



**TESIS SS142501**

**ALGORITMA GENETIKA UNTUK IDENTIFIKASI DAN  
ESTIMASI PARAMETER MODEL SELF-EXCITING  
THRESHOLD AUTOREGRESSIVE**

**MAULIDA NURHIDAYATI  
NRP. 1313 201 017**

**DOSEN PEMBIMBING  
Dr. Irhamah, S.Si., M.Si.**

**PROGRAM MAGISTER  
JURUSAN STATISTIKA  
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2015**



**THESIS SS142501**

**GENETICS ALGORITHM FOR IDENTIFICATION AND PARAMETER  
ESTIMATION OF SELF-EXCITING THRESHOLD  
AUTOREGRESSIVE MODEL**

**MAULIDA NURHIDAYATI  
NRP. 1313 201 017**

**SUPERVISOR  
Dr. Irhamah, S.Si., M.Si.**

**PROGRAM OF MAGISTER  
DEPARTMENT OF STATISTIC  
FACULTY OF MATHEMATICS AND NATURAL SCIENCES  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2015**

**ALGORITMA GENETIKA UNTUK IDENTIFIKASI DAN ESTIMASI  
PARAMETER MODEL *SELF-EXCITING THRESHOLD*  
*AUTOREGRESSIVE***

Tesis disusun untuk memenuhi salah satu syarat memperoleh gelar  
Magister Sains (M.Si)  
di  
Institut Teknologi Sepuluh Nopember

Oleh :

**MAULIDA NURHIDAYATI**  
**NRP. 1313201017**

Tanggal Ujian : 5 Juni 2015  
Periode Wisuda : September 2015

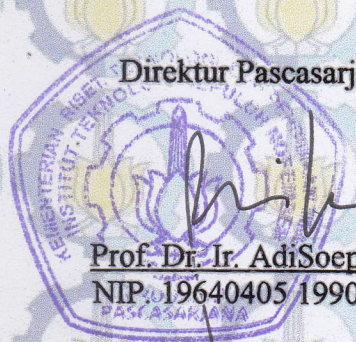
Disetujui Oleh :

1. Dr. Irhamah, S.Si., M.Si (Pembimbing)  
NIP. 19780406 200112 2 002

2. Dr. Brodjol Sutijio S.U., M.Si (Penguji)  
NIP. 19660125 199002 1 001

3. Dr. Santi Puteri Rahayu, S.Si., M.Si (Penguji)  
NIP. 19750115 199903 2 003

Direktur Pascasarjana ITS



Prof. Dr. Ir. Adi Soeprijanto, MT  
NIP. 19640405 199002 1 001  
PASCASARJANA

**ALGORITMA GENETIKA UNTUK IDENTIFIKASI DAN ESTIMASI  
PARAMETER MODEL *SELF-EXCITING THRESHOLD*  
*AUTOREGRESSIVE***

Nama Mahasiswa : Maulida Nurhidayati  
NRP : 1313 201 017  
Dosen Pembimbing : Dr. Irhamah, S.Si., M.Si

**ABSTRAK**

Model *Self Exciting Threshold Autoregressive* (SETAR) adalah model deret waktu yang dapat diterapkan pada data yang mengikuti model nonlinier. Model SETAR mempartisi data menjadi beberapa *regime* dengan masing-masing *regime* mengikuti suatu model *autoregressive* (AR). Pada model SETAR terdapat masalah dalam identifikasi model. Metode yang selama ini digunakan yaitu metode *Grid Search* (GS) tidak dapat memberikan hasil yang tepat pada saat model yang diidentifikasi merupakan suatu model subset SETAR. Berdasarkan hal tersebut, pada penelitian ini digunakan Algoritma Genetika (GA) yang merupakan suatu teknik pencarian berorientasi target yang diterapkan pada proses optimasi untuk mencari solusi optimum global. Hasil analisis yang diperoleh dari data simulasi menunjukkan bahwa metode GA memberikan hasil identifikasi model subset SETAR lebih baik dibandingkan dengan metode GS. Untuk data simulasi model SETAR yang bukan subset, metode GA dan GS memberikan hasil identifikasi yang sama. Hasil estimasi yang dilakukan dengan menggunakan metode GA pada data simulasi memberikan hasil parameter yang sama dengan metode GS. Hasil model terbaik pada data return saham BBTN dan WIKA dengan menggunakan metode GA memberikan hasil AIC yang lebih kecil dibandingkan metode GS subset pada data *in sample*. Untuk peramalan data out sample, metode GA memberikan hasil MSE paling kecil dibandingkan metode GS subset.

***Kata Kunci*** : Nonlinier, SETAR, Grid Search, Algoritma Genetika, Return saham

# **GENETICS ALGORITHM FOR IDENTIFICATION AND PARAMETER ESTIMATION OF SELF-EXCITING THRESHOLD AUTOREGRESSIVE MODEL**

Student Name : Maulida Nurhidayati  
Student Identity Number : 1313 201 017  
Supervisor : Dr. Irhamah, M.Si

## **ABSTRACT**

Self Exciting Threshold Autoregressive ( SETAR ) Model is a time series model that can be applied to data that follows the nonlinear model. SETAR model divides the data into several regime in which each regime follows a model of autoregressive ( AR ). There is a problem In the SETAR model,especially in model identification. The recently used method,Grid Search (GS ) method, can not give proper results when the identified model is a subset SETAR model. Accordingly, this study used Genetic Algorithm ( GA ), which is a target -oriented search techniques applied to the optimization process to find a global optimum solution. The analysis of the data simulation indicated that the GA method gives subset SETAR model identification result better than the GS method. For non-subset Setar model simulation data, GA and GS methods give the same identification results. The estimation results which were done using GA on data simulation gives the same parameter result to the GS method. The result of best model on the stock return data of BBTN and WIKA using the GA method gives smaller AIC results compared with GS method on a subset of the data in the sample. For forecasting the out of sample data , GA method gives smaller MSE results than the GS subset method

***Keywords*** : *Nonlinear, SETAR, Grid Search, Genetics Algorithm, Stock Return*

## KATA PENGANTAR

Alhamdulillah Rabbal'alamiin. Puji syukur yang tak terhingga penulis panjatkan kehadirat Allah SWT, yang mana atas segala limpahan karunia, ridha dan rahmat-Nya, penulis akhirnya dapat menyelesaikan penulisan tesis dengan judul ***ALGORITMA GENETIKA UNTUK IDENTIFIKASI DAN ESTIMASI PARAMETER MODEL SELF-EXCITING THRESHOLD AUTOREGRESSIVE*** ini dengan baik.

Tesis ini disusun untuk memenuhi salah satu syarat memperoleh gelas Magister Sains (M.Si) di Jurusan Statistika, Fakultas Matematika dan Ilmu Pengetahuan Alam, Institut Teknologi Sepuluh Nopember Surabaya.

Terselesainya tesis ini tidak terlepas dari dukungan berbagai pihak yang telah memberikan kontribusi kepada penulis. Oleh karena itu, pada kesempatan ini penulis ingin menyampaikan terima kasih sedalam-dalamnya kepada yang terhormat:

1. Kedua orang tua tercinta, Ibu Komsatun dan Ayah Winarto, serta kakak Kolis dan adik Bidin yang senantiasa mendoakan dan memberikan motivasi untuk kesuksesan penulis.
2. Ibu Irhamah selaku dosen pembimbing dalam penyusunan tesis ini, yang telah banyak memberikan saran dan masukan dengan penuh kesabaran dan bersedia meluangkan waktu untuk membimbing penulis selama ini.
3. Bapak Brodjol Sutijo S.U. dan Ibu Santi Puteri Rahayu selaku dosen penguji, yang juga telah banyak memberi saran dan masukan guna kesempurnaan tesis ini.
4. Bapak Muhammad Mashuri selaku Ketua Jurusan Statistika, yang telah banyak memberi saran dan masukan guna kesempurnaan tesis ini.
5. Bapak dan Ibu dosen Jurusan Statistika FMIPA ITS yang telah memberikan ilmu yang bermanfaat, serta seluruh staf administrasi akademik, laboratorium, dan ruang baca Statistika FMIPA ITS yang telah memberikan pelayanan dan fasilitas selama perkuliahan.

6. Teman-teman sebimbingan, Mei, Risky, dan Mbak Rini, terima kasih atas saran, kerjasama dan kebersamaannya pada saat bimbingan dan diskusi.
7. GK35 (Tiwi, Puput, mbak Pipit), Geng Kos Cantik (Nariza, Ina), Kiki, Zul, Lala, Mbak Afsah, Irun, Aya, Evelin terima kasih atas canda tawanya, persaudaraan yang terjalin semoga tak terhenti.
8. Rekan-rekan seperjuangan Magister Statistika angkatan 2013, terima kasih atas saran, kerjasama dan kebersamaannya.

Penulis menyadari bahwa dalam penulisan tesis ini masih banyak kekurangan. Sehingga saran dan kritik dari berbagai pihak sangat penulis harapkan demi perbaikan dan kesempurnaan karya ini. Semoga tulisan ini memberikan manfaat bagi semua. Harapan penulis, semoga hasil penelitian ini bermanfaat bagi pengembangan ilmu pengetahuan. Semoga Allah meridhai kita semua. Amiin.

Surabaya, Agustus 2015

Penulis,

## DAFTAR ISI

	halaman
<b>HALAMAN JUDUL .....</b>	<b>i</b>
<b>LEMBAR PENGESAHAN .....</b>	<b>v</b>
<b>ABSTRAK .....</b>	<b>vii</b>
<b>ABSTRACT .....</b>	<b>ix</b>
<b>KATA PENGANTAR.....</b>	<b>xi</b>
<b>DAFTAR ISI.....</b>	<b>xiii</b>
<b>DAFTAR TABEL .....</b>	<b>xv</b>
<b>DAFTAR GAMBAR.....</b>	<b>xvii</b>
<b>DAFTAR LAMPIRAN .....</b>	<b>xix</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang Masalah.....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian.....	4
1.4 Manfaat Penelitian.....	4
1.5 Batasan Masalah.....	4
<b>BAB 2 KAJIAN PUSTAKA .....</b>	<b>7</b>
2.1 Analisis Deret Waktu .....	7
2.2 Fungsi Autokorkovarian dan Fungsi Autokorelasi .....	7
2.3 Fungsi Autokorelasi Parsial .....	8
2.4 Proses Autoregressive (AR).....	9
2.5 Proses Deret Waktu Nonlinier.....	10
2.6 Uji Nonlinieritas .....	10
2.6.1 Uji Terasvirta.....	10
2.6.2 Uji White .....	11
2.7 Model <i>Threshold</i> Autoregressive (TAR) .....	12
2.8 Model Self-Exciting <i>Threshold</i> Autoregressive (SETAR) .....	12
2.8.1 Estimasi Model SETAR 2 <i>Regime</i> .....	14
2.8.2 Penentuan Threhold.....	15



2.9 Uji Signifikansi Parameter.....	15
2.10 Kriteria Pemilihan Model Terbaik.....	16
2.11 Algoritma Genetika (GA).....	16
2.11.1 Kromosom .....	17
2.11.2 Fungsi <i>Fitness</i> .....	17
2.11.3 Elitisme .....	17
2.11.4 Seleksi dengan Roda Roulette .....	18
2.11.5 <i>Crossover</i> atau Pindah Silang .....	18
2.11.6 Mutasi .....	19
2.11.7 Penggantian Populasi .....	19
2.11.8 Algoritma Genetika Standar .....	20
2.12 Indeks Pasar Saham LQ45.....	20
<b>BAB 3 METODOLOGI PENELITIAN .....</b>	<b>23</b>
3.1 Sumber Data .....	23
3.2 Metode Penelitian .....	23
3.2.1 Studi Simulasi Model SETAR.....	23
3.2.2 Aplikasi pada Data Return Saham .....	29
3.2.3 Perbandingan metode SETAR-GA.....	31
<b>BAB 4 HASIL DAN PEMBAHASAN .....</b>	<b>39</b>
4.1 Studi Simulasi.....	39
4.1.1 Model SETAR(2,1,1).....	39
4.1.2 Metode Subset SETAR.....	67
4.2 Aplikasi Metode SETAR dengan GA Pada Data Return Saham .....	86
4.2.1 Return Saham BBTN .....	88
4.2.2 Return Saham WIKA.....	92
4.3 Perbandingan Metode pada Data Out Sample .....	95
<b>BAB 5 KESIMPULAN DAN SARAN .....</b>	<b>97</b>
5.1 Kesimpulan.....	97
5.2 Saran .....	97
<b>DAFTAR PUSTAKA .....</b>	<b>99</b>
<b>LAMPIRAN .....</b>	<b>101</b>
<b>BIOGRAFI PENULIS.....</b>	<b>131</b>

## DAFTAR TABEL

	halaman
Tabel 3.1 Parameter Simulasi Model SETAR(2,1,1).....	24
Tabel 3.2 Parameter Simulasi Model Subset SETAR.....	25
Tabel 3.3 Parameter Algoritma Genetika untuk Simulasi Model SETAR .....	27
Tabel 4.1 Parameter Model Simulasi SETAR(2,1,1).....	39
Tabel 4.2 Statistik Diskriptif Data Simulasi Model SETAR(2,1,1).....	40
Tabel 4.3 Uji Nonlinieritas Terasvirta dan White Simulasi SETAR (2,1,1).....	41
Tabel 4.4 Identifikasi Data Simulasi Model 1 dengan GS.....	43
Tabel 4.5 Identifikasi Data Simulasi Model 1 dengan GA .....	43
Tabel 4.6 Estimasi Parameter Data Simulasi Model 1 Untuk n=200 .....	44
Tabel 4.7 Estimasi Parameter Data Simulasi Model 1 Untuk n=500 .....	45
Tabel 4.8 Estimasi Parameter Data Simulasi Model 1 Untuk n=1000 .....	46
Tabel 4.9 Identifikasi Data Simulasi Model 2 dengan GS.....	48
Tabel 4.10 Identifikasi Data Simulasi Model 2 dengan GA .....	48
Tabel 4.11 Estimasi Parameter Data Simulasi Model 2 untuk n=200 .....	49
Tabel 4.12 Estimasi Parameter Data Simulasi Model 2 untuk n=500 .....	50
Tabel 4.13 Estimasi Parameter Data Simulasi Model 2 untuk n=1000 .....	51
Tabel 4.14 Identifikasi Data Simulasi Model 3 dengan GS.....	53
Tabel 4.15 Identifikasi Data Simulasi Model 3 dengan GA .....	53
Tabel 4.16 Estimasi Parameter Data Simulasi Model 3 untuk n=200 .....	54
Tabel 4.17 Estimasi Parameter Data Simulasi Model 3 untuk n=500 .....	55
Tabel 4.18 Estimasi Parameter Data Simulasi Model 3 untuk n=1000 .....	56
Tabel 4.19 Identifikasi Data Simulasi Model 4 dengan GS.....	58
Tabel 4.20 Identifikasi Data Simulasi Model 4 dengan GA .....	58
Tabel 4.21 Estimasi Parameter Data Simulasi Model 4 untuk n=200 .....	59
Tabel 4.22 Estimasi Parameter Data Simulasi Model 4 untuk n=500 .....	60
Tabel 4.23 Estimasi Parameter Data Simulasi Model 4 untuk n=1000 .....	61
Tabel 4.24 Identifikasi Model 5 SETAR(2,1,1) dengan GS .....	63
Tabel 4.25 Identifikasi Model 5 SETAR(2,1,1) dengan Algoritma Genetika .....	63

