Watson</i> (DW)	74
Lampiran 7. Program Pemilihan Banyak Knot Optimal Regresi Penalized Spline Filter Orde 1	76
Lampiran 8. Program Estimasi Regresi Penalized Spline Filter Orde 1	87
Lampiran 9. Tabel Data Asli dan Data Prediksi.....	92



BAB I

PENDAHULUAN

1.1 Latar Belakang

Analisis regresi merupakan analisis statistika yang digunakan untuk menyelidiki pola hubungan fungsional antara variabel respon dan variabel prediktor (Budiantara, 2009). Tujuan analisis regresi yaitu untuk mengetahui sejauh mana hubungan sebuah variabel bebas dengan beberapa variabel tak bebas (Pujilestari, 2017). Pada umumnya, untuk mengestimasi fungsi regresi, terdapat dua macam pendekatan yaitu pendekatan regresi parametrik dan regresi nonparametrik. Pendekatan parametrik yaitu pendekatan yang digunakan apabila bentuk hubungan antara variabel prediktor dan variabel respon diketahui atau bentuk dari kurva regresinya diasumsikan mengikuti pola tertentu. Pendekatan nonparametrik yaitu pendekatan yang digunakan apabila bentuk hubungan antara variabel respon dan prediktornya tidak diketahui atau tidak adanya informasi mengenai bentuk fungsi regresinya. Regresi parametrik memiliki asumsi yang kuat dan kaku seperti residual berdistribusi normal dan memiliki variansi yang konstan. Selain itu, diketahui karakteristik data dari penelitian sebelumnya sangatlah penting agar diperoleh model yang baik. Dalam model regresi parametrik, estimasi kurva regresi ekuivalen dengan estimasi terhadap parameter-parameter dalam model (Budiantara, 2009).

Dalam menerapkan regresi parametrik sering kali asumsi-asumsinya tidak terpenuhi. Solusi dari masalah tersebut adalah mengestimasi fungsi regresi menggunakan pendekatan regresi nonparametrik. Regresi nonparametrik adalah suatu metode pemodelan yang tidak terikat akan asumsi-asumsi dari persamaan regresi tertentu yang memberikan fleksibilitas yang tinggi dalam menduga sebuah model. Regresi nonparametrik digunakan jika bentuk kurva data tidak diketahui sebelumnya. Regresi nonparametrik memiliki fleksibilitas yang tinggi karena bentuk estimasi kurva regresinya dapat menyesuaikan datanya tanpa dipengaruhi oleh faktor subyektivitas peneliti.

Estimasi fungsi pada regresi nonparametrik dilakukan berdasarkan data pengamatan dengan beberapa teknik *smooth* tertentu antara lain dengan menggunakan histogram, estimator kernel, deret orthogonal, estimator spline, estimator *k-Nearest Neighborhood*, estimator deret fourier, dan estimator wavelet. Dalam penelitian ini akan difokuskan pada pendekatan nonparametrik dengan estimator spline. Laome (2010) telah meneliti tentang perbandingan regresi spline dengan kernel untuk mengestimasi model regresi nonparametrik. Berdasarkan kriteria kebaikan model yaitu R^2 , MSE dan uji asumsi residual, pendekatan regresi spline lebih baik dari regresi kernel untuk pertumbuhan balita. Hal ini dapat dilihat dari uji asumsi residual yang menunjukkan bahwa pendekatan spline memenuhi semua asumsi residual.

Regresi spline merupakan suatu pendekatan ke arah pencocokan data dengan tetap memperhitungkan kemulusan kurva (Eubank, 1999). Spline mempunyai keunggulan dalam mengatasi pola data yang menunjukkan naik atau turun yang tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus. Titik knot merupakan perpaduan bersama yang menunjukkan pola perilaku fungsi spline pada selang yang berbeda (Hardle, 1990). Dalam regresi spline, pemilihan banyak dan letak knot merupakan hal yang penting. Kemudian dipilih model optimal berdasarkan kriteria tertentu, misalnya nilai *Generalized Cross Validation* (GCV) minimum. Hal ini membutuhkan waktu yang lama dan jika dilakukan menggunakan *software* memerlukan memori yang besar. Karena itu diperlukan alternatif untuk mengatasi masalah ini, yaitu regresi penalized spline dimana knot terletak di titik-titik kuantil dari nilai *unique* (tunggal) variabel prediktor (Ruppert et al., 2003).

Regresi penalized spline cocok digunakan untuk pemodelan data *time series* yang berfluktuatif. Seperti yang sudah diteliti oleh Agustina (2015) bahwa data *time series* IHSG biasanya dapat dimodelkan dengan model parametrik. Namun dalam model parametrik terdapat asumsi yang harus dipenuhi, yaitu asumsi stasioneritas dan *white noise*. Sedangkan data IHSG yang fluktuatif biasanya tidak memenuhi asumsi tersebut. Oleh karena itu dilakukan pemodelan data IHSG menggunakan analisis yang tidak memerlukan asumsi-asumsi yang harus dipenuhi, salah satunya adalah regresi nonparametrik penalized spline.

Selain asumsi stasioneritas dan *white noise*, data *time series* dapat juga diasumsikan erornya saling berkorelasi. Dalam hal ini untuk diaplikasikan pada pemodelan regresi nonparametrik dibutuhkan analisis tersendiri, salah satu pendekatan yang dapat digunakan adalah pendekatan nonparametrik Penalized Spline Filter (Krivobokova et.al., 2007). Blochl (2014) dan Paige et.al. (2010) telah meneliti tentang model regresi spline *smoothing* PSF untuk data *time series*. Penalized Spline Filter merupakan salah satu bentuk penaksir regresi nonparametrik, dimana fungsi penaltinya menggunakan fungsi filter. Estimasi model regresi nonparametrik Penalized Spline Filter diperoleh dengan meminimumkan fungsi *Penalized Least Square* (Kauerman et.al., 2008). *Penalized Least Square* merupakan sebuah metode optimasi yang memberikan komponen penghalus pada metode *least square* dengan kriteria optimasi yang menggabungkan antara kecocokan terhadap data dan kemulusan kurva (JiangQing et.al, 2016).

Dari uraian di atas, maka dalam skripsi ini peneliti akan mengambil judul **“Estimasi Regresi Nonparametrik Penalized Spline Filter Menggunakan Metode *Penalized Least Square* Pada Data *Time series*”**. Data *time series* yang dimodelkan berupa data nilai kurs rupiah terhadap dollar Amerika Serikat terhitung mulai tanggal 1 November 2016 sampai dengan 31 Januari 2017. Untuk penentuan titik knot yang optimal dengan estimasi model regresi terbaik menggunakan bantuan *software* R 3.3.2.

1.2 Rumusan Masalah

Permasalahan yang dibahas dalam penelitian skripsi ini adalah

1. Bagaimana estimasi parameter model regresi nonparametrik Penalized Spline Filter menggunakan metode *Penalized Least Square*?
2. Bagaimana aplikasi model regresi nonparametrik Penalized Spline Filter untuk menganalisis data nilai kurs rupiah terhadap dollar Amerika Serikat tanggal 1 November 2016 sampai dengan 31 Januari 2017?

1.3 Batasan Masalah

Permasalahan yang dibahas dalam penelitian pada skripsi ini akan dibatasi lingkup pembahasannya, yaitu

1. Metode yang digunakan adalah Penalized Spline Filter dengan metode optimasi *Penalized Least Square*. Parameter penghalus λ yang optimal untuk setiap nilai knot diperoleh dengan kriteria nilai *Generalized Cross Validation* (GCV) dan *Mean Squared Error* (MSE) terkecil, sedangkan pemilihan banyak knot optimal dilakukan dengan metode *Full Search*.
2. Data yang digunakan adalah data nilai kurs rupiah terhadap dollar Amerika Serikat pada tanggal 1 November 2016 sampai dengan 31 Januari 2017.