Tabel 4.26 Estimasi Parameter Model 5 SETAR(2,1,1) untuk n=200 .....	64
Tabel 4.27 Estimasi Parameter Model 5 SETAR(2,1,1) untuk n=500 .....	65
Tabel 4.28 Estimasi Parameter Model 5 SETAR(2,1,1) untuk n=1000 .....	66
Tabel 4.29 Parameter Simulasi Model Subset SETAR .....	67
Tabel 4.30 Statistik Diskriptif Data Simulasi Model Subset SETAR .....	68
Tabel 4.31 Uji Nonlinieritas Terasvirta dan White Simulasi Model Subset SETAR.....	68
Tabel 4.32 Identifikasi Model Subset 1 dengan Metode <i>GS</i> dan <i>GA</i> .....	70
Tabel 4.33 Estimasi Model Identifikasi 1 dengan metode CLS untuk n=200.....	71
Tabel 4.34 Estimasi Model Identifikasi 1 dengan metode CLS untuk n=500.....	72
Tabel 4.35 Identifikasi Model 2 SETAR(2,[2],[1,3]) dengan <i>GS</i> dan <i>GA</i> .....	75
Tabel 4.36 Estimasi Model Identifikasi 2 dengan metode CLS untuk n=200.....	76
Tabel 4.37 Estimasi Model Identifikasi 2 dengan metode CLS untuk n=500.....	77
Tabel 4.38 Identifikasi Model 3 SETAR(2,2,[3]) dengan <i>GS</i> dan <i>GA</i> .....	79
Tabel 4.39 Estimasi Model Identifikasi 3 dengan metode CLS untuk n=200.....	80
Tabel 4.40 Estimasi Model Identifikasi 3 dengan metode CLS untuk n=500.....	81
Tabel 4.41 Identifikasi Model 4 SETAR(2,[2],[1,3]) dengan <i>GS</i> dan <i>GA</i> .....	83
Tabel 4.42 Estimasi Model Identifikasi 4 dengan metode CLS untuk n=200.....	84
Tabel 4.43 Estimasi Model Identifikasi 4 dengan metode CLS untuk n=500.....	85
Tabel 4.44 Statistik Diskriptif Data Return Saham .....	87
Tabel 4.45 Uji Nonlinieritas Data Return Saham.....	87
Tabel 4.46 Identifikasi Model SETAR pada Data BBTN .....	89
Tabel 4.47 Identifikasi <i>GS</i> subset dan <i>GA</i> Data return Saham BBTN.....	90
Tabel 4.48 Estimasi parameter model subset Data Retun Saham BBTN.....	90
Tabel 4.49 Identifikasi Model SETAR pada Data Return Saham WIKA.....	93
Tabel 4.50 Identifikasi <i>GS</i> subset dan <i>GA</i> Data return Saham WIKA.....	93
Tabel 4.51 Estimasi parameter model subset Data Retun Saham WIKA .....	94
Tabel 4.52 Perbandingan kebaikan model Data Retun Saham BBTN .....	96
Tabel 4.53 Perbandingan kebaikan model Data Retun Saham WIKA.....	96

## DAFTAR GAMBAR

	halaman
Gambar 2.1 Contoh Crossover.....	18
Gambar 2.2 Contoh Mutasi.....	19
Gambar 3.1 Contoh Kromosom Tahap Identifikasi Simulasi Model SETAR.....	26
Gambar 3.2 Contoh Kromosom Tahap Estimasi Simulasi Model SETAR.....	26
Gambar 3.3 Flow Chart Program Secara Umum.....	32
Gambar 3.4 Flow Chart Sub Program Pemodelan SETAR.....	33
Gambar 3.5 Flow Chart Sub Program Identifikasi dengan GA.....	34
Gambar 3.6 Flow Chart Sub Program Identifikasi dengan GS.....	35
Gambar 3.7 Flow Chart Sub Program Identifikasi dengan GA.....	35
Gambar 3.8 Flow Chart Sub Program Estimasi dengan GA.....	36
Gambar 3.9 Flow Chart Sub Program Pencarian fitness pada estimasi GA.....	36
Gambar 3.10 Flow Chart Sub Program Simulasi model SETAR.....	37
Gambar 4.1 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 1.....	42
Gambar 4.2 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 2.....	47
Gambar 4.3 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 3.....	52
Gambar 4.4 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 4.....	57
Gambar 4.5 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 5.....	62
Gambar 4.6 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 1.....	70
Gambar 4.7 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 2.....	74
Gambar 4.8 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 3.....	78

Gambar 4.9 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi	
Model Subset 4 .....	82
Gambar 4.10 Time Series Plot Return Saham BBTN .....	88
Gambar 4.11 Plot PACF Data Return Saham BBTN .....	89
Gambar 4.12 Time Series Plot Return Saham WIKA .....	92
Gambar 4.13 Plot PACF Data Return Saham WIKA .....	92

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang Masalah

Data deret waktu adalah sebuah kumpulan observasi terhadap nilai-nilai sebuah variabel dari beberapa periode yang reguler, seperti harian, mingguan, bulanan, kuartalan, tahunan, dll. Data deret waktu yang merupakan data harian dapat berupa data harga saham dan laporan cuaca. Data mingguan dapat berupa informasi uang yang beredar. Data kuartalan dapat berupa data PDRB dan data tahunan dapat berupa data anggaran pemerintah.

Model untuk data deret waktu secara umum terdiri dari model deret waktu linier dan model deret waktu nonlinier. Model deret waktu linier adalah model deret waktu yang menggambarkan suatu hubungan yang linier. Sedangkan model deret waktu nonlinier adalah model deret waktu yang menggambarkan pola hubungan yang tidak linier. Model-model deret waktu seperti model *Autoregressive* (AR), *Moving Average* (MA), *Autoregressive Moving Average* (ARMA) serta *Autoregressive Integrated Moving Average* (ARIMA) adalah contoh model deret waktu linier. Untuk model deret waktu seperti *Threshold Autoregressive* (TAR), *Self Exiting Threshold Autoregressive* (SETAR), *Exponential Autoregressive* (EXPAR), *Smooth Transition Autoregressive* (STAR), dan *Markov switching* adalah contoh model deret waktu nonlinier. Model deret waktu nonlinier memberikan hasil peramalan yang lebih baik dibandingkan dengan model deret waktu linier ketika data diindikasikan mengikuti pola nonlinier.

Pada data-data keuangan seperti harga saham, kurs mata uang, dan tingkat inflasi biasanya menunjukkan fenomena kluster volatilitas, yaitu periode dimana harga menunjukkan perubahan yang bergantian untuk periode yang panjang dan diikuti periode yang menunjukkan keadaan yang stabil. Pengetahuan tentang volatilitas merupakan suatu yang sangat penting dalam berbagai area. Sebagai contoh, investor dalam pasar saham sangat tertarik dalam volatilitas harga saham karena dengan volatilitas yang tinggi berarti kerugian yang besar atau keuntungan

yang besar menciptakan ketidakjelasan yang besar (Gujarati dan Porter, 2012). Karena adanya suatu kluster volatilitas, model untuk data-data tersebut mulai didekati dengan menggunakan model deret waktu nonlinier. Kelebihan dari model ini adalah dapat menangkap fenomena seperti asimetri, lompatan data, serta fluktuasi yang tidak dapat ditangkap oleh model deret waktu linier. Salah satu model deret waktu nonlinier yang dapat digunakan adalah model *Self-Exiting Threshold Autoregressive* (SETAR). Model SETAR adalah kasus khusus dari model deret waktu *Threshold Autoregressive* (TAR) ketika *threshold* yang digunakan diambil dari nilai lag pada data deret waktu itu sendiri. (Franses & Dijk, 2003). Prinsip dasar dari model SETAR adalah membagi data menjadi beberapa *regime* berdasarkan suatu nilai *delay* dan *threshold*. Masing-masing *regime* yang terbentuk mengikuti suatu model AR atau biasa disebut model AR lokal.

Dalam melakukan identifikasi model, model SETAR memiliki beberapa masalah. Masalah tersebut terjadi pada saat menentukan nilai *threshold*, *delay*, serta order AR yang sesuai untuk masing-masing *regime*. Banyaknya jumlah parameter yang dicari mengakibatkan proses pencarian model terbaik memerlukan waktu yang lama. Berdasarkan masalah identifikasi model tersebut, Wu dan Chang (2002) serta Baragonna, Battagliab, dan Cucina (2004) menggunakan algoritma genetika (GA) untuk mendapatkan model SETAR terbaik.

Algoritma genetika (GA) adalah suatu teknik pencarian berorientasi target, biasa diterapkan pada proses optimasi mencari nilai-nilai ekstrim universal. Pada GA, teknik pencarian dilakukan sekaligus atas sejumlah solusi yang dikenal dengan populasi. Operasi yang digunakan pada GA antara lain: seleksi, crossover dan mutasi. GA menggunakan prosedur pencarian yang didasarkan pada nilai fungsi tujuan, tidak ada pemakaian gradien atau teknik kalkulus. GA memiliki beberapa keunggulan dalam proses optimasi antara lain: (1) GA bekerja pada kumpulan solusi; (2) GA mencari berdasarkan populasi dari solusi, bukan hanya satu solusi; dan (3) GA menggunakan informasi fitness yang ingin dicari, bukan turunan atau pengetahuan khusus lainnya (Sawaka, 2002).

Pasar modal di Indonesia mengalami fluktuasi yang signifikan di tengah gejolak perkembangan ekonomi dan politik Indonesia. Pasar ekuitas mengalami stagnansi pasca krisis ekonomi pada tahun 1997 dan kembali bergairah sejak 1999.

Hal ini terlihat pada saat pasar modal menjadi sarana perusahaan dalam melakukan restrukturisasi. Nilai kapasitas pasar pada tahun 2000-2002 sempat mengalami penurunan akibat kondisi ekonomi makro yang tidak stabil. Namun demikian, dengan membaiknya kondisi mikro ekonomi pada tahun 2003 memberikan pengaruh pada perdagangan di bursa hingga kapitalitas pasar mencapai Rp. 1.982,7 triliun pada akhir Desember 2007 (Tandelilin, 2010).

Pasar derivatif mulai berkembang sejak diperkenalkannya instrumen baru berupa *LQ45 index futures* atau kontrak berjangka indeks efek (KBIE) oleh BES (Bursa Efek Surabaya) pada tanggal 13 Agustus 2001. Selanjutnya pada tanggal 6 Oktober 2004, BEI memulai perdagangan kontrak opsi saham (KOS). Berkembangnya pasar derivatif diharapkan akan semakin memperkuat pasar modal Indonesia dan memperluas alternatif investasi bagi investor.

Tujuan investor dalam berinvestasi adalah memaksimalkan return, tanpa melupakan faktor resiko investasi yang harus dihadapi. Return merupakan salah satu faktor yang memotivasi investor berinvestasi dan juga merupakan imbalan atas keberanian investor menanggung resiko atas investasi yang dilakukan. Salah satu hal yang wajar jika investor menuntut tingkat return tertentu atas dana yang telah diinvestasikan. Untuk menghasilkan keputusan investasi yang tepat dan menguntungkan, belumlah cukup bagi investor jika hanya sekedar mengetahui apa yang sedang terjadi pada pasar modal saat ini dan mengapa hal itu bisa terjadi. Investor juga perlu mengetahui situasi yang akan terjadi pada pasar modal di masa yang akan datang sehingga mendapatkan return yang besar dengan resiko yang kecil.

Berdasarkan hal tersebut, pada penelitian ini akan dilakukan identifikasi serta estimasi pada data return saham dengan menggunakan model deret waktu nonlinier. Model deret waktu yang digunakan adalah SETAR dengan algoritma genetika untuk mendapatkan suatu model SETAR yang merupakan model yang optimum global.



## **1.2 Rumusan Masalah**

Berdasarkan latar belakang yang telah dijelaskan, rumusan masalah dalam penelitian ini adalah sebagai berikut

1. Bagaimana hasil identifikasi dan estimasi parameter model SETAR dengan algoritma genetika pada data simulasi?
2. Bagaimana hasil identifikasi dan estimasi parameter model SETAR dengan algoritma genetika pada data return saham?
3. Bagaimana perbandingan hasil performasi dari metode SETAR dan SETAR dengan algoritma genetika pada data return saham?

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan masalah, tujuan dari penelitian ini adalah sebagai berikut

1. Mendapatkan hasil identifikasi dan estimasi parameter model SETAR dengan algoritma genetika pada data simulasi.
2. Mendapatkan hasil identifikasi dan estimasi parameter model SETAR dengan menggunakan algoritma genetika pada data return saham.
3. Membandingkan hasil performasi dari metode SETAR dan SETAR dengan algoritma genetika pada data return saham.

## **1.4 Manfaat Penelitian**

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut :

1. Menerapkan metode SETAR dengan algoritma genetika untuk identifikasi serta estimasi parameter untuk diterapkan pada data-data deret waktu nonlinier terutama pada data-data keuangan.
2. Mengembangkan ilmu untuk mendapatkan model deret waktu nonlinier dengan tingkat akurasi ramalan terbaik.

## **1.5 Batasan Masalah**

Dalam penelitian ini, masalah dibatasi pada pengkajian data deret waktu dengan menggunakan metode SETAR dengan algoritma genetika. Pada metode

SETAR algoritma genetika dibatasi *regime* yang digunakan adalah 2. Data yang digunakan adalah data sekunder yang merupakan data return Bank Tabungan Negara (Persero) Tbk (BBTN) dan data return saham Wijaya Karya (Persero) Tbk (WIKA). Untuk data return saham BBTN dimulai dari 4 Januari 2010 sampai 31 Desember 2014 sedangkan untuk data return saham WIKA dimulai dari 1 November 2007 sampai 31 Desember 2014 data tersebut digunakan sebagai data *in sample*. Untuk data dari 1 Januari sampai 28 Februari 2015 akan digunakan sebagai validasi (data *out sample*).

Halaman ini sengaja dikosongkan

## BAB 2

### KAJIAN PUSTAKA

#### 2.1 Analisis Deret Waktu

Data deret waktu adalah sebuah kumpulan observasi terhadap nilai-nilai sebuah variabel dari beberapa periode yang reguler, seperti harian, mingguan, bulanan, kuartalan, tahunan, dll. Data deret waktu yang merupakan data harian dapat berupa data harga saham, dan laporan cuaca. Data mingguan dapat berupa informasi uang yang beredar. Data kuartalan dapat berupa data PDRB dan data tahunan dapat berupa data anggaran pemerintah.

Suatu data yang dimodelkan dengan analisis deret waktu dapat diasumsikan bahwa data tersebut dalam keadaan stasioner. Artinya tidak terjadi trend dalam mean dan varian. Dalam analisis deret waktu, data diharapkan mengikuti proses stokastik yaitu suatu proses yang dinyatakan dalam suatu variabel random  $Z(t)$  dinotasikan dengan  $Z_t$  yang mempunyai fungsi kepadatan  $f(Z_t)$ . Artinya data deret waktu pada  $t_1, t_2, \dots, t_n$  mempunyai nilai  $Z_{t_1}, Z_{t_2}, \dots, Z_{t_n}$  secara acak dari suatu distribusi probabilitas  $f(Z_{t_i})$  (Wei, 2006).

#### 2.2 Fungsi Autokorkovarian dan Fungsi Autokorelasi

Suatu proses yang stasioner  $\{Z_t\}$  memiliki mean  $E(Z_t) = \mu$  dan varian  $Var(Z_t) = E(Z_t - \mu)^2 = \sigma^2$  yang masing-masing merupakan suatu konstanta kemudian kovarian  $Cov(Z_t, Z_s)$  merupakan suatu fungsi dari perbedaan waktu  $|t - s|$ . Kovarian antara  $Z_t$  dan  $Z_{t+k}$  didefinisikan sebagai

$$\gamma_k = Cov(Z_t, Z_{t+k}) = E(Z_t - \mu)(Z_{t+k} - \mu) \quad (2.1)$$

Korelasi antara  $Z_t$  dan  $Z_{t+k}$  adalah

$$\rho_k = \frac{Cov(Z_t, Z_{t+k})}{\sqrt{Var(Z_t)}\sqrt{Var(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (2.2)$$

dengan  $Var(Z_t) = Var(Z_{t+k}) = \gamma_0$ .