1.4 Tujuan Penelitian

Tujuan dari penelitian pada skripsi ini adalah

1. Menentukan estimator parameter dari model regresi nonparametrik Penalized Spline Filter menggunakan metode *Penalized Least Square*.
2. Mengetahui penerapan regresi nonparametrik Penalized Spline Filter terbaik untuk menganalisis data nilai kurs rupiah terhadap dollar Amerika Serikat tanggal 1 November 2016 sampai dengan 31 Januari 2017.

1.5 Manfaat Penelitian

Manfaat yang diperoleh dari penelitian pada skripsi ini sebagai berikut

- a. Bagi Mahasiswa

Manfaat yang didapat dari penelitian ini sebagai pengembangan ilmu teoritis yang dipelajari di perkuliahan, penambahan wawasan tentang estimasi regresi nonparametrik Penalized Spline Filter menggunakan metode *Penalized Least Square* pada data *time series* yaitu mendapatkan model terbaik untuk menduga atau memprediksi nilai kurs harian rupiah terhadap dollar Amerika Serikat, dan referensi untuk mengembangkan estimasi regresi nonparametrik spline lainnya

pada penerapan kasus lain atau menggunakan pendekatan yang berbeda.

b. Bagi Perpustakaan Jurusan Matematika Unnes

Manfaat yang didapat dari penelitian ini adalah menambah referensi mengenai estimasi regresi nonparametrik Penalized Spline Filter dengan metode *Penalized Least Square* untuk pemodelan data *time series*.

1.6 Sistematika Penulisan

Adapun sistematika penulisan skripsi ini terdiri dari tiga bagian yaitu bagian awal, bagian isi, dan bagian akhir, yang masing-masing diuraikan sebagai berikut

1.6.1 Bagian Awal

Bagian ini terdiri dari halaman judul, halaman pengesahan, motto dan persembahan, kata pengantar, abstrak, daftar isi, daftar tabel, daftar gambar, dan daftar lampiran.

1.6.2 Bagian Isi

Bagian isi terdiri dari lima bab. Adapun bab tersebut sebagai berikut

Bab 1 Pendahuluan

Pada bab Pendahuluan berisi tentang latar belakang masalah, rumusan masalah, pembatasan masalah, tujuan, manfaat, dan sistematika penulisan.

Bab 2 Tinjauan Pustaka

Pada bab Tinjauan Pustaka berisi tentang landasan teori seperti regresi parametrik, regresi nonparametrik, pemodelan data *time series* dengan regresi nonparametrik, pengujian asumsi, spline, fungsi spline linier, penalized spline filter, pemilihan parameter penghalus λ optimal, pemilihan knot optimal, kurs. Selain itu bab ini berisi tentang penelitian sebelumnya yang relevan dan kerangka berpikir.

Bab 3 Metode Penelitian

Pada bab Metode Penelitian berisi indentifikasi masalah, fokus penelitian, jenis dan sumber data, analisis dan pemecahan masalah dan penarikan kesimpulan.

Bab 4 Hasil dan Pembahasan

Pada bab Hasil dan Pembahasan berisi hasil penelitian dan bahasan sebagai jawaban atas permasalahan.

Bab 5 Penutup

Pada bab ini bersisi tentang kesimpulan dari pembahasan dan saran terkait dengan kesimpulan.

1.6.3 Bagian Akhir

Bagian ini meliputi daftar pustaka dan lampiran-lampiran yang mendukung.

BAB 2

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 Regresi Parametrik

Misalkan dimiliki sekumpulan data berpasangan (x, y) dan hubungan antara kedua variabel diasumsikan mengikuti model regresi

$$y = f(x) + \varepsilon \quad (2.1)$$

dengan $f(x)$ kurva regresi dan ε *error random* yang diasumsikan berdistribusi normal independen dengan mean nol dan variansi σ^2 . Dalam regresi parametrik terdapat asumsi yang sangat kuat dan kaku yaitu bentuk kurva regresi diketahui, misalnya linier, kuadratik, kubik, polinomial derajat p , eksponen dan lain-lain. Untuk memodelkan data menggunakan regresi parametrik linier, kuadratik, kubik atau yang lain-lain, umumnya dimulai dengan membuat *scatterplot*. Apabila *scatterplot* ini terdapat kecenderungan data mengikuti pola linier maka digunakan model regresi parametrik linier, sebaliknya jika *scatterplot* data terdapat kecenderungan pola kuadratik maka digunakan model regresi parametrik kuadratik, dan seterusnya (Wahba, 1990 dan Eubank, 1999).

2.1.2 Regresi Nonparametrik

Regresi nonparametrik merupakan model regresi yang digunakan untuk mengetahui hubungan antara variabel respon dengan variabel prediktor dengan asumsi bentuk kurva regresi tidak diketahui. Regresi nonparametrik hanya memiliki asumsi bahwa kurva regresi yang dibentuk bersifat *smooth* atau mulus. Model regresi nonparametrik secara umum dapat ditulis sebagai berikut

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

dengan keterangan sebagai berikut:

y_i : variabel respon pengamatan ke- i .

x_i : variabel prediktor pengamatan ke- i .

$f(x_i)$: fungsi regresi pada pengamatan ke- i yang tidak diketahui bentuknya.

ε_i : residual atau eror untuk pengamatan ke- i yang diasumsikan independen dengan mean nol dan variansi σ^2

Dalam kasus nonparametrik pemenuhan asumsi kenormalan eror tidak diperlukan, karena penggunaan regresi nonparametrik tidak terikat akan asumsi-asumsi seperti yang mendasari pendekatan parametrik. Pendekatan nonparametrik sering disebut pendekatan bebas distribusi.

Model regresi nonparametrik lebih mendekati kenyataan dibandingkan dengan model regresi parametrik, karena $f(x_i)$ diasumsikan berada pada fungsi

dengan dimensi yang tidak terbatas. Regresi nonparametrik mempunyai keunggulan dibandingkan regresi parametrik, yaitu mempunyai tingkat fleksibilitas yang lebih tinggi, artinya data mampu mencari bentuk estimasi kurva regresi secara mandiri tanpa dipengaruhi faktor dari peneliti. Pada regresi nonparametrik, estimasi $f(x_i)$ sama dengan estimasi koefisien-koefisien regresi, sehingga dapat membantu menentukan bagaimana menduga fungsi regresi yang lebih umum (Eubank, 1999).

2.1.3 Pemodelan Data *Time series* dengan Regresi Nonparametrik

Peramalan merupakan memperkirakan besarnya atau jumlah sesuatu pada waktu yang akan datang berdasarkan data masa lampau yang dianalisis secara alamiah khususnya menggunakan metode statistika (Sudjana, 1984:254). Ada dua jenis model peramalan yang utama, yaitu model *time series* dan model regresi. Pada model *time series*, peramalan masa depan dilakukan berdasarkan nilai pada masa lalu dari suatu variabel dan kesalahan pada masa lalu, tujuan model *time series* adalah menemukan pola dalam runtun data histories dan menerapkan pola tersebut ke masa depan. Sedangkan model regresi mengasumsikan bahwa faktor yang diramalkan menunjukkan suatu hubungan sebab akibat dengan satu atau lebih variabel bebas.

Hardle (1990) menemukan suatu gagasan baru yaitu dengan memodifikasi data *time series* menjadi dua variabel sehingga dapat digunakan untuk melakukan estimasi data dengan pendekatan regresi melalui dua variabel tersebut. Dua

variabel tersebut merupakan variabel prediktor (data hari lalu) dan variabel respon (nilai yang ingin diprediksi pada hari ini).

Pada dasarnya (X_i, Y_i) , $i = 1, 2, \dots, n$ dalam pemodelan regresi adalah saling independen. Namun dalam prakteknya sering dijumpai bahwa asumsi independensi data tersebut tidak dipenuhi, misalnya dalam kasus pengamatan data yang telah dicatat dalam urutan waktu dari suatu obyek penelitian yang mana respon obyek sekarang tergantung dari respon sebelumnya. Oleh karena itu perlu disusun suatu pemodelan data yang asumsi independensi datanya tidak dipenuhi. Ada 3 konsep dasar matematika yang mendasari pemodelan ini (Hardle, 1990), yaitu

- a. Model (S): suatu barisan stasioner $\{(X_i, Y_i), i = 1, 2, \dots, n\}$ (boleh dependen stokastik) adalah hasil observasi dan akan diestimasi $f(x) = E(Y|X = x)$.
- b. Model (T): suatu *time series* $\{Z_i, i \geq 1\}$ adalah hasil observasi dan digunakan untuk memprediksi Z_{n+1} dengan $f(x) = E(Z_{n+1}|Z_n = x)$.
- c. Model (C): pengamatan error $\{\varepsilon_{in}\}$ dalam model regresi membentuk barisan variabel random yang berkorelasi.

Untuk memprediksi masalah *time series* (T) dapat digambarkan ke model pertama. Dengan menetapkan *time series* stasioner $\{Z_i, i \geq 1\}$. Nilai *lag* Z_{i-1} sebagai X_i dan Z_i sebagai Y_i . Kemudian untuk masalah estimasi Z_{n+1} dari $\{Z_i\}$,

$i = 1, 2, \dots, n$ dapat dianggap sebagai masalah regresi pemulusan untuk $\{X_i, Y_i\}_{i=2}^n = \{Z_{i-1}, Z_i\}_{i=2}^n$. Permasalahan prediksi untuk *time series* $\{Z_i\}$ adalah sama seperti estimasi $f(x) = E(Y|X = x)$ untuk dua dimensi *time series* $\{X_i, Y_i\}_{i=1}^n$. Selanjutnya fungsi f diestimasi dengan menggunakan regresi nonparametrik.

Suatu urutan pengamatan memiliki model *time series* jika memenuhi dua hal berikut

1. Interval waktu t dinyatakan dalam satuan waktu yang sama.
2. Adanya ketergantungan antara pengamatan Z_t dengan Z_{t-k} yang dipisahkan oleh jarak waktu berupa kelipatan Δt sebanyak k kali (dinyatakan sebagai lag k).

Dasar pemikiran *time series* adalah pengamatan sekarang Z_t tergantung pada satu atau beberapa pengamatan sebelumnya Z_{t-1} .

2.1.4 Pengujian Asumsi

a. Uji Normalitas Residual

Uji Normalitas bertujuan untuk mengetahui apakah residual berdistribusi normal atau tidak. Statistik uji yang digunakan adalah *Kolmogorov-Smirnov*. Uji *Kolmogorov-Smirnov* merupakan uji normalitas menggunakan fungsi kumulatif. Hipotesis yang digunakan adalah sebagai berikut

H_0 : Residual berdistribusi normal

H_1 : Residual tidak berdistribusi normal

Kriteria Uji : Terima H_0 jika *nilai Sig.* $> \alpha$, maka nilai residual berdistribusi normal (Sukestiyarto, 2013).

b. Uji Stasioneritas

Stasioneritas mengasumsikan bahwa proses yang berlangsung ada pada kondisi *equilibrium* (tetap) pada tingkat rata-rata dan variansi konstan. Pendeteksi kestasioneran data secara formal dapat dilakukan dengan menggunakan uji akar-akar unit (*Unit Root Test*) metode *Augmented Dickey-Fuller* (ADF).

Hipotesis yang digunakan adalah sebagai berikut

H_0 : Data stasioner

H_1 : Data tidak stasioner

Kriteria Uji : Terima H_0 jika nilai kritis $\alpha <$ nilai statistik t dari ADF *Unit Root Test*, maka data bersifat stasioner (Hendikawati, 2015).

c. Uji Proses *White Noise*

Dalam analisis *time series*, residual harus mengikuti proses *white noise*, yang berarti residual harus independen dan berdistribusi normal dengan mean nol dan varian konstan. Suatu model dikatakan baik apabila uji independensi antar lag terpenuhi, parameter-parameternya signifikan dan mempunyai MSE terkecil. Uji independensi residual antar lag menggunakan kolerogram dari residual pada model ARIMA (p, d, q).

Hipotesis yang digunakan adalah sebagai berikut

H_0 : Residual *white noise*

H_1 : Residual tidak *white noise*

Kriteria Uji : Terima H_0 jika nilai probabilitas $> \alpha$, maka residual *white noise* (Hendikawati, 2015).

d. Uji Autokorelasi

Autokorelasi dalam konsep regresi berarti komponen error berkorelasi berdasarkan pada data *time series*. Metode pengujian yang sering digunakan adalah uji *Durbin-Watson*.

Hipotesis yang digunakan adalah sebagai berikut

H_0 : Tidak ada autokorelasi

H_1 : Ada autokorelasi

Uji *Durbin-Watson* akan menghasilkan nilai *Durbin-Watson* (DW) yang nantinya akan dibandingkan dengan dua nilai *Durbin-Watson* tabel, yaitu *Durbin Upper* (DU) dan *Durbin Lower* (DL). Dasar pengambilan ketentuan sebagai berikut

- 1) Jika $(4-DL) < DW < DL$ maka H_0 ditolak, yang artinya terdapat autokorelasi.
- 2) Jika $DW > (4-DU)$ dan $DW < DU$ maka H_0 ditolak, yang artinya terdapat autokorelasi.
- 3) Jika $DL < DW < DU$ atau $(4-DL) < DW < (4-DU)$, maka tidak menghasilkan kesimpulan yang pasti.

2.1.5 Spline

Salah satu model regresi dengan pendekatan nonparametrik yang dapat digunakan untuk menduga kurva regresi adalah regresi spline. Regresi spline adalah suatu pendekatan ke arah plot data dengan tetap memperhitungkan kemulusan kurva, kurva dibentuk berdasarkan titik-titik knot yang kemudian saling dihubungkan. Regresi spline memungkinkan untuk berbagai macam orde sehingga dapat dibentuk regresi spline linier, kuadratik, kubik maupun orde p . Spline adalah potongan-potongan polinomial (fungsi *truncated*) yang mulus yang masih memungkinkan memiliki sifat tersegmen (Yuedong, 2011). Penggunaan spline difokuskan kepada adanya pola data, yang pada daerah tertentu mempunyai karakteristik yang berbeda dengan daerah lain. Kemampuan ini ditunjukkan oleh fungsi *truncated* (potongan-potongan polinomial) yang melekat pada estimator dan potongan-potongan tersebut yang disebut titik knot (Bintariningrum, 2014). Salah satu kelebihan fungsi *truncated* adalah model ini mengikuti pola data sesuai pergerakannya dengan adanya titik-titik knot. Titik knot merupakan titik perpaduan bersama yang menunjukkan perubahan pola perilaku fungsi pada selang yang berbeda (Hardle, 1990). Knot diartikan sebagai suatu titik fokus dalam fungsi spline, sehingga kurva yang dibentuk tersegmen atau terbagi pada titik tersebut (Budiantara, 2015). Fungsi spline merupakan suatu gabungan fungsi polinomial, dimana penggabungan beberapa polinomial tersebut dilakukan dengan suatu cara yang menjamin sifat kontinuitas pada knot-knot. Regresi spline

cenderung mencari sendiri estimasi data kemana pun data tersebut bergerak. Spline mempunyai keunggulan dalam mengatasi pola data yang menunjukkan naik atau turun yang tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus (Hardle, 1990).