Untuk suatu proses yang stasioner, fungsi autokovarian  $\gamma_k$  dan fungsi autokorelasi  $\rho_k$  memiliki sifat-sifat sebagai berikut

1.  $\gamma_0 = \text{Var}(Z_t); \rho_0 = 1.$
2.  $|\gamma_k| \leq \gamma_0; |\rho_k| \leq 1.$
3.  $\gamma_k = \gamma_{-k}$  dan  $\rho_k = \rho_{-k}$  untuk semua k.  $\gamma_k$  dan  $\rho_k$  adalah fungsi genap yang simetris pada lag k=0. Berdasarkan sifat tersebut, fungsi autokorelasi selalu digambarkan pada lag yang tidak negatif yang selanjutnya disebut sebagai *correlogram* (Wei, 2006).

### 2.3 Fungsi Autokorelasi Parsial

Autokorelasi parsial digunakan untuk mengukur korelasi antara  $Z_t$  dan  $Z_{t+k}$  setelah menghilangkan atau memisahkan dependensi linier pada variabel  $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$  terhadap  $Z_{t+k}$  dan dinyatakan sebagai

$$\text{Corr}(Z_t, Z_{t+k} | Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}) \quad (2.3)$$

Autokorelasi parsial dapat diperoleh berdasarkan model regresi dengan variabel dependen adalah  $Z_{t+k}$  dan variabel independen adalah  $Z_{t+k-1}, Z_{t+k-2} + \dots + Z_t$  sehingga model yang terbentuk adalah

$$Z_{t+k} = \phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \dots + \phi_{kk} Z_t + e_{t+k} \quad (2.4)$$

dengan

$\phi_{ki}$  : parameter regresi yang ke-i

$e_{t+k}$  : residual dengan mean nol dan tidak berkorelasi dengan  $Z_{t+k-j}$  untuk  $j = 1, 2, \dots, k$

Dari persamaan (2.4), kalikan kedua ruas dengan  $Z_{t+k-j}$  dan ekspektasikan kedua ruas diperoleh

$$\begin{aligned} Z_{t+k} Z_{t+k-j} &= Z_{t+k-j} (\phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \dots + \phi_{kk} Z_{t+k}) \\ E(Z_{t+k} Z_{t+k-j}) &= E(Z_{t+k-j} (\phi_{k1} Z_{t+k-1} + \phi_{k2} Z_{t+k-2} + \dots + \phi_{kk} Z_{t+k})) \\ \gamma_j &= \frac{\phi_{k1} \gamma_{j-1} + \phi_{k2} \gamma_{j-2} + \dots + \phi_{kk} \gamma_{j-k}}{\gamma_0} \\ \rho_j &= \phi_{k1} \rho_{j-1} + \phi_{k2} \rho_{j-2} + \dots + \phi_{kk} \rho_{j-k} \end{aligned} \quad (2.5)$$

Untuk  $j = 1, 2, \dots, k$  persamaan (2.5) dapat dituliskan sebagai berikut

$$\begin{aligned}
\rho_1 &= \phi_{k1}\rho_0 + \phi_{k2}\rho_1 + \dots + \phi_{kk}\rho_{k-1} \\
\rho_2 &= \phi_{k1}\rho_1 + \phi_{k2}\rho_0 + \dots + \phi_{kk}\rho_{k-2} \\
&\dots \\
\rho_k &= \phi_{k1}\rho_{k-1} + \phi_{k2}\rho_{k-2} + \dots + \phi_{kk}\rho_0
\end{aligned} \tag{2.6}$$

Dengan menggunakan aturan Cramer untuk  $k=1,2,\dots$  dan  $\rho_0=1$  diperoleh

$$\begin{aligned}
\phi_{11} &= \rho_1 \\
\phi_{22} &= \frac{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & \rho_2 \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 \\ \rho_1 & 1 \end{vmatrix}} \\
&\dots \\
\phi_{kk} &= \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}
\end{aligned} \tag{2.7}$$

Persamaan (2.7) adalah suatu fungsi dari  $k$  dimana  $\phi_{kk}$  disebut sebagai fungsi autokorelasi parsial (Wei, 2006).

#### 2.4 Proses Autoregressive (AR)

Proses autoregressive digunakan untuk menggambarkan suatu kondisi dimana nilai sekarang dari suatu deret waktu bergantung pada nilai sebelumnya ditambah dengan *random stock*. Menurut Wei (2006) bentuk umum dari model autoregressive orde  $p$  adalah

$$\dot{Z}_t = \phi_1 \dot{Z}_{t-1} + \phi_2 \dot{Z}_{t-2} + \dots + \phi_p \dot{Z}_{t-p} + a_t \tag{2.8}$$

atau

$$\phi_p(B) \dot{Z}_t = a_t \tag{2.9}$$

dengan

$$\phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$$

$$\dot{Z}_t = Z_t - \mu$$

## 2.5 Proses Deret Waktu Nonlinier

Menurut Wei(2006) proses deret waktu dapat dituliskan sebagai berikut

$$Z_t = \sum_{i=0}^{\infty} \psi_i a_{t-i} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \psi_{i,j} a_{t-i} a_{t-j} + \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} \psi_{i,j,k} a_{t-i} a_{t-j} a_{t-k} + \dots \quad (2.10)$$

Jika proses  $Z_t$  pada persamaan (2.10) hanya terdiri dari suku pertama maka proses yang terjadi adalah proses linier. Jika proses  $Z_t$  pada persamaan (2.10) terdiri dari suku pertama dan suku yang lain maka proses yang terjadi adalah proses nonlinier (Wei, 2006).

## 2.6 Uji Nonlinieritas

Uji nonlinieritas dilakukan untuk mengetahui apakah suatu data mengikuti pola linier atau nonlinier. Beberapa uji nonlinieritas yaitu uji terasvirta dan uji white. Kedua uji tersebut termasuk uji nonlinieritas yang tidak mensyaratkan suatu model nonlinieritas tertentu.

### 2.6.1 Uji Terasvirta

Terasvirta, Lin, dan Granger (1993) memperkenalkan suatu uji nonlinieritas yang termasuk uji kelompok *Lagrange Multiplier* yang dikembangkan dari model *neural network*. Prosedur untuk mendapatkan nilai statistik uji F adalah sebagai berikut

1. Meregresikan  $Z_t$  dengan  $1, Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$  dan menghitung residual  $\hat{a}_t$  serta menghitung jumlah kuadrat residual.

$$SSR_0 = \sum_{t=1}^T \hat{a}_t^2 \quad (2.11)$$

2. Meregresikan  $\hat{a}_t$  dengan  $1, Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$  dan m prediktor tambahan yang merupakan suku kubik atau kuadrat yang merupakan hasil pendekatan

ekspansi Taylor. Menghitung residual  $\hat{v}_t$  serta menghitung jumlah kuadrat residual.

$$SSR = \sum_{t=1}^T \hat{v}_t^2 \quad (2.12)$$

3. Menghitung nilai F

$$F = \frac{(SSR_0 - SSR) / m}{SSR / (N - p - 1 - m)} \quad (2.13)$$

dengan  $N$  adalah jumlah pengamatan.

Dibawah hipotesis linieritas dalam mean, nilai F didekati dengan distribusi F dengan derajat bebas  $m$  dan  $N - p - 1 - m$  (Terasvirta, et al., 1993).

### 2.6.2 Uji White

Uji white adalah salah satu uji nonlinier yang dikembangkan oleh Lee, White, dan Granger (1993). Uji tersebut didasarkan pada teknik pemodelan *neural network* dan merupakan uji dalam kelompok *Lagrange Multiplier(LM)* dengan menambahkan unit tersembunyi (*hidden units*) untuk jaringan linier. Prosedur untuk mendapatkan nilai statistik uji F adalah sebagai berikut

1. Meregresikan  $Z_t$  dengan  $1, Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$  dan menghitung residual  $\hat{a}_t$  serta menghitung jumlah kuadrat residual.

$$SSR_0 = \sum_{t=1}^T \hat{a}_t^2 \quad (2.14)$$

2. Meregresikan  $\hat{a}_t$  dengan  $1, Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$  dan  $m$  prediktor tambahan yang merupakan nilai dari fungsi distribusi kumulatif dari logistik yaitu  $\psi(\gamma' w_t) = (1 + e^{-\gamma' w_t})^{-1} - \frac{1}{2}$ . Menghitung residual  $\hat{v}_t$  serta menghitung jumlah kuadrat residual.

$$SSR = \sum_{t=1}^T \hat{v}_t^2 \quad (2.15)$$

3. Menghitung nilai F



$$F = \frac{(SSR_0 - SSR) / m}{SSR / (N - p - 1 - m)} \quad (2.16)$$

dengan  $N$  adalah jumlah pengamatan.

Dibawah hipotesis linieritas dalam mean, nilai  $F$  didekati dengan distribusi  $F$  dengan derajat bebas  $m$  dan  $N - p - 1 - m$  (Lee, et al., 1993).

## 2.7 Model *Threshold Autoregressive* (TAR)

Model *Threshold Autoregressive* (TAR) pertama kali diusulkan oleh Tong dan Lim (1980). TAR adalah alternatif model untuk mendeskripsikan deret waktu periodik. Model ini dimotivasi oleh beberapa karakteristik nonlinier yang biasa ditemukan dalam kehidupan sehari-hari seperti adanya asimetri dalam pola turun dan naik suatu proses, fenomena lompatan, serta frekuensi ketergantungan amplitudo yang tidak dapat ditangkap oleh model deret waktu linier. Model ini menggunakan *threshold* untuk meningkatkan pendekatan linier.

Model TAR dapat dituliskan sebagai berikut (Wei, 2006)

$$Z_t = \phi_0 + \sum_{i=1}^p \phi_i Z_{t-i} + \left( \alpha_0 + \sum_{i=1}^p \alpha_i Z_{t-i} \right) I \left( \frac{Z_{t-d} - \tau}{\delta} \right) + a_t \quad (2.17)$$

dengan

$Z_t$  : data deret waktu

$d$  : parameter *delay*

$\tau$  : parameter lokasi (*threshold*)

$\delta$  : parameter skala

$I(\cdot)$  : fungsi penghalus yang dapat berupa suatu fungsi logistik, eksponensial, maupun suatu fungsi indikator.

## 2.8 Model *Self-Exciting Threshold Autoregressive* (SETAR)

Model *Self-Exciting Threshold Autoregressive* (SETAR) adalah pengembangan dari model TAR. Suatu proses  $Z_t$  mengikuti suatu proses SETAR jika mengikuti model (Tsay, 1989)

$$Z_t = \phi_{0,j} + \sum_{i=1}^{p_j} \phi_{i,j} Z_{t-i} + a_{t,j}; \text{ jika } Z_{t-d} \in R_j \quad (2.18)$$

dengan

$$j = 1, 2, \dots, k$$

$d$  : bilangan bulat positif dan merupakan parameter delay

$a_{t,j}$ : barisan peubah acak yang identik, independen, dan mengikuti distribusi tertentu dengan mean nol dan varian  $\sigma_j^2$

$$R = \sum_{i=1}^k R_i \text{ dimana } R_1 = (-\infty, r_1), R_i = [r_{i-1}, r_i) \text{ untuk } i = 2, \dots, k-1, \text{ dan } R_k = [r_k, \infty)$$

serta  $-\infty < r_1 < \dots < r_{k-1} < \infty$  adalah *threshold* (Wei, 2006).

Persamaan (2.18) dapat ditulis sebagai model  $SETAR(k; p_1, p_2, \dots, p_k)$  dengan  $k$  adalah banyaknya *regime* yang dipisahkan oleh  $(k-1)$  *threshold* dan  $p_j$  menyatakan orde dari model autoregressive pada *regime* ke- $j$ . Proses yang terjadi pada masing-masing *regime* mengikuti model linier autoregressive.

Penjabaran dari model  $SETAR(k, p_1, p_2, \dots, p_k)$  pada persamaan (2.18) dapat dilihat pada persamaan (2.19).

$$Z_t = \begin{cases} \phi_{0,1} + \sum_{i=1}^{p_1} \phi_{i,1} Z_{t-i} + a_{t,1} & , \text{jika } Z_{t-d} \leq r_1 \\ \phi_{0,2} + \sum_{i=1}^{p_2} \phi_{i,2} Z_{t-i} + a_{t,2} & , \text{jika } r_1 < Z_{t-d} \leq r_2 \\ \dots & \\ \phi_{0,k} + \sum_{i=1}^{p_k} \phi_{i,k} Z_{t-i} + a_{t,k} & , \text{jika } Z_{t-d} > r_{k-1} \end{cases} \quad (2.19)$$

Berdasarkan persamaan (2.19) dapat pula dibentuk model  $SETAR(2, p_1, p_2)$  yang merupakan model SETAR dengan 2 *regime* dimana  $p_1$  menunjukkan order AR pada *regime* bawah sedangkan  $p_2$  menunjukkan order AR pada *regime* atas. Bentuk persamaan model  $SETAR(2, p_1, p_2)$  dapat dilihat pada persamaan (2.20) dengan  $d$  adalah *delay* dan  $r_1$  adalah *threshold*.

$$Z_t = \begin{cases} \phi_{0,1} + \sum_{i=1}^{p_1} \phi_{i,1} Z_{t-i} + a_{t,1} & , \text{jika } Z_{t-d} \leq r_1 \\ \phi_{0,2} + \sum_{i=1}^{p_2} \phi_{i,2} Z_{t-i} + a_{t,2} & , \text{jika } Z_{t-d} > r_1 \end{cases} \quad (2.20)$$

### 2.8.1 Estimasi Model SETAR 2 Regime

Untuk melakukan estimasi model dari persamaan (2.20) dilakukan dengan cara membagi data menjadi dua *regime* sesuai dengan *threshold* yang memisahkan kedua *regime* tersebut. Diperoleh persamaan untuk *regime* atas dan bawah sebagai berikut

$$Z_{t,1} = \phi_{0,1} + \sum_{i=1}^p \phi_{i,1} Z_{t-i} + a_{t,1} \quad , \text{ untuk } \textit{regime} \text{ bawah} \quad (2.21)$$

$$Z_{t,2} = \phi_{0,2} + \sum_{i=1}^p \phi_{i,2} Z_{t-i} + a_{t,2} \quad , \text{ untuk } \textit{regime} \text{ atas} \quad (2.22)$$

Jumlah data secara keseluruhan adalah N. Karena data dibagi kedalam 2 *regime* maka  $T_1 + T_2 = N$ . Sehingga dapat ditulis dalam bentuk matrik sebagai berikut

$$\begin{bmatrix} Z_{t,1} \\ Z_{t+1,1} \\ \dots \\ Z_{T_1,1} \\ Z_{t,2} \\ Z_{t+1,2} \\ \dots \\ Z_{T_2,2} \end{bmatrix} = \begin{bmatrix} 1 & Z_{t-1,1} & Z_{t-2,1} & \dots & Z_{t-p,1} & & & & & & \\ 1 & Z_{t,1} & Z_{t-1,1} & \dots & Z_{t-p+1,1} & & & & & & \\ \dots & \dots & \dots & \dots & \dots & & & & & & \\ 1 & Z_{T_1-1,1} & Z_{T_1-2,1} & \dots & Z_{T_1-p,1} & & & & & & \\ & & & & & 1 & Z_{t-1,2} & Z_{t-2,2} & \dots & Z_{t-p,2} & \\ & & & & & 1 & Z_{t,2} & Z_{t-1,2} & \dots & Z_{t-p+1,2} & \\ \dots & & & & & \dots & \dots & \dots & \dots & \dots & \\ & & & & & 1 & Z_{T_2-1,2} & Z_{T_2-2,2} & \dots & Z_{T_2-p,2} & \end{bmatrix} \begin{bmatrix} \phi_{0,1} \\ \phi_{1,1} \\ \dots \\ \phi_{p,1} \\ \phi_{0,2} \\ \phi_{1,2} \\ \dots \\ \phi_{p,2} \end{bmatrix} + \begin{bmatrix} a_{t,1} \\ a_{t+1,1} \\ \dots \\ a_{T_1,1} \\ a_{t,2} \\ a_{t+1,2} \\ \dots \\ a_{T_2,2} \end{bmatrix} \quad (2.23)$$

Dengan asumsi bahwa  $p = \max\{p_1, p_2\}$ ,  $\phi_{i,j} = 0$  untuk  $i > p_j$ ,  $\mathbf{a}^* = (a_1, a_2, \dots, a_p)'$  diketahui dan  $a_1, a_2, \dots, a_p = 0$ . Bentuk persamaan (2.23) dapat ditulis sebagai bentuk persamaan (2.24).

$$\mathbf{Z} = \mathbf{X}\Phi + \mathbf{a} \quad (2.24)$$

Penyelesaian untuk bentuk persamaan (2.24) dapat dilakukan dengan menggunakan OLS.

$$\begin{aligned} \mathbf{a} &= \mathbf{Z} - \mathbf{X}\Phi \\ \mathbf{a}'\mathbf{a} &= (\mathbf{Z} - \mathbf{X}\Phi)'(\mathbf{Z} - \mathbf{X}\Phi) \end{aligned} \quad (2.25)$$

Dengan meminimumkan  $\mathbf{a}'\mathbf{a}$  sama dengan  $\frac{\partial \mathbf{a}'\mathbf{a}}{\partial \Phi} = 0$  sehingga

$$\begin{aligned}
\frac{\partial \mathbf{a}'\mathbf{a}}{\partial \Phi'} &= \frac{\partial}{\partial \Phi'} [(\mathbf{Z} - \mathbf{X}\Phi)'(\mathbf{Z} - \mathbf{X}\Phi)] \\
\mathbf{0} &= \frac{\partial}{\partial \Phi'} [(\mathbf{Z}' - \Phi' \mathbf{X}')(\mathbf{Z} - \mathbf{X}\Phi)] \\
\mathbf{0} &= \frac{\partial}{\partial \Phi'} [\mathbf{Z}'\mathbf{Z} - \mathbf{Z}'\mathbf{X}\Phi - \Phi' \mathbf{X}'\mathbf{Z} + \Phi' \mathbf{X}'\mathbf{X}\Phi] \\
\mathbf{0} &= \frac{\partial}{\partial \Phi'} [\mathbf{Z}'\mathbf{Z} - 2\Phi' \mathbf{X}'\mathbf{Z} + \Phi' \mathbf{X}'\mathbf{X}\Phi] \tag{2.26} \\
\mathbf{0} &= \mathbf{0} - 2\mathbf{X}'\mathbf{Z} + 2\mathbf{X}'\mathbf{X}\Phi \\
\mathbf{X}'\mathbf{X}\Phi &= \mathbf{X}'\mathbf{Z} \\
\Phi &= (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Z}
\end{aligned}$$

### 2.8.2 Penentuan Threshold

Pada model SETAR, estimasi *threshold* membutuhkan suatu perhatian khusus. Misalkan  $k = 2$  maka hanya ada 1 *threshold* misalkan  $r_1$  yang memenuhi  $Z_{\pi_s} < r_1 < Z_{\pi_{s+1}}$ . Sehingga sebarang nilai pada interval  $[Z_{\pi_s}, Z_{\pi_{s+1}}]$  dapat digunakan untuk menemukan  $r_1$ . Pada umumnya, salah satu cara yang dapat dilakukan untuk menentukan estimasi untuk *threshold* yaitu dengan menggunakan estimasi persentil pada estimasi titik (Tsay, 1989).

### 2.9 Uji Signifikansi Parameter

Model deret waktu dibangun dengan melakukan identifikasi dan estimasi parameter dari model. Misalkan  $\hat{\phi}_i$  adalah estimasi dari  $\phi_i$ . Uji signifikan parameter dapat dilakukan dengan tahapan berikut

1. Menetapkan hipotesis

$$H_0 : \phi_i = 0$$

$$H_0 : \phi_i \neq 0, \text{ dengan } i = 1, 2, \dots, k$$

2. Statistik uji

$$t = \frac{\hat{\phi}_i}{\sqrt{\text{Var}_{\hat{\phi}_i}}}$$

Dengan derajat bebas=N-M, N adalah banyaknya pengamatan dan M adalah banyaknya parameter dalam model.

### 3. Kriteria penolakan $H_0$

Tolak  $H_0$  jika  $|t| > t_{\alpha, df}$  dengan  $df = N - M$ ,  $N$  adalah banyaknya pengamatan dan  $M$  adalah banyaknya parameter dalam model.

## 2.10 Kriteria Pemilihan Model Terbaik

*Akaike's Information Criteria (AIC)* adalah suatu kriteria pemilihan model terbaik yang diperkenalkan oleh Akaike dengan mempertimbangkan banyaknya parameter dalam model. Semakin kecil nilai AIC yang diperoleh semakin baik model yang digunakan. Kriteria AIC dapat dirumuskan sebagai berikut (Wei, 2006)

$$\begin{aligned} AIC(M) &= -2 \ln[\text{maksimum likelihood}] + 2M \\ &= N \ln\left(\frac{SSE}{N}\right) + 2M \end{aligned} \quad (2.27)$$

dengan

$SSE$  : *sum square error*

$M$  : banyak parameter dalam model

## 2.11 Algoritma Genetika (GA)

Algoritma genetika (GA) pertama kali dikembangkan oleh John Holland dari Universitas Michigan pada tahun 1975. John Holland mengatakan bahwa setiap masalah yang berbentuk adaptasi (alami maupun buatan) dapat diformulasikan dalam terminologi genetika. GA adalah simulasi dari proses evolusi Darwin dan operasi genetika atas kromosom (Dewi & Purnomo, 2005).

GA masuk dalam kelompok *evolutionary Algorithm*. Algoritma genetika didasarkan pada genetika dan seleksi alam. Elemen-elemen ini yang dipakai dalam prosedur GA. GA banyak dipakai dalam menyelesaikan masalah kombinatorial seperti TSP (*traveling salesman problem*), *crew scheduling* untuk *airline* hingga masalah kontrol. GA merupakan temuan penting dalam bidang optimasi, dimana suatu algoritma diciptakan dengan meniru mekanisme evolusi dalam perkembangan makhluk hidup. Dalam GA prosedur pencarian hanya didasarkan pada nilai fungsi tujuan, tidak ada pemakaian gradient atau teknik kalkulus (Santosa & Willy, 2011).

### 2.11.1 Kromosom

Dalam GA, kromosom merupakan bagian penting dari algoritma. Satu kromosom atau individu mewakili satu vektor solusi. Terkadang vektor solusi dapat digunakan dalam implementasi GA atau dapat juga dilakukan pengkodean. Pengkodean dilakukan untuk mewakili suatu solusi dengan menggunakan bilangan biner. Hal ini tergantung pada optimasi yang dihadapi. Setiap anggota kromosom disusun oleh gen-gen, dimana masing-masing gen mewakili elemen dari vektor solusi. Dengan dibangkitkannya populasi ini, maka akan tersedia banyak pilihan solusi.

### 2.11.2 Fungsi *Fitness*

Fungsi *fitness* digunakan untuk mengukur tingkat kebaikan atau kesesuaian suatu solusi dengan solusi yang dicari. Fungsi *fitness* bisa berhubungan langsung dengan fungsi tujuan, atau bisa juga sedikit modifikasi terhadap fungsi tujuan. Sejumlah solusi yang dibangkitkan dalam populasi akan dievaluasi menggunakan nilai *fitness*.

Untuk kasus minimasi diharapkan diperoleh nilai tujuan  $f(x)$  yang nilainya kecil sehingga digunakan  $fitness \frac{1}{f(x)}$ . Sebaliknya, untuk kasus maksimasi diharapkan diperoleh nilai tujuan  $f(x)$  yang nilainya besar sehingga digunakan  $fitness f(x)$ .

Setelah setiap solusi dievaluasi dengan fungsi *fitness*, perlu dilakukan proses seleksi terhadap kromosom. Proses seleksi dilakukan untuk memilih diantara kromosom anggota populasi ini, mana yang bisa menjadi induk (*parent*) atau melakukan identifikasi diantara populasi ini, kromosom yang akan menjadi anggota populasi berikutnya. Ada beberapa cara melakukan seleksi ini. Sebagian anggota populasi bisa dipilih untuk proses reproduksi. Cara umum yang digunakan adalah melalui *roulette wheel selection*.

### 2.11.3 Elitisme

Proses seleksi dilakukan secara random sehingga tidak ada jaminan bahwa suatu individu yang bernilai *fitness* tinggi akan selalu terpilih. Walaupun individu

dengan nilai *fitness* tertinggi terpilih, mungkin saja individu tersebut akan rusak (nilai *fitness* menurun) karena proses pindah silang atau mutasi. Oleh karena itu, untuk menjaga agar individu bernilai *fitness* tinggi tidak hilang selama evolusi, maka perlu dibuat suatu kopiannya. Prosedur ini dikenal sebagai *elitisme*.

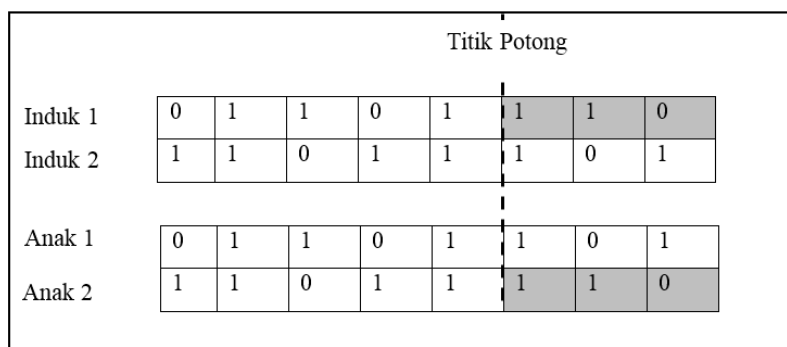
Konsep *elitisme* dalam GA berusaha untuk mempertahankan individu terbaik yang telah diperoleh di suatu generasi ke dalam generasi selanjutnya. Sehingga individu terbaik akan muncul di populasi berikutnya. *Elitisme* dimaksudkan untuk menjaga individu terbaik untuk tetap muncul di dalam populasi pada iterasi berikutnya.

#### 2.11.4 Seleksi dengan Roda Roulette

Metode seleksi roda roulette merupakan metode yang paling sederhana atau sering juga dikenal dengan nama *stochastic sampling with replacement*. Pada metode ini, individu-individu dipetakan dalam suatu segmen garis secara beraturan sedemikian hingga tiap-tiap segmen individu memiliki ukuran yang sama dengan ukuran fitnessnya. Sebuah bilangan random dibangkitkan dan individu yang memiliki segmen dalam kawasan bilangan random tersebut akan terseleksi. Proses diulang hingga diperoleh sejumlah individu yang diharapkan.

#### 2.11.5 Crossover atau Pindah Silang

Salah satu komponen yang paling penting dalam GA adalah *crossover* atau pindah silang. Sebuah kromosom yang mengarah pada solusi yang baru bisa diperoleh dari proses memindah-silangkan dua buah kromosom. Gambar 2.1 memperlihatkan contoh dari proses crossover.

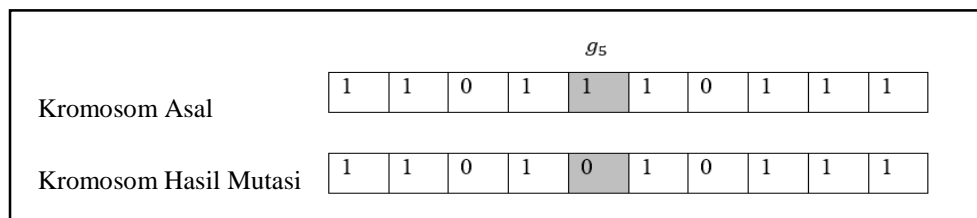


Gambar 2.1 Contoh Crossover

Pindah silang bisa juga berakibat buruk jika ukuran populasi sangat kecil. Dalam satu populasi yang sangat kecil, suatu kromosom dengan gen-gen yang mengarah ke solusi akan sangat cepat menyebar ke kromosom-kromosom lainnya. Untuk mengatasi masalah ini digunakan suatu aturan bahwa pindah silang hanya bisa dilakukan dengan suatu probabilitas tertentu  $P_c$ . Artinya, pindah silang bisa dilakukan jika suatu bilangan random  $[0,1)$  yang dibangkitkan kurang dari  $P_c$  yang ditentukan. Pada umumnya,  $P_c$  ditentukan mendekati 1, misalnya 0,8.

### 2.11.6 Mutasi

Prosedur mutasi sangatlah sederhana. Untuk semua gen yang ada, jika bilangan random yang dibangkitkan kurang dari probabilitas mutasi  $P_{mut}$  yang ditentukan maka ubah gen menjadi nilai kebalikannya (dalam binary encoding, 0 diubah 1, dan 1 diubah 0). Biasanya  $P_{mut}$  diset sebagai  $\frac{1}{n}$ , dimana  $n$  adalah jumlah gen dalam kromosom. Dengan  $P_{mut}$  sebesar ini berarti mutasi hanya terjadi pada sekitar satu gen saja. Pada GA sederhana, nilai  $P_{mut}$  tetap sama selama evolusi. Gambar menggambarkan proses mutasi yang terjadi pada gen  $g_5$ .



Gambar 2.2 Contoh Mutasi

### 2.11.7 Penggantian Populasi

Dalam GA dikenal skema pergantian populasi yang disebut *generational replacement*, yang berarti semua individu (misal  $N$  individu dalam satu populasi) dari suatu generasi digantikan sekaligus oleh  $N$  individu baru hasil pindah silang dan mutasi. Skema penggantian ini tidak realitas dari sudut pandang biologi. Di dunia nyata, individu-individu dari generasi berbeda bisa berada dalam waktu bersamaan. Fakta lain adalah individu-individu muncul dan hilang secara konstan, tidak pada generasi tertentu.



### **2.11.8 Algoritma Genetika Standar**

Secara garis besar GA standar dapat dijelaskan sebagai berikut

1. membangkitkan populasi awal
2. Set iterasi  $t=1$
3. Pilih individu terbaik untuk disalin sejumlah tertentu untuk menggantikan individu lain (*elitisme*)
4. Melakukan seleksi kompetitif untuk memilih anggota populasi sebagai induk untuk dilakukan pindah silang.
5. Lakukan pindah silang untuk induk yang terpilih.
6. Menentukan beberapa individu dalam populasi untuk mengalami proses mutasi.
7. Jika belum mencapai konvergensi set iterasi  $t=t+1$ .
8. Kembali ke langkah 2.

Dalam implementasi GA sering digunakan bilangan biner untuk mewakili nilai variabel yang dicari. Umumnya akan ditentukan terlebih dahulu beberapa interval nilai kontinyu yang dicari. Sesudah itu akan ditentukan berapa jumlah bilangan biner yang diperlukan untuk mewakili (berapa jumlah bit).

### **2.12 Indeks Pasar Saham LQ45**

Interaksi transaksi setiap sekuritas di pasar modal berbeda-beda. Sebagian besar memiliki frekuensi yang sangat tinggi dan aktif diperdagangkan di pasar modal, namun sebagian sekuritas lainnya relatif sedikit frekuensi transaksi dan cenderung bersifat pasif. Hal ini menyebabkan perkembangan dan tingkat likuiditas IHSG menjadi kurang mencerminkan kondisi real yang terjadi di bursa efek. Di Indonesia, persoalan tersebut dipecahkan dengan menggunakan indeks LQ45. LQ45 terdiri dari 45 saham BEI dengan likuiditas yang tinggi dan kapasitas pasar yang besar serta lolos seleksi menurut beberapa kriteria pemilihan. Kriteria-kriteria yang digunakan untuk memilih ke-45 saham yang masuk dalam indeks LQ45 sebagai berikut

1. Masuk dalam urutan 60 terbesar dari total transaksi saham di pasar reguler (rata-rata nilai transaksi selama 12 bulan terakhir).
2. Urutan berdasarkan kapitalitas pasar (rata-rata nilai kapitalitas pasar selama 12 bulan terakhir).