Menurut Ruppert, et al., (2003), *truncated power basis* dengan orde p dan m knot dapat dituliskan $1, x, \dots, x^p, (x - k_1)_+^p, \dots, (x - k_m)_+^p$.

Secara umum fungsi spline dengan orde p dan m knot dinyatakan sebagai

$$f(x_i) = \beta_0 + \sum_{i=1}^p \beta_i x_i + \sum_{m=1}^K \beta_{p+m} (x_i - k_m)_+^p \quad (2.3)$$

dimana

β_0 = konstanta

β_i = koefisien pada variabel x_i

β_{p+m} = koefisien pada variabel x_i *truncated* knot ke- m

x_i = variabel prediktor ke- i

k_m = knot ke- m pada peubah x_i

i = $1, 2, \dots, p$

m = banyaknya knot

dengan $(x_i - k_m)_+^p = \begin{cases} (x_i - k_m)^p, & x_i \geq k_m \\ 0 & , x_i < k_m \end{cases}$ sebagai fungsi *truncated* dan

k_1, k_2, \dots, k_m merupakan titik-titik knot. Nilai dari $(x_i - k_m)_+$ haruslah bilangan bulat positif. Jadi, ketika hasil pengurangan dari nilai amatan variabel prediktor x terhadap nilai titik knot berupa bilangan negatif, maka nilai tersebut akan dituliskan menjadi 0.

Sehingga dari fungsi (2.3) diperoleh model regresi spline sebagai berikut

$$y_i = f(x_i) + \varepsilon_i$$

$$y_i = \beta_0 + \sum_{i=1}^p \beta_i x_i + \sum_{m=1}^K \beta_{p+m} (x_i - k_m)_+^p + \varepsilon_i \quad (2.4)$$

Dari bentuk matematis fungsi spline di atas, dapat dikatakan bahwa spline merupakan model polinomial yang tersegmen (*piecewise polinomial*). Hanya saja, spline justru bersifat kontinu pada knot-knotnya. Spline orde ke- p dapat juga diartikan sebagai model polinomial orde ke- p pada setiap segmennya. Hal ini berarti fungsi spline merupakan suatu fungsi potongan polinomial yang setiap fungsinya tergabung dengan titik knot yang menjamin sifat kontinuitas (Eubank, 1999).

2.1.6 Fungsi Spline Linier

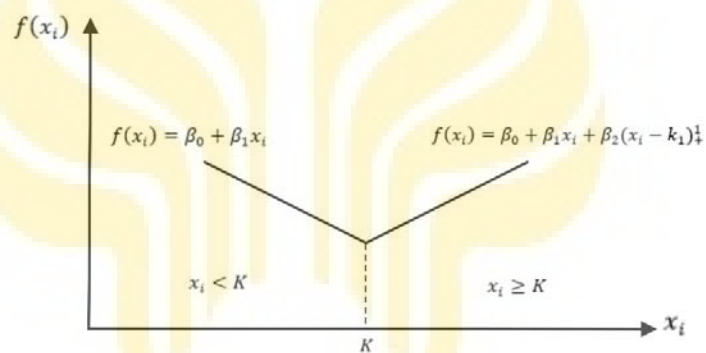
Fungsi spline linier merupakan fungsi spline dengan satu orde. Fungsi spline linier dengan satu titik knot (K) dapat disajikan dalam bentuk

$$f(x_i) = \beta_0 + \beta_1 x_i + \beta_2 (x_i - k_1)_+^1$$

Fungsi ini dapat pula disajikan menjadi *fungsi truncated* (Tripena, 2005)

$$f(x_i) = \begin{cases} \beta_0 + \beta_1 x_i & , x_i < K \\ \beta_0 + \beta_1 x_i + \beta_2 (x_i - k_1)_+^1, & x_i \geq K \end{cases}$$

Grafik spline linier dengan satu titik knot pada dapat disajikan seperti pada Gambar 2.1 berikut



Gambar 2.1. Fungsi spline linier dengan 1 titik knot

2.1.7 Penalized Spline Filter

Penalized Spline Filter (PSF) digunakan pada data *time series* yang diasumsikan erornya saling berkorelasi (Kauerman et.al., 2008). Dengan menggunakan fungsi Penalized Spline Filter (PSF), yaitu

$$\sum_{i=1}^n (y - f(x_i))^2 + \lambda \sum_{i=3}^n [(f(x_i) - f(x_{i-1})) - (f(x_{i-1}) - f(x_{i-2}))]^2 \quad (2.5)$$

dimana

$\sum_{i=1}^n (y - f(x_i))^2$: ukuran kesesuaian *time series*.

$\sum_{i=3}^n [(f(x_i) - f(x_{i-1})) - (f(x_{i-1}) - f(x_{i-2}))]^2$: ukuran kemulusan kurva dalam memetakan data.

λ : untuk nilai $\lambda \geq 0$ parameter pemulus, yaitu pengontrol keseimbangan antara kecocokan terhadap data (*goodness of fit*) dan kemulusan kurva (*penalty*).

Jadi, yang diboboti dengan λ (parameter pemulus), merupakan penalti kekasaran (*roughness penalty*) yang memberikan ukuran kemulusan atau kekasaran kurva dalam memetakan data, melalui parameter penghalus $\lambda \geq 0$. Dengan kata lain, suku kedua akan mempenalti kurva dari fungsi f . Nilai λ bervariasi dari 0 sampai $+\infty$, jika $\lambda \rightarrow +\infty$, maka penalti mendominasi persamaan (2.5) dan estimasi akan menginterpolasi titik-titik data. Sebaliknya, jika $\lambda \rightarrow 0$, maka hasil estimasi mendekati kuadrat terkecil, maka penalti tidak lagi ada dalam persamaan (2.5) dan estimasi spline menjadi konstan. Dengan demikian, parameter penghalus λ memainkan peran sentral dalam mengendalikan perimbangan (*trade off*) antara ketepatan model (*goodness of fit*) dan mulusnya penduga (Wahba 1990).

Dengan kata lain, Jika λ besar (interval kecil), maka akan diperoleh penaksir dengan bias yang besar tetapi memiliki variansi yang kecil (*oversmoothing*) atau penaksir kurva yang diperoleh akan semakin mulus. Sebaliknya jika λ kecil (interval besar), maka akan diperoleh penaksir dengan bias yang kecil namun variansinya besar (*undersmoothing*). Dengan kata lain ukuran standar jumlah kuadrat eror dan mendominasi kriteria penaksiran kurva, sehingga mengakibatkan kurva menjadi sangat fluktuatif (Adisantoso, 2010).

Sedangkan $f(x_i)$, $f(x_{i-1})$, dan $f(x_{i-2})$ sebagai fungsi filter, dimana fungsi filter ini merupakan fungsi linier dari x_1 , x_{i-1} , dan x_{i-2} dengan $i = 3, 4, \dots, x$. Sehingga dapat ditulis sebagai berikut

$$f(x_i) = X_i\beta_i, i = 3, 4, \dots, x$$

$$f(x_{i-1}) = X_{i-1}\beta_{i-1}, i = 2, 3, \dots, x$$

$$f(x_{i-2}) = X_{i-2}\beta_{i-2}, i = 1, 2, \dots, x$$

Penalized Spline Filter merupakan salah satu bentuk regresi spline yang memuat fungsi dengan memperhitungkan parameter penghalus λ . Metode optimasi yang digunakan adalah *Penalized Least Square* (PLS) sebuah metode yang memberikan komponen penghalus pada metode *least square*. PLS yaitu kriteria optimasi yang menggabungkan antara kecocokan terhadap data dan kemulusan kurva. Estimasi fungsi yang mampu memetakan data dengan baik serta mempunyai ragam eror yang kecil.

2.1.8 Pemilihan Parameter Penghalus λ Optimal

Parameter λ merupakan pengontrol keseimbangan antara kecocokan terhadap data (*goodness of fit*) dan kemulusan kurva (*penalty*). Parameter penghalus λ mempunyai pengaruh yang sangat besar dalam model regresi penalized spline (Ruppert et al., 2003). Pemilihan λ optimal dalam regresi spline pada hakekatnya merupakan pemilihan letak titik knot. Pemilihan knot dan λ pada spline dilakukan secara *trial error*. Jika λ besar maka estimasi fungsi yang diperoleh akan semakin mulus, namun kemampuan untuk memetakan data tidak

terlalu baik. Sebaliknya, jika λ kecil maka estimasi fungsi yang diperoleh akan semakin kasar (Fahrmeir dan Tuhtz, 1994). Pada nilai λ kecil kurvanya akan kasar, atau sebaliknya λ yang besar maka kurvanya akan mulus (*smooth*), dimana fungsi yang mulus terlihat jelas secara geometrik, ketika gradien kurva pada titik-titik knot tertentu tidak berubah dengan cepat (Eubank, 1999). Oleh karena itu, dalam memilih nilai diharapkan nilainya optimal agar diperoleh estimasi fungsi yang mulus dan pemetaan data yang baik. Model spline terbaik dapat dilihat dari beberapa kriteria tertentu yaitu mempunyai nilai *Mean Squared Error* (MSE) dan nilai *Generalized Cross Validation* (GCV) yang minimum. Untuk mendapatkan kurva mulus yang mempunyai λ optimal menggunakan data amatan sebanyak n , diperlukan ukuran kinerja atas estimator yang dapat diterima secara universal. Eubank, (1999) menyebutkan, ukuran kinerja atas estimator tersebut adalah

a. *Mean Squared Error* (MSE)

Ukuran kinerja atas estimator yang sederhana adalah kuadrat dari sisaan yang dirata-rata atau *Mean Squared Error*.

$$\begin{aligned} MSE(\lambda) &= n^{-1} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \\ &= n^{-1} \sum_{i=1}^n (y_i - f_\lambda(t_i))^2 \end{aligned} \quad (2.6)$$

b. *Generalized Cross Validation* (GCV)

Menurut Budiantara, (2005), GCV merupakan modifikasi dari *Cross Validation* (CV). *Cross Validation* (CV) merupakan suatu metode untuk memilih

model berdasarkan pada kemampuan prediksi dari model tersebut. CV adalah metode untuk memilih λ yang meminimumkan

$$CV(\lambda) = n^{-1} \sum_{i=1}^n \left(\frac{y_i - f_{\lambda}(t_i)}{1 - s_{ii-\lambda}} \right)^2 \quad (2.7)$$

dengan $s_{ii-\lambda}$ adalah elemen diagonal ke- i dari matriks S_{λ} . GCV diperoleh dengan mengganti $s_{ii-\lambda}$ pada persamaan (2.7) dengan $n^{-1} \sum_{i=1}^n s_{ii-\lambda} = n^{-1} \text{tr}(S_{\lambda})$

Fungsi GCV dapat dinyatakan sebagai berikut:

$$\begin{aligned} GCV(\lambda) &= n^{-1} \frac{\sum_{i=1}^n (y_i - f_{\lambda}(t_i))^2}{(1 - n^{-1} \text{tr}(S_{\lambda}))^2} \\ &= \frac{MSE(\lambda)}{(1 - n^{-1} \text{tr}(S_{\lambda}))^2} \end{aligned} \quad (2.8)$$

untuk $\text{tr}(S_{\lambda}) < n$. S_{λ} adalah matrik yang definit positif berukuran $n \times n$, dan n adalah jumlah pengamatan. $\text{tr}(S_{\lambda})$ adalah jumlah diagonal utama matrik S_{λ} perluasan dari parameter penghalus λ .

Demmler-Reinsch Orthogonalization merupakan suatu metode yang dapat mempermudah dalam memilih GCV optimal untuk setiap parameter penghalus λ . Dengan metode ini, mencari nilai *Mean Squared Error* (MSE) akan lebih cepat sehingga nilai GCV terkecil dapat diketahui dengan cepat juga.

Algoritma *Demmler-Reinsch Otrogonalization* adalah sebagai berikut

1. Menghitung dekomposisi cholesky dari $X^T X$ untuk mendapatkan matriks R
 $\ni X^T X = R^T R$.
2. Menghitung dekomposisi nilai singular matriks $(R^{-1})^T D R^{-1}$ untuk
mendapatkan $U \text{diag}(s) U^T \ni (R^{-1})^T D R^{-1} = U \text{diag}(s) U^T$
3. Menghitung matriks $\Lambda = X R^{-1} U$ dan vector $b = \Lambda^T Y$.
4. Menghitung $Y = \Lambda \frac{b}{1 + \lambda^{2p} S} = k+1$

$$\text{dimana } 1 = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1_{(p+k+1) \times 1} \end{bmatrix}$$

$$\text{dan } S = \begin{bmatrix} S_1 \\ S_2 \\ \vdots \\ S_{p+k+1} \end{bmatrix}$$

$$\text{dengan } tr(S_\lambda) = \sum_{i=1}^{p+k+1} (1 + \lambda^{2p} S_i)^{-1}$$

Kedua kriteria tersebut, baik $MSE(\lambda)$ ataupun $GCV(\lambda)$ diharapkan memiliki nilai yang minimum sehingga model regresi spline dapat dikatakan memiliki nilai λ yang optimal.

2.1.9 Pemilihan Knot Optimal

Titik knot adalah titik yang menunjukkan perubahan data pada sub-sub interval (Budiantara, 2009). Bentuk estimator spline sangat dipengaruhi oleh nilai parameter penghalus λ . Maka dari itu, pemilihan λ optimal mutlak diperlukan

untuk memperoleh estimator spline yang sesuai dengan data. Bentuk estimator spline juga dipengaruhi oleh lokasi dan banyaknya titik-titik knot. Pemilihan λ optimal dalam regresi spline pada hakikatnya merupakan pemilihan lokasi titik knot (Eubank, 1999). Sehingga dalam regresi spline, titik knot berperan penting.

Dalam menentukan titik knot, tidak digunakan nilai data terkecil dan terbesar, melainkan digunakan kombinasi linier dari $(n - 2)$ data. Penentuan titik knot dapat dilihat dari *scatterplot* dengan melihat setiap segmen yang terbentuk dari kurva. Banyaknya titik knot akan digunakan untuk menentukan model regresi spline. Letak knot dalam Penalized Spline Filter terletak pada sampel kuantil dari nilai *unique* (tunggal) variabel prediktor. Dengan knot ke- k adalah kuantil ke- j dari nilai *unique* variabel prediktor dimana $j = \frac{k}{1+k}$. Dengan kata lain, titik knot pada Penalized Spline Filter terletak pada nilai *unique* variabel prediktor. Sehingga penentuan banyak knot sangat berpengaruh untuk menentukan titik knot dalam fungsi Penalized Spline Filter tersebut.