3. Telah tercatat di BEI selama paling sedikit 3 bulan.
4. Kondisi keuangan dan prospek pertumbuhan perusahaan, frekuensi dan jumlah hari transaksi di pasar reguler.

Indeks LQ45 pertama kali diluncurkan pada tanggal 24 Februari 1997. Hari dasar untuk perhitungannya adalah 13 Juli 1994 dengan nilai dasar 100. Selanjutnya bursa efek secara rutin memantau perkembangan kinerja masing-masing ke-45 saham yang masuk dalam perhitungan Indeks LQ45 (Tandelilin, 2010).

Halaman ini sengaja dikosongkan

## **BAB 3**

### **METODOLOGI PENELITIAN**

#### **3.1 Sumber Data**

Data yang digunakan dalam penelitian ini adalah data sekunder yang diperoleh dari *yahoo finance* yang merupakan data return saham Bank Tabungan Negara (Persero) Tbk (BBTN) dan data return saham Wijaya Karya (Persero) Tbk (WIKA). Data return saham tersebut selanjutnya dibagi dalam data in-sample dan data out-sample. Untuk data in-sampel digunakan data return saham BBTN dimulai dari 4 Januari 2010 sampai 31 Desember 2014 dan data return saham WIKA dimulai dari 1 November 2007 sampai 31 Desember. Untuk data *out sample* sebagai data validasi digunakan data dari 1 Januari sampai 28 Februari 2015.

#### **3.2 Metode Penelitian**

Langkah-langkah analisis dalam penelitian ini dijabarkan dalam 2 tahap. Tahap pertama adalah studi simulasi model SETAR. Pada tahapan ini, identifikasi untuk model dilakukan dengan menggunakan *grid search* dan algoritma genetika sedangkan estimasi parameter dilakukan dengan menggunakan metode *conditional least square*(CLS) dan algoritma genetika(GA). Tahap kedua akan diaplikasikan metode SETAR untuk analisis data return saham BBTN dan return saham WIKA serta menentukan model terbaik berdasarkan ramalan *in-sample* dan *out-sample*. Tahap identifikasi model dilakukan dengan menggunakan dengan *grid search* dan algoritma genetika. Selanjutnya untuk estimasi parameter dilakukan dengan menggunakan metode *CLS* dan GA.

##### **3.2.1 Studi Simulasi Model SETAR**

Studi simulasi model SETAR dilakukan untuk membangkitkan beberapa model SETAR. Model SETAR yang akan dibangkitkan adalah model SETAR(2,1,1) dan model subset SETAR. Model SETAR(2,1,1) adalah model SETAR 2 *regime* dan setiap *regime* mengikuti AR(1) dengan *delay*=1. Nilai parameter model SETAR(2,1,1) serta *threshold* yang digunakan dapat dilihat pada

Tabel 3.1. Berdasarkan tabel tersebut dapat dilihat bahwa terdapat 5 model yang dibangkitkan dengan perbedaan koefisien pada masing-masing model. Standar deviasi yang digunakan pada 5 model tersebut adalah 2. Masing-masing model dibangkitkan dengan jumlah data 200,500, dan 1000.

Tabel 3.1 Parameter Simulasi Model SETAR(2,1,1)

Model	$\varphi_{0,1}$	$\varphi_{1,1}$	$\varphi_{0,2}$	$\varphi_{1,2}$	Threshold
1	8	-0,2	8	0,2	8
2	1	0,2	4	0,2	3
3	5	0,5	9	-0,5	6
4	8	-0,8	9	-0,2	6
5.	-7	-0,5	-2	0,8	-2,5

Secara matematis model pada Tabel 3.1 dapat dituliskan sebagai berikut

a. Model 1

$$Z_t = \begin{cases} 8 - 0,2Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 8 \\ 8 + 0,2Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 8 \end{cases}$$

b. Model 2

$$Z_t = \begin{cases} 1 + 0,2Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 3 \\ 4 + 0,2Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 3 \end{cases}$$

c. Model 3

$$Z_t = \begin{cases} 5 + 0,5Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 6 \\ 9 - 0,5Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 6 \end{cases}$$

d. Model 4

$$Z_t = \begin{cases} 8 - 0,8Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 6 \\ 9 - 0,2Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 6 \end{cases}$$

e. Model 5

$$Z_t = \begin{cases} -7 - 0,5Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq -2,5 \\ -2 + 0,8Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > -2,5 \end{cases}$$

Nilai parameter model subset SETAR serta *threshold* yang digunakan dapat dilihat pada Tabel 3.2. Model 1 dan model 2 merupakan model subset SETAR dengan *delay*=1 dan merupakan model yang simulasikan oleh Barogona, Battaglia

an Cucina (2004). Model 3 adalah model yang sama dengan model 1 dengan  $delay=2$ . Seperti pada model 3, model 4 merupakan model yang sama dengan model 2 dengan  $delay=2$ . Standar deviasi yang digunakan pada 4 model tersebut adalah 1. Masing-masing model yang dibangkitkan dengan jumlah data 200 dan 500.

Tabel 3.2 Parameter Simulasi Model Subset SETAR

	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
<b>Regime 1</b>	$\phi_{1,1} = -1,2$ $\phi_{2,1} = -0,7$	$\phi_{2,1} = -0,6$	$\phi_{1,1} = -1,2$ $\phi_{2,1} = -0,7$	$\phi_{2,1} = -0,6$
<b>Regime 2</b>	$\phi_{3,2} = 0,8$	$\phi_{1,2} = 0,75$ $\phi_{3,2} = -0,5$	$\phi_{3,2} = 0,8$	$\phi_{1,2} = 0,75$ $\phi_{3,2} = -0,5$
<b>Threshold</b>	0	0	0	0
<b>Delay</b>	1	1	2	2

Secara matematis model subset SETAR pada Tabel 3.2 dapat dituliskan sebagai berikut

a. Model 1

$$Z_t = \begin{cases} -1,2Z_{t-1} - 0,7Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq 0 \\ 0,8Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > 0 \end{cases}$$

b. Model 2

$$Z_t = \begin{cases} -0,6Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq 0 \\ 0,75 - 0,5Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > 0 \end{cases}$$

c. Model 3

$$Z_t = \begin{cases} -1,2Z_{t-1} - 0,7Z_{t-2} + a_{t,1} & \text{jika } Z_{t-2} \leq 0 \\ 0,8Z_{t-3} + a_{t,2} & \text{jika } Z_{t-2} > 0 \end{cases}$$

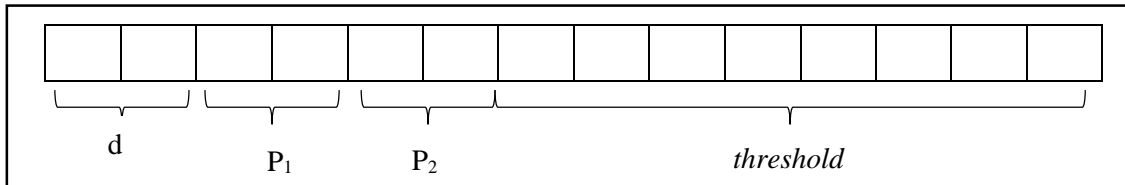
d. Model 4

$$Z_t = \begin{cases} -0,6Z_{t-2} + a_{t,1} & \text{jika } Z_{t-2} \leq 0 \\ 0,75 - 0,5Z_{t-3} + a_{t,2} & \text{jika } Z_{t-2} > 0 \end{cases}$$

Dalam melakukan analisis dengan menggunakan algoritma genetika pada model SETAR, terlebih dahulu perlu dilakukan penentuan panjang kromosom yang merupakan solusi dari model serta parameter-parameter dalam algoritma genetika. Terdapat 2 kromosom yang akan digunakan pada analisis ini. Kromosom pertama

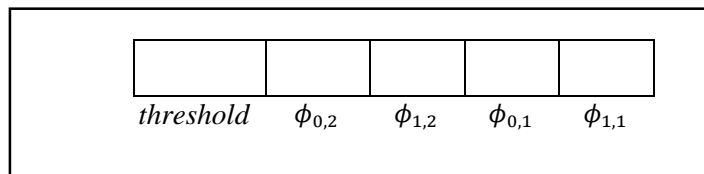
adalah kromosom yang digunakan dalam identifikasi model sedangkan kromosom kedua adalah kromosom yang digunakan untuk estimasi parameter.

Kromosom untuk tahap identifikasi terdiri dari solusi untuk orde  $d, p_1, p_2$  serta *threshold* yang selanjutnya dikodekan dalam bentuk string biner. Dalam melakukan pengkodean untuk  $d, p_1, p_2$  didasarkan pada orde tertinggi dari plot PACF yang mengalami *cuts off* dan *threshold* merupakan suatu nilai dari data yang terletak pada interval 5%-95% atau 10%-90% atau 15%-85% atau 20%-80%. Kriteria yang digunakan untuk mendapatkan solusi terbaik adalah dengan menggunakan AIC minimum. Salah satu kromosom yang digunakan dalam identifikasi model SETAR dapat dilihat pada Gambar 3.1.



Gambar 3.1 Contoh Kromosom Tahap Identifikasi Simulasi Model SETAR

Untuk kromosom pada tahap estimasi, panjang kromosom disesuaikan dengan orde dari  $p_1, p_2$ . Pada tahap estimasi, kromosom yang digunakan adalah kromosom dengan *real value*. Kromosom tersebut berisi nilai *threshold* serta parameter untuk *regime 1* dan parameter untuk *regime 2*. Kriteria yang digunakan pada tahap ini adalah meminimumkan SSE dan bias parameter. Salah satu kromosom yang digunakan dapat dilihat pada Gambar 3.2.



Gambar 3.2 Contoh Kromosom Tahap Estimasi Simulasi Model SETAR

Dalam melakukan analisis pada algoritma genetika, perlu didefinisikan terlebih dahulu parameter-parameter yang digunakan serta fungsi *fitness* yang dipakai dalam menentukan kromosom yang merupakan solusi terbaik. Nilai parameter serta *fitness* yang digunakan dapat dilihat pada Tabel 3.3. Pada tabel

tersebut dapat dilihat parameter-parameter yang digunakan pada tahap identifikasi sama dengan pada tahap estimasi. Perbedaannya terletak pada *fitness* yang digunakan pada masing-masing tahap.

Tabel 3.3 Parameter Algoritma Genetika untuk Simulasi Model SETAR

Parameter	Nilai
PopSize	200
Pc	0,8
Pm	0,1
Maksimum iterasi	200
<i>Fitness</i> tahap Identifikasi	$N * \log\left(\frac{SSE}{N}\right) + 2(p_1 + p_2 + 2)$
<i>Fitness</i> tahap Estimasi	$SSE = (Z - \bar{Z})^2$

Langkah-langkah analisis pada data simulasi pada penelitian ini adalah sebagai berikut

1. Melakukan simulasi model SETAR berdasarkan model yang ditentukan.
2. Melakukan uji nonlinieritas Terasvirta dan White.
3. Jika data memenuhi asumsi nonlinier, maka dilanjutkan pada langkah 4. Jika data tidak memenuhi asumsi nonlinier maka ulangi langkah 1.
4. Membuat plot PACF untuk menentukan order maksimum dari AR.
5. Menentukan nilai maksimum untuk  $p_1 = p$ ,  $p_2 = p$ , dan  $d = p$
6. Jika identifikasi dilakukan dengan menggunakan metode GA maka lanjut ke langkah 7. Jika identifikasi dilakukan dengan menggunakan metode GS maka lanjut langkah 8.
7. Melakukan identifikasi model SETAR dengan GA.
  - a. Mengurutkan data.
  - b. Menentukan daerah pencarian *threshold* yaitu 5%-95% daerah pencarian, 10%-90% daerah pencarian, 15%-85% daerah pencarian, dan 20%-80% daerah pencarian *threshold*.
  - c. Menentukan batas pencarian *threshold* berdasarkan persentil dari data.
  - d. Menentukan banyaknya gen yang digunakan untuk mewakili *threshold*.
  - e. Menyusun kromosom dari identifikasi model yang terdiri dari orde maksimum dari  $d, p_1, p_2$ , subset untuk  $p_1$ , subset untuk  $p_2$ , dan *threshold*.



- f. Menentukan *fitness*,  $P_c$ ,  $P_m$ , dan banyak generasi.
  - g. Melakukan inisialisasi dari generasi ke-i.
  - h. Melakukan decode kromosom dari populasi
  - i. Melakukan estimasi dengan metode CLS.
  - j. Menghitung nilai AIC yang merupakan nilai *fitness* yang digunakan.
  - k. Melakukan seleksi dengan *roulette wheel*, *crossover*, mutasi, serta *elitism* pada populasi yang dimiliki.
  - l. Diperoleh kromosom terbaik dari generasi ke-i.
  - m. Jika banyak generasi sama dengan K maka lanjut ke langkah n. Jika banyak generasi kurang dari K maka ulangi langkah g.
  - n. Diperoleh kromosom dengan AIC minimum dan dilakukan decode kromosom.
  - o. Melakukan estimasi parameter dengan metode GA.
  - p. Diperoleh model terbaik dengan AIC terkecil.
8. Melakukan identifikasi model SETAR dengan GS.
- a. Mengurutkan data.
  - b. Menentukan batas *threshold* yaitu persentil 15 dan persentil 85.
  - c. Melakukan kombinasi  $d, p_1, p_2$ , dan *threshold* dimulai dari  $d, p_1, p_2 = 1$  dan *threshold* pada batas minimum.
  - d. Melakukan estimasi dengan menggunakan metode CLS.
  - e. Menghitung nilai AIC.
  - f. Apabila semua kombinasi telah dilakukan maka diperoleh 10 model dengan AIC terkecil.
9. Diperoleh model SETAR terbaik dengan nilai AIC terkecil.

Berikut adalah langkah-langkah estimasi parameter dengan menggunakan metode CLS.

1. Berdasarkan nilai  $d, p_1, p_2$ , dan *threshold* selanjutnya ditentukan lag dari data.
2. Menentukan regime dari masing-masing data.

3. Mengalikan lag data dengan regime yang dimiliki. Hasil perkalian tersebut selanjutnya diregresikan dengan data..
4. Menghitung residual dan menentukan SSE dari data.

Untuk langkah-langkah estimasi parameter dengan menggunakan metode GA dijabarkan sebagai berikut.

1. Menentukan panjang kromosom berdasarkan jumlah parameter yang diestimasi ditambah dengan *threshold*.
2. Menentukan *fitness*,  $P_c$ ,  $P_m$ , dan banyak generasi.
3. Melakukan inisialisasi populasi dari generasi ke-i.
4. Menghitung nilai AIC yang merupakan nilai *fitness* yang digunakan.
5. Melakukan seleksi dengan *roulette wheel*, *crossover*, mutasi, serta *elitism* pada populasi yang dimiliki.
6. Diperoleh kromosom terbaik dari generasi ke-i.
7. Jika selisih antara fitness pada generasi ke-(i-1) dan generasi ke-i kurang dari  $10^{-8}$  maka lanjut langkah 8. Jika selisih lebih dari itu maka ulangi langkah 3.
8. Diperoleh kromosom dengan SSE minimum.

### 3.2.2 Aplikasi pada Data Return Saham

Aplikasi metode SETAR pada data return saham dilakukan dengan 2 tahapan. Tahap pertama adalah identifikasi model SETAR dengan menggunakan metode GS dan GA. Setelah diperoleh model SETAR selanjutnya dilakukan estimasi parameter dengan menggunakan metode CLS dan GA. Hasil yang diperoleh digunakan untuk melakukan ramalan *k-step* untuk data *out sampel*.

Tahapan analisis data return saham adalah sebagai berikut

1. Analisis Diskriptif data.
2. Melakukan uji nonlinieritas Terasvirta dan White.
3. Jika data return saham memenuhi asumsi nonlinier, maka dilanjutkan pada langkah 4. Jika data tidak memenuhi asumsi nonlinier maka data return saham dianalisis dengan menggunakan analisis untuk model-model linier seperti AR, MA, ARMA, atau ARIMA.
4. Membuat plot PACF untuk menentukan order maksimum dari AR.