Terdapat algoritma yang digunakan untuk mempermudah menentukan banyak knot yang optimal yaitu dengan menggunakan algoritma *Full Search*. Pemilihan banyak knot yang akan dicobakan metode tersebut berlaku untuk $K < n_{unique} - p - 1$, dengan n_{unique} adalah banyaknya nilai *unique* (tunggal) dari variabel prediktor $\{x_i\}_{i=1}^n$ dan p adalah orde polinomial (Ruppert, et al., 2003).

Dalam penulisan ini, algoritma yang digunakan untuk memilih banyak knot (K) optimal adalah algoritma *Full Search*. Kriteria algoritma *Full Search*

dalam memilih banyak knot yang optimal yaitu dengan memilih banyak knot dan parameter penghalus optimal yang dicobakan yang menghasilkan nilai GCV dan MSE terkecil (Sayuti, 2013).

2.1.10 Kurs

Menurut Krugman dan Maurice (1994), kurs (*exchange rate*) adalah harga sebuah mata uang dari satu negara yang diukur atau dinyatakan dalam mata uang lainnya. Kurs mempunyai peran sentral dalam hubungan perdagangan internasional karena kurs memungkinkan untuk membandingkan harga-harga semua barang dan jasa yang dihasilkan oleh berbagai negara. Mata uang selalu menghadapi kemungkinan penurunan nilai tukar (kurs) atau depresiasi terhadap mata uang lainya dan sebaliknya mengalami kenaikan nilai tukar. Menurut Sukirno (1994), perbedaan maupun pergerakan nilai tukar mata uang suatu negara (kurs) pada prinsipnya ditentukan oleh besarnya permintaan dan penawaran mata uang tersebut. Hal ini mengakibatkan perlunya dilakukan prediksi kurs mata uang untuk mengetahui seberapa besar nilai tukar mata uang pada masa mendatang yang bersifat harian. Dari hasil prediksi yang diperoleh, pihak-pihak yang berkepentingan dalam perdagangan internasional baik impor maupun ekspor dapat mengambil langkah-langkah strategis yang sekiranya perlu dilakukan agar tidak mengalami kerugian yang cukup besar. Faktor utama yang mempengaruhi tinggi rendahnya nilai tukar mata uang dalam negeri terhadap mata uang asing

adalah tingkat inflasi, aktifitas neraca pembayaran, perbedaan suku bunga di berbagai negara, tingkat pendapatan relatif, kontrol pemerintah, ekspektasi.

Dalam penelitian ini, metode statistika sangat berperan penting dalam memprediksi maupun menduga estimasi nilai tukar kurs rupiah terhadap dollar Amerika Serikat. Salah satu metode yang digunakan dalam memprediksi data kurs adalah analisis *time series* klasik (parametrik). Asumsi yang harus dipenuhi dalam metode ini adalah stasioneritas dan proses *white noise*. Namun data *time series* yang berfluktuasi seperti kurs sering kali tidak memenuhi asumsi stasioneritas. Apabila asumsi dari pendekatan regresi parametrik tidak terpenuhi maka estimasian dapat dilakukan dengan pendekatan nonparametrik.

Hardle dan Wahba menyarankan penggunaan regresi spline sebagai alternatif pendekatan non parametrik. Spline mempunyai keunggulan dalam mengatasi pola data yang menunjukkan naik atau turun yang tajam dengan bantuan titik-titik knot, serta kurva yang dihasilkan relatif mulus. Titik knot merupakan perpaduan bersama yang menunjukkan pola perilaku fungsi spline pada selang yang berbeda (Hardle, W., 1990). Salah satu bentuk regresi spline adalah Penalized Spline Filter.

2.2 Penelitian yang Relevan

Penelitian-penelitian sebelumnya yang relevan dengan penelitian ini, sebagai berikut

1. Agustina (2015) mengkaji pemodelan data Indeks Harga Saham Gabungan (IHSG) menggunakan regresi nonparametrik. Data *time series* IHSG biasanya dapat dimodelkan dengan model parametrik. Namun dalam model parametrik terdapat asumsi yang harus dipenuhi, yaitu asumsi stasioneritas dan *white noise*. Sedangkan data IHSG yang fluktuatif biasanya tidak memenuhi asumsi tersebut. Oleh karena itu dilakukan pemodelan data IHSG menggunakan analisis yang tidak memerlukan asumsi-asumsi yang harus dipenuhi, salah satunya adalah regresi nonparametrik penalized spline.
2. Kartikaningtiyas (2015) yaitu berupa data kurs rupiah terhadap dollar Amerika Serikat. Pada penelitian ini, fungsi basis polinomial pada estimator penalized spline kurang mampu menangani suatu data yang berubah-ubah dan terjadi ketidakstabilan numerik ketika jumlah titik-titik knot yang besar dan nilai parameter penghalus (λ) kecil atau bernilai 0. Untuk menangani ketidakstabilan numerik dan data yang berubah-ubah, peneliti melakukan perubahan terhadap fungsi basis polinomial pada estimator penalized spline dengan fungsi basis radial.

3. Ahmad (2014) mengkaji tentang estimasi model regresi nonparametrik dengan menggunakan radial smoothing berdasarkan estimator penalized spline untuk diterapkan pada data Indeks Harga Konsumen (IHK) dan inflasi bulanan Indonesia. Data tersebut dimodifikasi menjadi variabel respon yaitu data IHK, dan variabel prediktor yaitu data inflasi bulanan. Untuk mendapatkan parameter penghalus dan jumlah knot optimal dilakukan dengan cara meminimumkan GCV menggunakan algoritma *Demmer-Reinsch Orthogonalization* dengan bantuan software S-PLUS.
4. Tripena (2011) mengkaji tentang pemilihan parameter penghalus untuk estimasi regresi spline linier pada data beda potensial listrik dalam limbah cair. Metode yang digunakan adalah *Mean Squared Error* (MSE) dan *Generalized Cross Validation* (GCV). Hasil penelitian menunjukkan bahwa dalam pemilihan MSE memberikan nilai parameter penghalus lebih kecil dari pada metode GCV. Ini berarti bahwa untuk kasus data beda potensial listrik dalam limbah cair metode MSE merupakan metode yang terbaik untuk mengestimasi parameter penghalus dari regresi spline linier.
5. Laome (2010) mengkaji perbandingan model regresi nonparametrik spline dan regresi kernel. Berdasarkan kriteria kebaikan model yaitu R^2 , MSE dan uji asumsi residual, pendekatan regresi spline lebih baik dari regresi kernel untuk pertumbuhan balita. Hal ini dapat dilihat dari uji

asumsi residual yang menunjukkan bahwa pendekatan spline memenuhi semua asumsi residual.

6. Suparti (2013) menganalisis data inflasi tahunan Indonesia bulan Desember 2006 - Desember 2011 menggunakan model regresi spline. Diperoleh hasil target inflasi di Indonesia tahun 2012 yang telah ditetapkan oleh pemerintah sebesar $4.5+1\%$ dapat tercapai. Sedangkan target inflasi tahun 2013 yang telah ditetapkan oleh pemerintah sebesar $4.5+1\%$ sulit tercapai.
7. Ningrum (2017) mengkaji tentang pemodelan data nilai ekspor rempah-rempah Indonesia bulan Januari 2011 sampai bulan September 2016 yang merupakan data *time series* yang mengalami kenaikan relatif disetiap tahunnya dengan menggunakan regresi penalized spline. Model tersebut memberikan nilai *Root Mean Square Error* (RMSE) yang tergolong kecil, sehingga model tersebut merupakan model yang baik digunakan dalam prediksi.