5. Menentukan nilai maksimum untuk  $p_1 = p, p_2 = p$ , dan  $d = p$
6. Jika identifikasi dilakukan dengan menggunakan metode GA maka lanjut ke langkah 7. Jika identifikasi dilakukan dengan menggunakan metode GS maka lanjut langkah 8.
7. Melakukan identifikasi model SETAR dengan GA.
  - a. Mengurutkan data.
  - b. Menentukan daerah pencarian *threshold* yaitu 5%-95% daerah pencarian, 10%-90% daerah pencarian, 15%-85% daerah pencarian, dan 20%-80% daerah pencarian *threshold*.
  - c. Menentukan batas pencarian *threshold* berdasarkan persentil dari data.
  - d. Menentukan banyaknya gen yang digunakan untuk mewakili *threshold*.
  - e. Menyusun kromosom dari identifikasi model yang terdiri dari orde maksimum dari  $d, p_1, p_2$ , subset untuk  $p_1$ , subset untuk  $p_2$ , dan *threshold*.
  - f. Menentukan *fitness*, Pc, Pm, dan banyak generasi.
  - g. Melakukan inisialisasi dari generasi ke-i.
  - h. Melakukan decode kromosom dari populasi
  - i. Melakukan estimasi dengan metode CLS.
  - j. Menghitung nilai AIC yang merupakan nilai *fitness* yang digunakan.
  - k. Melakukan seleksi dengan *roulette wheel*, *crossover*, mutasi, serta *elitism* pada populasi yang dimiliki.
  - l. Diperoleh kromosom terbaik dari generasi ke-i.
  - m. Jika banyak generasi sama dengan K maka lanjut ke langkah n. Jika banyak generasi kurang dari K maka ulangi langkah g.
  - n. Diperoleh kromosom dengan AIC minimum dan dilakukan decode kromosom.
  - o. Melakukan estimasi parameter dengan metode GA.
  - p. Diperoleh model terbaik dengan AIC terkecil.
8. Melakuakn identifikasi model SETAR dengan GS.
  - a. Mengurutkan data.
  - b. Menentukan batas *threshold* yaitu persentil 15 dan persentil 85.

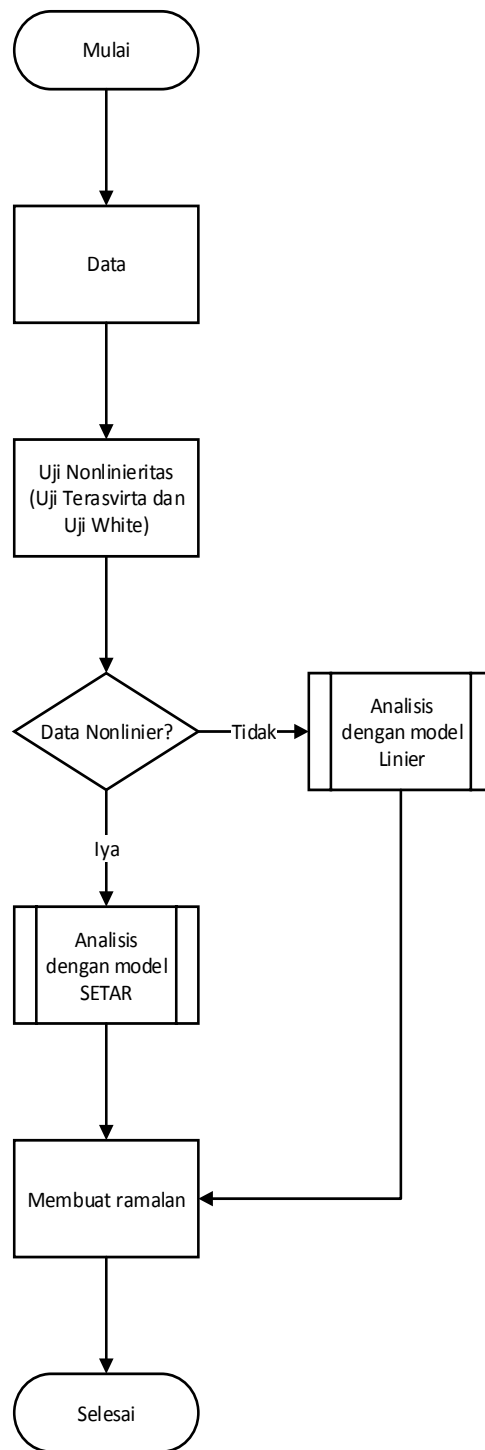
- c. Melakukan kombinasi  $d, p_1, p_2$ , dan *threshold* dimulai dari  $d, p_1, p_2 = 1$  dan *threshold* pada batas minimum.
  - d. Melakukan estimasi dengan menggunakan metode CLS.
  - e. Menghitung nilai AIC.
  - f. Apabila semua kombinasi telah dilakukan maka diperoleh 10 model dengan AIC terkecil.
9. Diperoleh model SETAR terbaik dengan nilai AIC terkecil.
  10. Membuat ramalan *in-sample* dan menghitung MSE serta AIC dari model SETAR-CLS dan model SETAR-GA.
  11. Membuat ramalan *out-sample* dan membandingkan hasilnya dengan data yang sesungguhnya. Kemudian menghitung nilai MSE dari model SETAR-CLS dan model SETAR-GA.

### 3.2.3 Perbandingan metode SETAR-GA

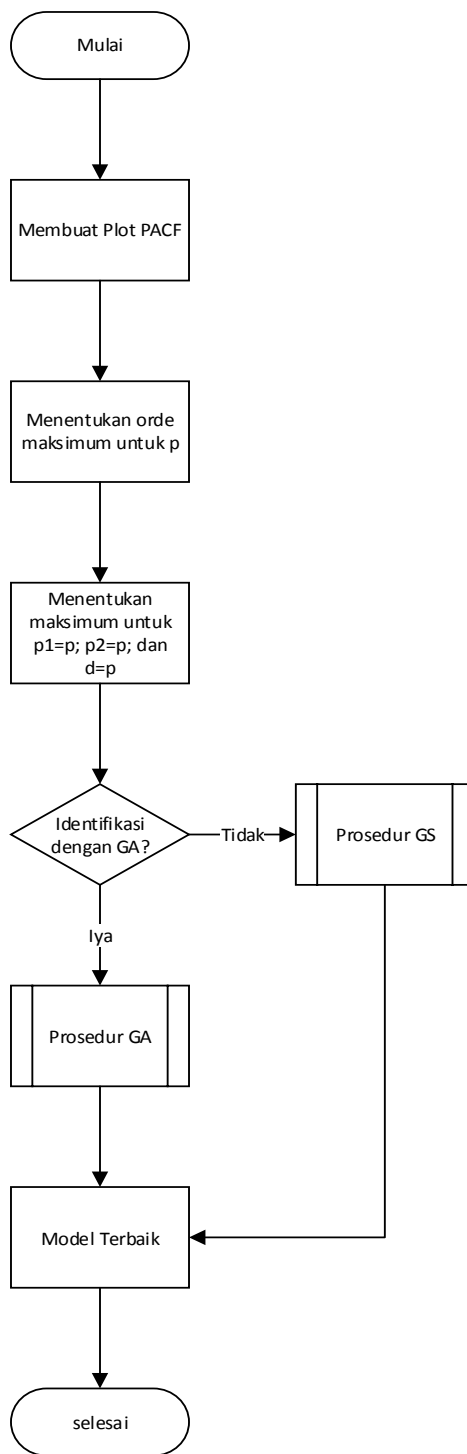
Kriteria-kriteria yang digunakan untuk membandingkan metode SETAR adalah sebagai berikut

1. Pada data *in-sample* : kriteria AIC terkecil
2. Pada data *out-sample* : kriteria MSE terkecil

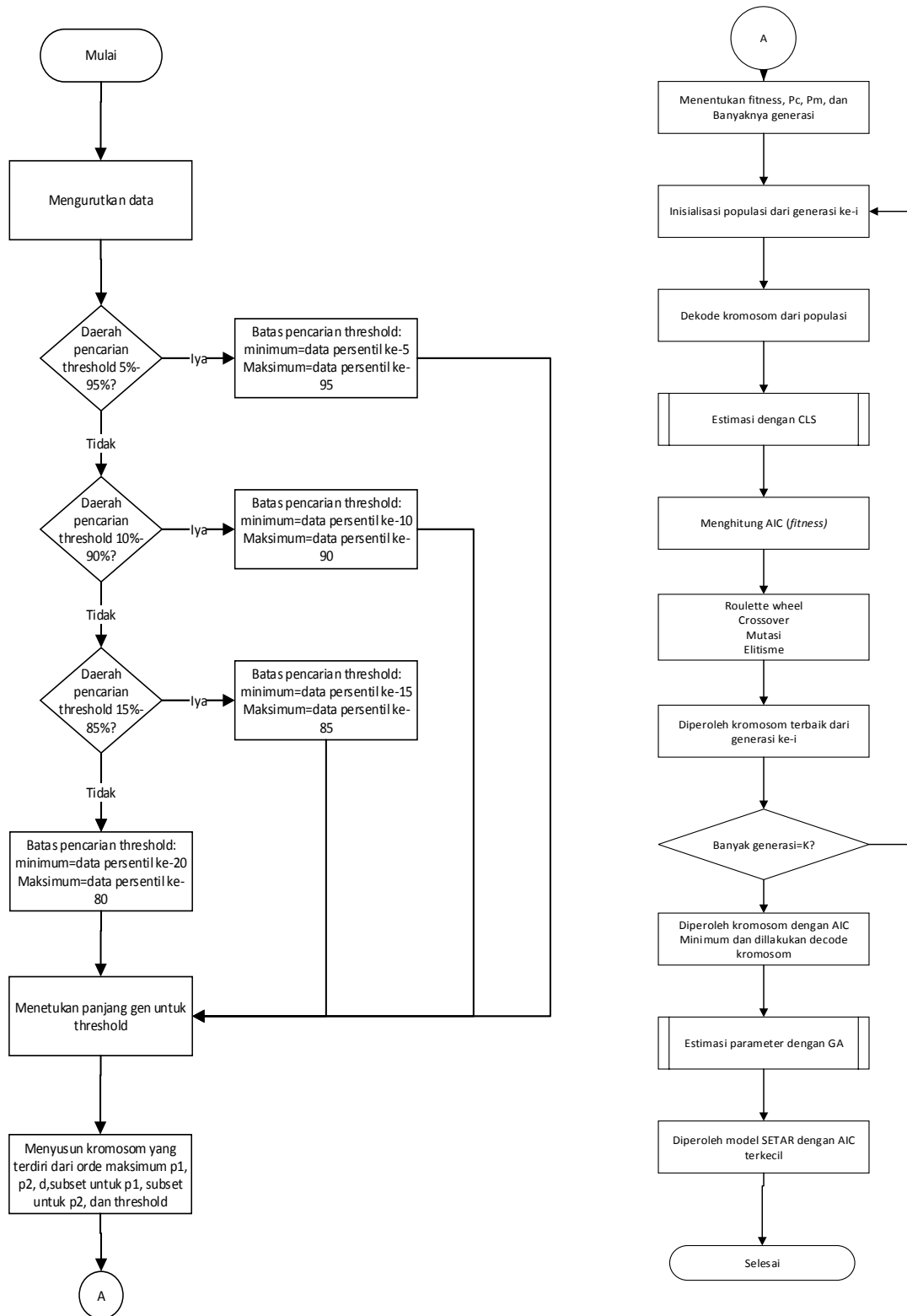
Model terbaik yang diperoleh adalah model yang memiliki nilai MSE dan AIC paling kecil.



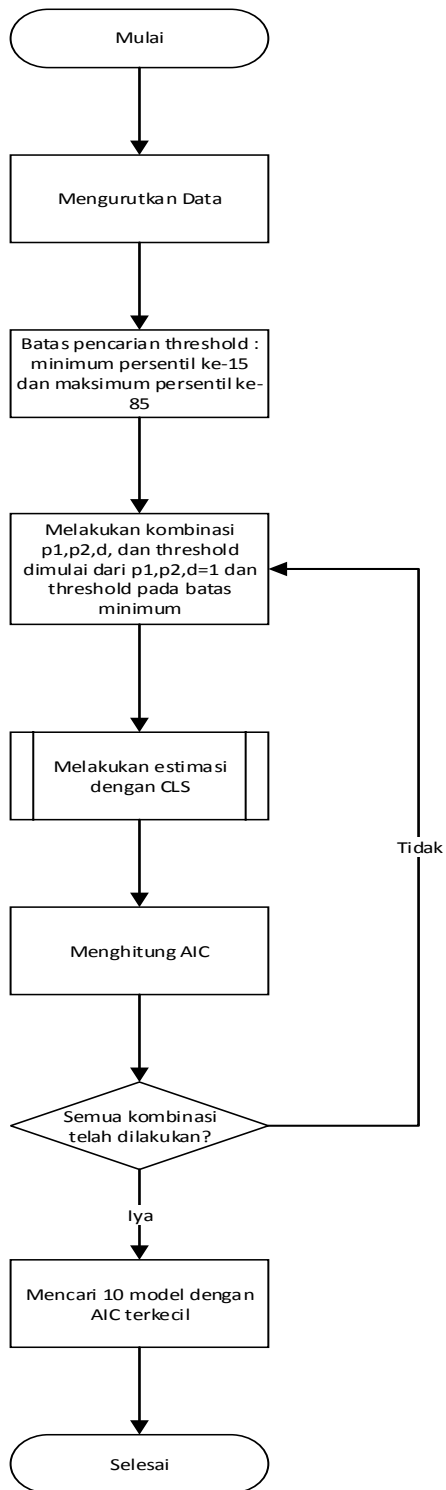
Gambar 3.3 Flow Chart Program Secara Umum



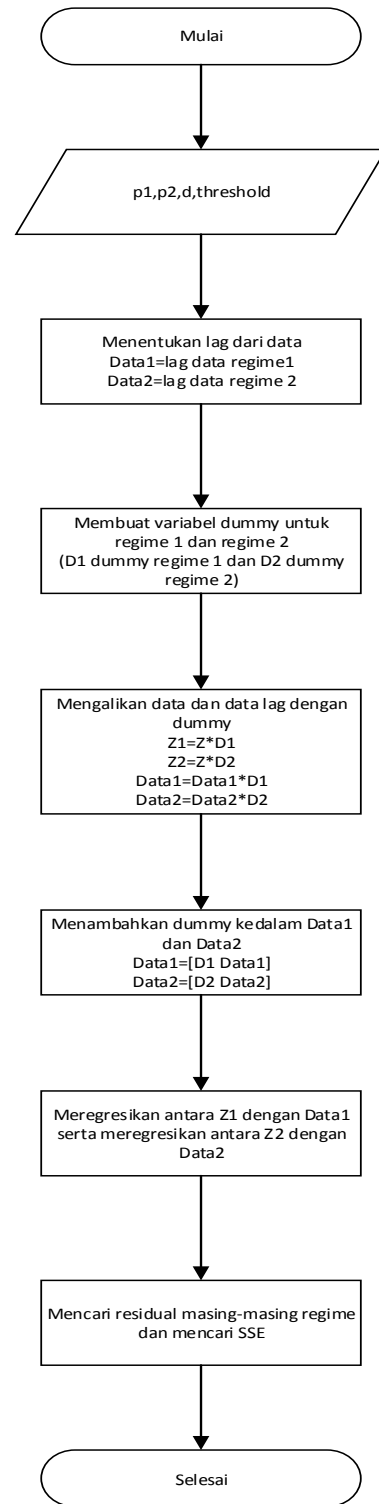
Gambar 3.4 Flow Chart Sub Program Pemodelan SETAR