2.3 Kerangka Berpikir

Analisis regresi telah diterapkan pada berbagai bidang, seperti administrasi bisnis, ekonomi, ilmu-ilmu pengetahuan sosial, kesehatan, dan biologi. Keberhasilan dalam menerapkan model regresi membutuhkan pemahaman yang mendalam, baik dalam teori yang mendasarinya maupun masalah-masalah yang dijumpai ketika menerapkan model regresi dalam

kehidupan nyata. Analisis regresi merupakan metode analisis data yang menggambarkan hubungan antara variabel respon dan variabel prediktor. Pada umumnya terdapat dua pendekatan untuk mengestimasi fungsi regresi yaitu pendekatan parametrik dan nonparametrik. Pendekatan parametrik mempertimbangkan jenis sebaran atau distribusi data, yaitu apakah data berdistribusi normal atau tidak. Ciri-ciri data parametrik, yaitu data berdistribusi normal, data merupakan data interval, dan jumlah data lebih dari sama dengan 30. Regresi parametrik diasumsikan bahwa bentuk kurva regresi diketahui berdasarkan teori, informasi sebelumnya, atau sumber-sumber lain yang dapat memberikan pengetahuan secara terperinci, dan juga memiliki asumsi yang kaku dan kuat. Apabila model dari pendekatan parametrik diasumsikan benar maka hasil estimasi parametrik akan efisien. Dalam menerapkan regresi parametrik sering kali asumsi-asumsinya tidak terpenuhi dan tidak semua pola hubungan dapat didekati dengan pendekatan parametrik, karena adanya suatu informasi mengenai bentuk hubungan variabel prediktor dan respon (Asriani, 2017). Untuk mengatasi penyimpangan asumsi pada regresi parametrik salah satu alternatif yang dapat digunakan adalah pendekatan regresi nonparametrik. Dalam kasus nonparametrik pemenuhan asumsi kenormalan eror tidak diperlukan, karena penggunaan regresi nonparametrik tidak terikat akan asumsi-asumsi seperti yang mendasari statistika parametrik. Statistika nonparametrik sering disebut statistika bebas distribusi. Regresi nonparametrik adalah suatu metode pemodelan yang tidak terikat akan asumsi-asumsi dari persamaan regresi tertentu yang

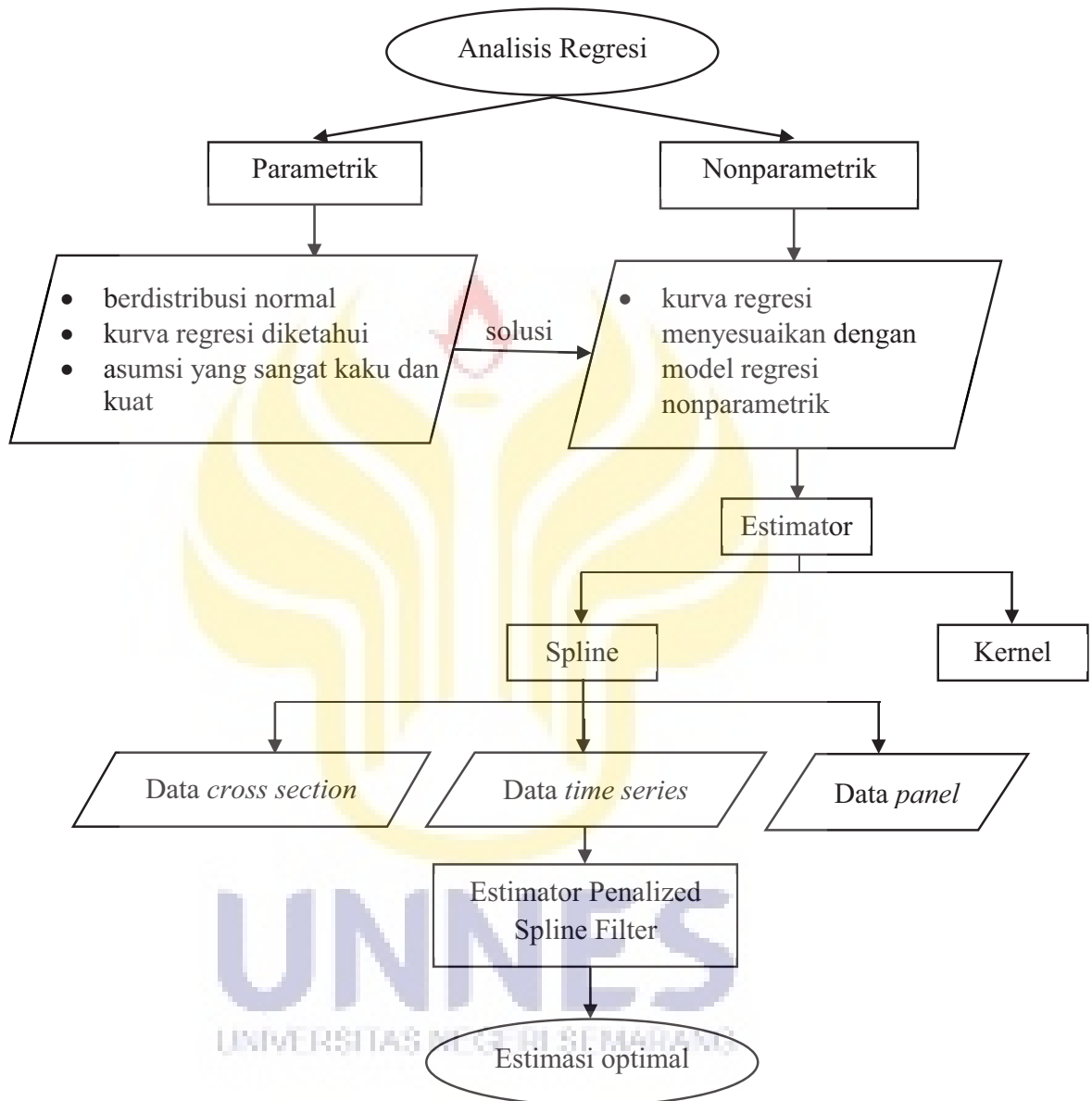
memberikan fleksibilitas yang tinggi dalam menduga sebuah model. Beberapa metode estimasi regresi nonparametrik yang unggul digunakan adalah estimator spline dan kernel. Dari penelitian-penelitian sebelumnya, beberapa peneliti memang menggunggulkan estimator spline dibandingkan kernel untuk pemodelan beberapa jenis data yaitu data *cross section*, data *time series*, dan data panel. Spline mempunyai keunggulan dalam mengatasi pola data yang menunjukkan naik atau turun yang tajam, pola data tersebut biasanya terjadi pada data *time series* yang berfluktuatif. Dengan bantuan titik-titik knot, kurva yang dihasilkan relatif mulus..

Salah satu data *time series* yang mengalami fluktuasi adalah data kurs harian. Hal ini mengakibatkan perlunya dilakukan prediksi atau pendugaan kurs mata uang untuk mengetahui seberapa besar nilai tukar mata uang pada masa mendatang yang bersifat harian. Dari hasil prediksi yang diperoleh, pihak-pihak yang berkepentingan dalam perdagangan internasional baik impor maupun ekspor dapat mengambil langkah-langkah strategis yang sekiranya perlu dilakukan agar tidak mengalami kerugian yang cukup besar. Pada data *time series* serig ditemui data dengan erornya saling berkolerasi.

Salah satu bentuk regresi spline untuk pemodelan data *time series* dengan data yang diasumsikan erornya saling berkolerasi atau autokorelasi adalah Penalized Spline Filter (PSF). PSF memiliki fungsi penalti berupa fungsi filter dan sifat tersegmen kontinu yang memberikan fleksibilitas yang lebih baik

sehingga memungkinkan untuk menyesuaikan diri secara efektif terhadap karakteristik fungsi regresi.