Gambar 3.5 Flow Chart Sub Program Identifikasi dengan GA

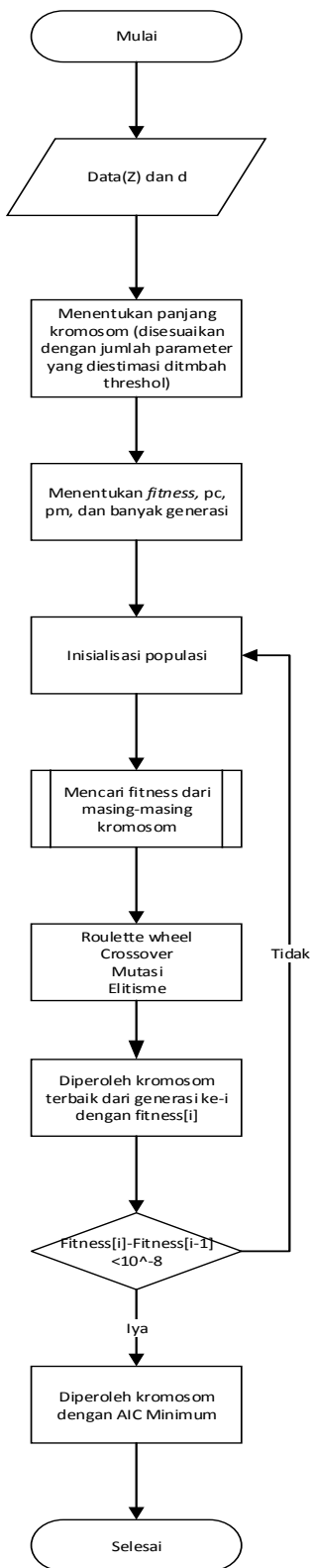


Gambar 3.6 Flow Chart Sub Program Identifikasi dengan GS

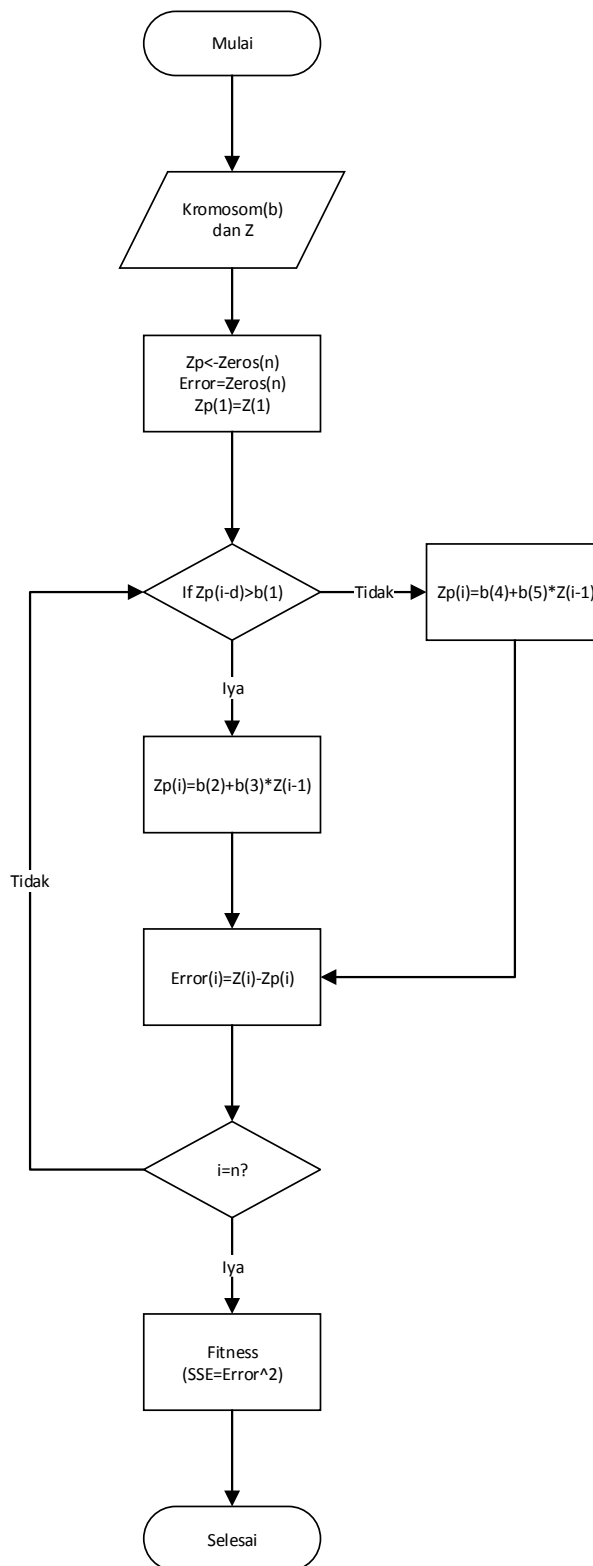


Gambar 3.7 Flow Chart Sub Program Identifikasi dengan GA

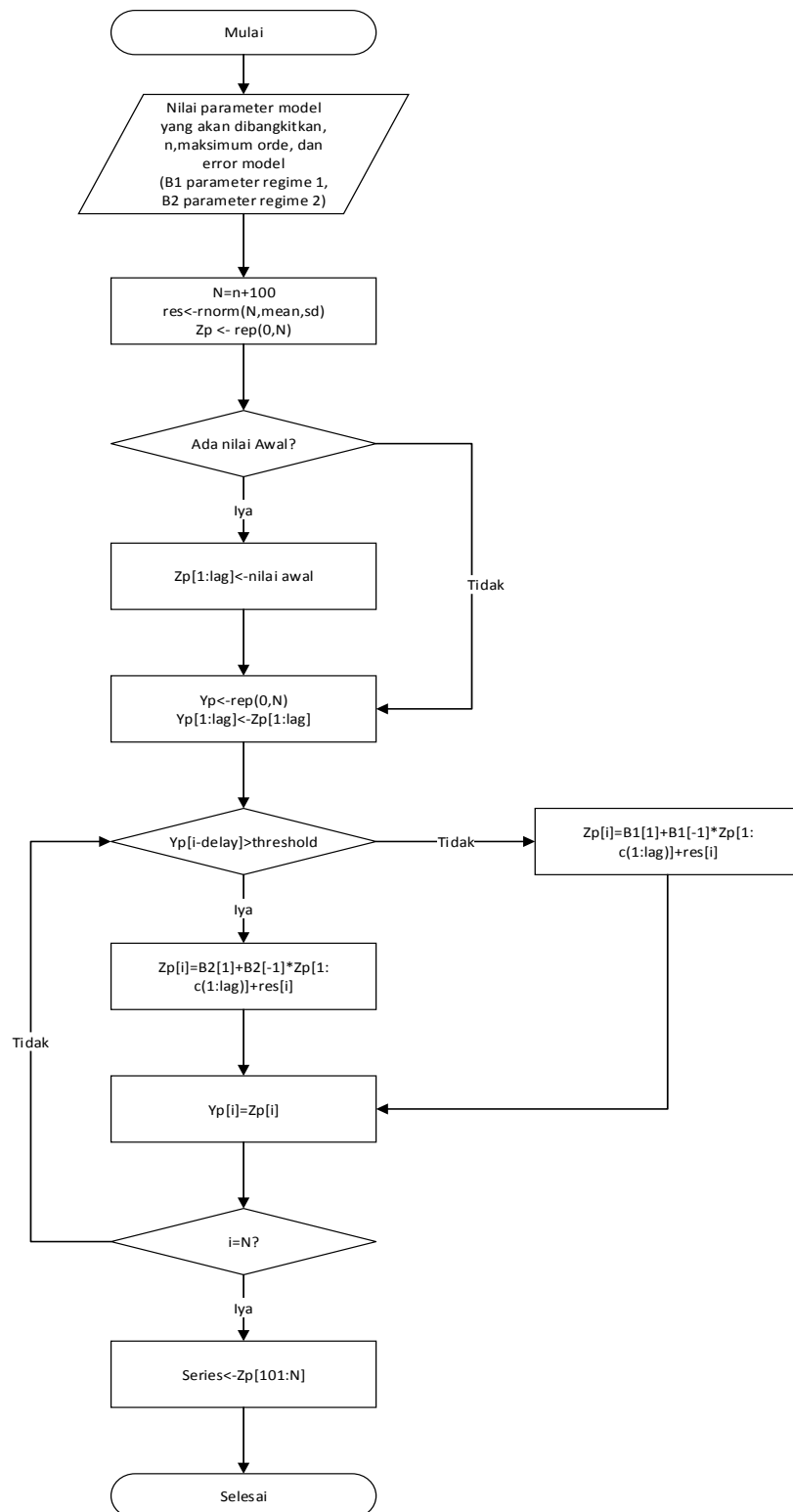




Gambar 3.8 Flow Chart Sub Program Estimasi dengan GA



Gambar 3.9 Flow Chart Sub Program Pencarian fitness pada estimasi GA



Gambar 3.10 Flow Chart Sub Program Simulasi model SETAR

Halaman ini sengaja dikosongkan

## BAB 4

### HASIL DAN PEMBAHASAN

#### 4.1 Studi Simulasi

Studi simulasi dilakukan dengan membangkitkan 2 model SETAR yaitu model SETAR(2,1,1) dan model subset SETAR. Model SETAR(2,1,1) adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti AR(1). Pada model ini, delay yang digunakan adalah 1. Masing-masing model dibangkitkan dengan jumlah data 200, 500, serta 1000.

Model subset SETAR adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti subset dari model AR(3). Model ini dibangkitkan dengan menggunakan delay 1 dan 2 dengan jumlah data 200 dan 500.

##### 4.1.1 Model SETAR(2,1,1)

Studi Simulasi model SETAR dilakukan dengan membangkitkan 5 model SETAR(2,1,1) yaitu model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model AR(1). Masing-masing model dibangkitkan dengan  $n=200, 500, \text{ dan } 1000$ . Delay yang digunakan pada 5 model tersebut adalah 1.  $a_t$  dibangkitkan mengikuti distribusi normal dengan rata-rata 0 dan standar deviasi 2. Parameter-parameter yang digunakan pada simulasi model SETAR(2,1,1) disajikan pada Tabel 4.1.

Tabel 4.1 Parameter Model Simulasi SETAR(2,1,1)

Model	$\varphi_{0,1}$	$\varphi_{1,1}$	$\varphi_{0,2}$	$\varphi_{1,2}$	Threshold
1	8	-0,2	8	0,2	8
2	1	0,2	4	0,2	3
3	5	0,5	9	-0,5	6
4	8	-0,2	9	-0,8	6
5.	-7	-0,5	-2	0,8	-2,5

Berdasarkan Tabel 4.1 terlihat bahwa model SETAR(2,1,1) dibangkitkan dengan menggunakan nilai koefisien yang berbeda. Tujuannya adalah untuk mengetahui pola model yang terbentuk berdasarkan nilai parameter yang ditentukan.

Pemodelan data dengan model SETAR dilakukan dengan beberapa tahapan. Tahapan pertama adalah melakukan uji nonlinieritas Terasvirta dan White. Kedua uji ini dilakukan sebagai syarat suatu data dapat dianalisis dengan model deret waktu nonlinier SETAR.

Sebelum melakukan analisis data simulasi model SETAR(2,1,1), data simulasi digambarkan dalam bentuk statistik diskriptif. Tabel 4.2 menunjukkan hasil diskriptif statistik dari masing-masing model simulasi SETAR(2,1,1). Hasil yang diperoleh menunjukkan bahwa data simulasi memiliki rata-rata di sekitar nilai *threshold* yang ditetapkan pada simulasi model dengan standar deviasi berkisar antara 2,2 hingga 2,7.

Tabel 4.2 Statistik Diskriptif Data Simulasi Model SETAR(2,1,1)

<b>Model</b>	<b>N</b>	<b>Min</b>	<b>Max</b>	<b>Mean</b>	<b>Sd deviasi</b>
1	200	1,958137	16,70481	8,943063	2,601472
	500	2,434098	15,4678	9,120108	2,560442
	1000	0,5548727	17,57465	8,858463	2,660521
2	200	-3,615723	10,42311	3,398611	2,79411
	500	-5,380371	10,86767	3,167653	2,734227
	1000	-4,877442	10,60689	3,533353	2,726566
3	200	1,143064	12,59017	6,020875	2,327357
	500	0,0456608	12,61189	6,12523	2,404263
	1000	0,1228242	12,69044	6,087188	2,344977
4	200	0,6395118	11,38772	6,159697	2,361783
	500	-0,067396	12,43845	6,311662	2,407534
	1000	-2,449929	14,34716	6,013438	2,522213
5	200	-9,759219	3,436348	-3,48608	2,616065
	500	-10,28239	3,183239	-3,79766	2,351326
	100	4,38398	-11,43161	-4,16511	2,257208

Rangkuman hasil uji nonlinieritas Terasvirta dan White untuk data simulasi model 1 sampai model 5 ditunjukkan pada Tabel 4.3. Berdasarkan hasil yang ditunjukkan pada Tabel 4.3, data simulasi model 1 sampai 5 memenuhi sifat nonlinieritas data karena nilai  $F_{hitung}$  pada masing-masing data lebih dari 3,04175 atau semua nilai p value lebih kecil dari  $\alpha$  (0,05).

Tabel 4.3 Uji Nonlinieritas Terasvirta dan White Simulasi SETAR (2,1,1)

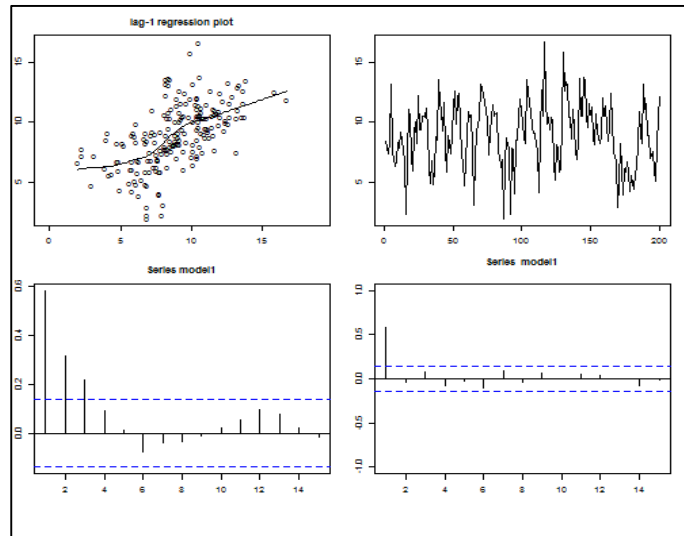
Model	N	Terasvirta		White	
		F	p-value	F	p-value
1	200	4,0325	0,01921	4,1119	0,0178
	500	11,0068	2,102e-05	16,3221	1,363e-07
	1000	34,5203	3,22e-15	35,9145	8,882e-16
2	200	9,5186	0,0001132	10,9055	3,222e-05
	500	13,3473	2,258e-06	20,4458	2,93e-09
	1000	27,2266	3,071e-12	37,8345	2,22e-16
3	200	12,7506	6,204e-06	11,2204	2,427e-05
	500	32,3544	6,173e-14	27,5489	4,504e-12
	1000	75,2898	2,2e-16	79,4572	2,2e-16
4	200	20,8003	6,372e-09	23,66	6,18e-10
	500	43,8475	2,2e-16	52,8242	2,2e-16
	1000	78,0555	2,2e-16	136,5835	2,2e-16
5	200	21,3183	4,159e-09	25,3138	1,643e-10
	500	60,0387	2,2e-16	67,7437	2,2e-16
	100	97,9692	2,2e-16	109,5513	2,2e-16

Pada studi simulasi model SETAR(2,1,1), identifikasi model dilakukan dengan menggunakan metode *grid search* (GS) pada *software R package tsDyn* dan metode algoritma genetika (GA). Untuk estimasi parameter dilakukan metode *conditional least square* (CLS) dan metode algoritma genetika (GA).