Gambar 2.2. Kerangka Berpikir

BAB 5

PENUTUP

5.1 Simpulan

Berdasarkan penelitian yang telah dilakukan, maka dapat ditarik kesimpulan yaitu sebagai berikut

1. Hasil estimasi parameter model regresi nonparamterik Penalized Spline Filter menggunakan *Penalized Least Square* diperoleh

$$\hat{y} = X\hat{\beta} = X((X^T X + \lambda P f)^{-1} + X^T$$

dengan,

$$\hat{\beta} = ((X^T X + \lambda P f)^{-1} + X^T$$

2. Untuk aplikasi regresi Penalized Spline Filter pada data kurs rupiah terhadap dollar Amerika Serikat yaitu diperoleh model terbaik dengan menggunakan 1 titik knot dan parameter penghalus λ sebesar 100000

$$\hat{y} = 1560,984 + 0,8830713x_i - 0,03445666(x_i - 13296)_+$$

Dengan nilai GCV sebesar 2544,012 dan R^2 sebesar 84,3%. Hal ini menunjukkan bahwa model di atas merupakan model Penalized Spline Filter yang optimal dalam menjelaskan kurs rupiah terhadap dollar AS tanggal 1 November 2016 sampai 31 Januari 2017 pada setiap harinya, termasuk pola perubahan yang teridentifikasi dalam model tersebut. Hasil

prediksi menunjukkan adanya suatu kesamaan pola terhadap data kurs yang sebenarnya.

5.2 Saran

Dari hasil penelitian ini, ada beberapa saran yang digunakan untuk penelitian selanjutnya yaitu sebagai berikut

1. Data yang diambil untuk simulasi harus sudah memenuhi pola regresi nonparametrik spline.
2. Penggunaan algoritma *Full Search* dan *Demmler-Reinsch Orthogonalization* mempermudah dalam proses pencarian knot optimal dan parameter penghalus λ optimal, agar pencarian tidak terlalu lama dan meminimumkan terjadi eror.
3. Memilihan titik knot dan parameter penghalus λ optimal harus tepat yaitu cukup dengan melihat nilai GCV terkecil saja dan juga tidak perlu melihat nilai MSE karena MSE bagian dari GCV.

DAFTAR PUSTAKA

- Adisantoso, J. 2010. Menentukan Parameter Pemulus pada Model Regresi Smoothing Spline.
- Agustina, N. 2015. Pemodelan Data Indeks Harga Saham Gabungan Menggunakan Regresi Penalized Spline. Skripsi. Semarang: FSM Universitas Diponegoro.
- Ahmad, N.R.R. 2014. Estimasi Model Regresi Nonparametrik Menggunakan Radial Smoothing Berdasarkan Estimator Penalized Spline. Skripsi. Universitas Airlangga.
- Asriani, E.D., Sugiman, Hendikawati P. 2016. Estimasi Multivariate Adaptive Regression Spline (MARS) pada Indeks Harga Saham Gabungan (IHSG). *Unnes Journal of Mathematics*. Vol 5. No 1.
- Bintariningrum, M. F., dan I. N. Budiantara. 2014. Pemodelan Regresi Nonparametrik Spline Truncated dan Aplikasinya pada Angka Kelahiran Kasar Di Surabaya. *Jurnal Sains dan Seni Pomits*. Vol 3. No 1. Hal. 7-12.
- Bloch, A. 2014. Penalized Splines as Time Series Filters in Economics Theoretical and Practical Aspects in the Frequency and Time Domain. *Munich Discussion Papers in Economics*
- Budiantara, I. N. 2005. Penentuan Titik-titik Knot dalam Regresi Spline. *Jurnal Berkala*. Jurusan Statistika FMIPA ITS. Surabaya. Vol 15. No 3.
- Budiantara, I. N. 2009. Spline Dalam Regresi Nonparametrik dan Semiparametrik: Sebuah pemodelan Statistika Masa Kini dan Masa Mendatang, Surabaya: ITS Press.
- Eubank, R. L. 1999. *Spline Smoothing and Nonparametric Regression*. New York: Marcel Dekker.
- Fahrmeir, L. & Tutz, G. 1994. *Multivariate Statistical Modelling Based on Generalized Linear Models*. Springer-Verlag. New York
- Hardle, W. 1990. *Applied Nonparametric Regression*. New York: Cambridge University Press.
- Hendikawati, P. 2015. Peramalan Data Runtun Waktu. Semarang: FMIPA Universitas Negeri Semarang.
- JianQing, Qi Lei dan Tong Xin. 2016. Penalized Least Squares Estimation with Weakly Dependent Data. *MSC*. Vol 59. No 12. Hal 2335-2354

- Kartikaningtiyas, H.S. 2015. Pemodelan Kurs Rupiah Terhadap Dollar Amerika Serikat Menggunakan Regresi Penalized Spline Berbasis Radial. *Jurnal Gaussian*. Vol 4. No 3. Hal 533-541.
- Kauerman, Goran dan Tatyana Krivobokova. 2008. *Filtering Time series with Penalized Splines*. New York: New School.
- Krivobokova, T dan Kauerman, G. 2007. A Note on Penalized Spline Smoothing with Correlated Errors. *Journal of the American Statistical Association*. Vol. 102. Hal. 1328-1337.
- Krugman R. P. dan Maurice O. 1994. *Ekonomi Internasional. Teori dan Kebijakan*. Jakarta: PT. Raja Grafindo Persada
- Laome, L. 2010. Perbandingan Model Regresi Nonparametrik dengan Regresi Spline dan Kernel. *Jurnal Ilmiah Matematika dan Terapan*. Volume 7. Nomor 1. Halaman 1-7.
- Ningrum, B.M. 2017. Model Regresi Penalized Spline pada Data Nilai Ekspor Rempah-Rempah Indonesia. Skripsi. Surakarta : MIPA Universitas Negeri Surakarta.
- Paige, R.L dan Trindade. A. 2010. The Hodrick-Prescott Filter: A special case of penalized spline smoothing. *Electronic Journal of Statistics*. Vol 4. Hal 856-874
- Polasek, W. 2011. *The Hodrick-Prescott (HP) Filter as a Bayesian Regression Model*. Department of Economics and Finance Institute for Advanced Studies.
- Pujilestari, S., Dwidayati N., Sugiman. 2017. Pemilihan Model Regresi Linier Berganda Terbaik pada Kasus Multikolinieritas Berdasarkan Metode Principal Component Analisis (PCA) dan Metode Stepwise. *Unnes Journal of Mathematics*. Vol 6. No 1. Hal 70-81.
- Ruppert, D., Wand, M. P. and Carroll, R. J. 2003. *Semiparametric Regression*. New York: Cambridge University Press.
- Sudjana. 1984. *Metode Statistika*. Bandung: Penerbit Tarsito. Sugiyono, Wibowo Eri., Juli 2004.
- Sukestiyarno, Y.L. 2013. *Statistika Dasar*. Yogyakarta: ANDI.
- Sukirno, S. 1994. *Teori Makro Ekonomi*. Jakarta: Raja Grafindo Persada
- Suparti. 2013. Analisis Data Inflasi di Indonesia Menggunakan Model Regresi Spline. *Media Statistika*. Volume 6. Nomor 1. Halaman 1-9.

Tripena, A. 2011. Pemilihan Parameter Penghalus Dalam Regresi Spline Linier. *JMP*. Volume 3. Nomor 1.

Wahba, G. 1990. *Spline Models for Observational Data*. Philadelphia: SIAM.

www.bi.go.id

Yuedong, W. 2011. *Smoothing Splines Methods and Applications*. Santa Barbara, California, USA: University of California.