Dalam melakukan pemodelan dengan metode GA perlu didefinisikan *fitness* yang digunakan. *Fitness* yang digunakan sangat mempengaruhi hasil yang akan diperoleh. Selain *fitness*, parameter-parameter seperti peluang crossover, mutasi, banyak generasi, dan panjang kromosom yang digunakan juga harus ditentukan. Nilai dari masing-masing parameter yang digunakan pada pemodelan dengan GA dapat dilihat pada Tabel 3.3.

#### 4.1.1.1 Model 1 SETAR(2,1,1)

Model 1 adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model AR(1). Pada model 1, konstanta *regime* 1 dan *regime* 2 ditentukan bernilai sama yaitu  $\phi_{0,1} = \phi_{0,2} = 8$ . *Threshold* yang digunakan adalah 8. Koefisien pada *regime* 1 adalah  $\phi_{1,1} = -0,2$  dan koefisien *regime* 2 adalah  $\phi_{1,2} = 0,2$ .



Gambar 4.1 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 1

Salah satu cara untuk mengetahui pola data simulasi model 1 yaitu menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$  seperti terlihat pada Gambar 4.1. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk cenderung nonlinier. Sehingga pemodelan data simulasi model 1 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier.

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.3. Hasil yang diperoleh menunjukkan bahwa data simulasi model 1 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan pengujian nonlinieritas, data simulasi model 1 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari plot PACF yang dibuat. Plot PACF pada Gambar 4.1 yang menunjukkan bahwa data simulasi model 1 *cuts off* pada lag 1 sehingga pendefinisian orde awal identifikasi menggunakan nilai  $d = 1$ ,  $p_1 = 1$ , dan  $p_2 = 1$ . Identifikasi model dilakukan dengan menggunakan 2 metode yaitu GS dan GA.

Tabel 4.4 Identifikasi Data Simulasi Model 1 dengan GS

<b>N</b>	<b>d</b>	<b>p<sub>1</sub></b>	<b>p<sub>2</sub></b>	<b>Threshold</b>	<b>AIC</b>
200	1	1	1	7,962354	267,7504
500	1	1	1	7,891665	730,1770
1000	1	1	1	7,999871	1452,933

Tabel 4.5 Identifikasi Data Simulasi Model 1 dengan GA

<b>N</b>	<b>Region</b>	<b>d</b>	<b>p<sub>1</sub></b>	<b>p<sub>2</sub></b>	<b>Threshold</b>	<b>AIC</b>
200	5%	1	1	1	8,050797	267,7504
	10%	1	1	1	8,011055	267,7504
	15%	1	1	1	8,036657	267,7504
	20%	1	1	1	8,058372	267,7504
500	5%	1	1	1	7,898805	730,1770
	10%	1	1	1	7,899268	730,1770
	15%	1	1	1	7,897378	730,1770
	20%	1	1	1	7,899035	730,1770
1000	5%	1	1	1	8,007837	1452,933
	10%	1	1	1	8,007653	1452,933
	15%	1	1	1	8,003181	1452,933
	20%	1	1	1	8,008440	1452,933

Tabel 4.4 menunjukkan hasil identifikasi data simulasi model 1 dengan menggunakan metode GS. Kriteria yang digunakan untuk menentukan model terbaik pada metode tersebut adalah meminimumkan AIC. Hasil yang diperoleh menunjukkan bahwa model yang terbentuk untuk  $n=200$  adalah model SETAR(2,1,1) dengan  $threshold=7,962354$ . Untuk  $n=500$  model yang terbentuk adalah model SETAR(2,1,1) dengan  $threshold=7,891665$ . Sedangkan untuk  $n=1000$  model yang terbentuk adalah model SETAR(2,1,1) dengan  $threshold=7,999871$ . Ketiga model yang terbentuk adalah model yang meminimumkan AIC.

Tabel 4.5 menunjukkan hasil identifikasi data simulasi model 1 dengan metode GA. Pada tabel tersebut dapat dilihat bahwa pencarian nilai  $threshold$  pada metode GA dengan menggunakan 4 daerah pencarian  $threshold$  yaitu 5% – 95% , 10% – 90%, 15% – 85%, dan 20% – 80%. Kriteria untuk mendapatkan model terbaik juga dilakukan dengan meminimumkan AIC. Hasil identifikasi model yang diperoleh dari 4 daerah pencarian  $threshold$  memberikan hasil AIC yang sama



dikarenakan nilai *threshold* dari data simulasi model 1 berada disekitar nilai rata-rata.

Dari Tabel 4.4 dan Tabel 4.5 diketahui bahwa hasil identifikasi data simulasi model 1 dengan metode GS dan GA memberikan hasil AIC yang sama dan diperoleh model SETAR(2,1,1). Perbedaan antara kedua metode ini terletak pada nilai *threshold* yang diperoleh. Perbedaan *threshold* terjadi karena metode GA menggunakan semua nilai antara selang pencarian untuk mendapatkan nilai *threshold*. Sedangkan pada metode GS, *threshold* diperoleh berdasarkan data deret waktu itu sendiri. Hasil estimasi parameter menggunakan metode CLS dan GA untuk n=200, 500, dan 1000 dapat dilihat pada Tabel 4.6, Tabel 4.7, dan Tabel 4.8.

Tabel 4.6 Estimasi Parameter Data Simulasi Model 1 Untuk n=200

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	7,342047	0,657954	7,342037	0,657963
$\phi_{1,1}$	-0,120399	0,079601	-0,120397	0,079603
$\phi_{0,2}$	6,920649	1,079351	6,920656	1,079344
$\phi_{1,2}$	0,314627	0,114627	0,314627	0,114627
<i>Threshold</i>	7,962354	0,037646	8,036657	0,036657
SSE	725,65124930		725,65124923	

Berdasarkan hasil pada Tabel 4.6 diketahui bahwa estimasi parameter data simulasi model 1 untuk n=200 memberikan beberapa hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* dan hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS. Secara matematis, model terbaik dari data simulasi model 1 untuk n=200 dapat dituliskan seperti pada persamaan (4.1) berikut.

$$Z_t = \begin{cases} 7,342037 - 0,120397Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 8,036657 \\ 6,920656 + 0,314627Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 8,036657 \end{cases} \quad (2.28)$$

Model terbaik pada persamaan (2.28) menunjukkan bahwa data simulasi model 1 untuk n=200 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 8,036657$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 8,036657$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 7,342037 dan data pada pengamatan sebelumnya sebesar -0,120397 sedangkan untuk *forecast* data pada *regime* 2

dipengaruhi oleh konstanta sebesar 6,920656 dan data pada pengamatan sebelumnya sebesar 0,314627.

Berdasarkan hasil pada Tabel 4.7 diketahui bahwa estimasi parameter data simulasi model 1 untuk  $n=500$  memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Pada data simulasi model 1 untuk  $n=500$ , hasil estimasi parameter yang memberikan hasil SSE paling kecil adalah hasil estimasi dengan metode GA.

Tabel 4.7 Estimasi Parameter Data Simulasi Model 1 Untuk  $n=500$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	8,949545	0,949545	8,949559	0,949559
$\phi_{1,1}$	-0,345643	0,145643	-0,345645	0,145645
$\phi_{0,2}$	7,717369	0,282631	7,717362	0,282638
$\phi_{1,2}$	0,224225	0,024225	0,224226	0,024226
<i>Threshold</i>	7,892000	0,108000	7,897378	0,102622
SSE	2111,09511350		2111,09511338	

Secara matematis, model terbaik dari data simulasi model 1 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.2) berikut.

$$Z_t = \begin{cases} 8,949559 - 0,345645Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 7,897378 \\ 7,717362 + 0,224226Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 7,897378 \end{cases} \quad (2.29)$$

Model terbaik pada persamaan (2.29) menunjukkan bahwa data simulasi model 1 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 7,897378$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 7,897378$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 8,949559 dan data pada pengamatan sebelumnya sebesar -0,345645 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 7,717362 dan data pada pengamatan sebelumnya sebesar 0,224226.

Tabel 4.8 Estimasi Parameter Data Simulasi Model 1 Untuk n=1000

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	8,277040	0,277040	8,277041	0,277041
$\phi_{1,1}$	-0,262240	0,062240	-0,262240	0,062240
$\phi_{0,2}$	8,223179	0,223179	8,223177	0,223177
$\phi_{1,2}$	0,180244	0,019756	0,180244	0,019756
<i>Threshold</i>	8,000000	0,000000	8,003181	0,003181
SSE	4233,09381710		4233,09381700	

Berdasarkan hasil pada Tabel 4.8 diketahui bahwa estimasi parameter data simulasi model 1 untuk n=1000 memberikan beberapa hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Pada data simulasi model 1 untuk n=1000, hasil estimasi parameter yang memberikan hasil SSE lebih kecil adalah hasil estimasi dengan metode GA.

Secara matematis, model terbaik dari data simulasi model 1 untuk n=1000 dapat dituliskan seperti pada persamaan (4.3) berikut.

$$Z_t = \begin{cases} 8,277041 - 0,262240Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 8,003181 \\ 8,223177 + 0,180244Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 8,003181 \end{cases} \quad (2.30)$$

Model terbaik pada persamaan (4.3) menunjukkan bahwa data simulasi model 1 untuk n=500 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 8,003181$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 8,003181$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 8,949559 dan data pada pengamatan sebelumnya sebesar -0,345645 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 7,717362 dan data pada pengamatan sebelumnya sebesar 0,224226.

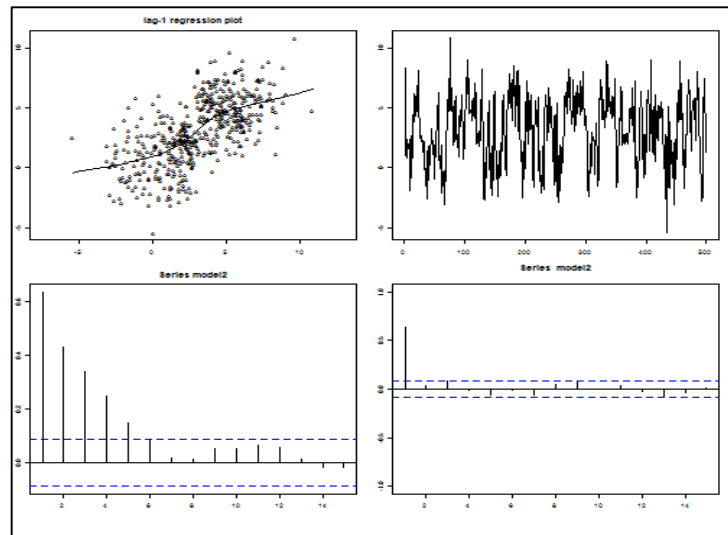
Berdasarkan hasil estimasi parameter model dari data simulasi 1 untuk n=200, 500, dan 1000 dapat disimpulkan bahwa estimasi parameter dengan metode GA memberikan hasil SSE lebih kecil dibandingkan metode CLS.

#### 4.1.1.2 Model 2 SETAR(2,1,1)

Model 2 adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model AR(1). Pada model 2 koefisien *regime* 1 dan *regime* 2 ditentukan

bernilai sama yaitu  $\phi_{1,1} = \phi_{1,2} = 0,2$ . *Threshold* yang digunakan adalah 3 dengan konstanta pada *regime 1* adalah  $\phi_{0,1} = 1$  dan konstanta *regime 2* adalah  $\phi_{0,2} = 4$ .

Salah satu cara untuk mengetahui pola data simulasi model 2 dapat dilakukan dengan menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$  seperti terlihat pada Gambar 4.2. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk cenderung nonlinier. Pola yang terbentuk terlihat seperti 2 garis linier yang terlihat sejajar apabila dipotong pada garis  $Z_{t-1} = 3$  yang merupakan nilai *threshold* pada simulasi data model 2. Berdasarkan uraian diatas, pemodelan dari data simulasi model 2 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier.



Gambar 4.2 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 2

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ , untuk melihat pola dari data simulasi model 2 dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.3. Hasil yang diperoleh menunjukkan bahwa data simulasi model 2 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan pengujian nonlinieritas, selanjutnya data simulasi model 2 dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari

plot PACF yang dibuat. Plot PACF pada Gambar 4.2 menunjukkan bahwa data simulasi model 2 *cuts off* pada lag 1 sehingga pendefinisian orde awal identifikasi menggunakan nilai  $d = 1, p_1 = 1, \text{ dan } p_2 = 1$ . Identifikasi model dilakukan dengan menggunakan metode yaitu *GS* dan *GA*.

Tabel 4.9 Identifikasi Data Simulasi Model 2 dengan *GS*

<b>N</b>	<b>d</b>	<b><math>p_1</math></b>	<b><math>p_2</math></b>	<b><i>Threshold</i></b>	<b>AIC</b>
200	1	1	1	2,900298	250,6056
500	1	1	1	3,051782	672,6851
1000	1	1	1	3,010300	1322,742

Tabel 4.10 Identifikasi Data Simulasi Model 2 dengan *GA*

<b>N</b>	<b>Region</b>	<b>d</b>	<b><math>p_1</math></b>	<b><math>p_2</math></b>	<b><i>Threshold</i></b>	<b>AIC</b>
200	5%	1	1	1	2,908943	250,6056
	10%	1	1	1	2,966442	250,6056
	15%	1	1	1	3,000248	250,6056
	20%	1	1	1	2,994548	250,6056
500	5%	1	1	1	3,057123	672,6851
	10%	1	1	1	3,054263	672,6851
	15%	1	1	1	3,052453	672,6851
	20%	1	1	1	3,055585	672,6851
1000	5%	1	1	1	3,014561	1322,742
	10%	1	1	1	3,011878	1322,742
	15%	1	1	1	3,013241	1322,742
	20%	1	1	1	3,013339	1322,742

Tabel 4.9 menunjukkan hasil identifikasi data simulasi model 2 dengan menggunakan metode *GS*. Kriteria yang digunakan untuk menentukan model terbaik dengan menggunakan minimum AIC. Hasil yang diperoleh menunjukkan bahwa model yang terbentuk untuk  $n=200$  adalah model SETAR(2,1,1) dengan  $threshold=2,900298$ . Untuk  $n=500$  model yang terbentuk adalah model SETAR(2,1,1) dengan  $threshold=3,051782$ . Sedangkan untuk  $n=1000$  model yang terbentuk adalah model SETAR(2,1,1) dengan  $threshold=3,010300$ . Ketiga model yang terbentuk adalah model yang meminimumkan AIC.

Tabel 4.10 menunjukkan hasil identifikasi data simulasi model 2 dengan metode *GA*. Pada tabel tersebut dapat dilihat bahwa pencarian nilai *threshold* pada metode *GA* dengan menggunakan 4 daerah pencarian *threshold* yaitu 5% – 95% ,

10% – 90%, 15% – 85%, dan 20% – 80%. Kriteria untuk mendapatkan model terbaik dilakukan dengan meminimumkan AIC. Hasil identifikasi model yang diperoleh dari 4 daerah pencarian *threshold* memberikan hasil yang sama karena nilai *threshold* yang disimulasikan berada disekitar nilai rata-rata.

Dari Tabel 4.9 dan Tabel 4.10 diketahui bahwa hasil identifikasi data simulasi model 2 dengan metode GS dan GA memberikan hasil yang sama yaitu model SETAR(2,1,1). Perbedaan antara kedua metode ini terletak pada nilai *threshold* yang diperoleh. Perbedaan nilai *threshold* terjadi karena metode GA menggunakan semua nilai antara selang pencarian untuk mendapatkan nilai *threshold*. Sedangkan pada metode GS, *threshold* diperoleh berdasarkan data deret waktu itu sendiri.

Hasil estimasi parameter menggunakan metode CLS dan GA untuk n=200, 500, dan 1000 dapat dilihat pada Tabel 4.11, Tabel 4.12, dan Tabel 4.13.

Tabel 4.11 Estimasi Parameter Data Simulasi Model 2 untuk n=200

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	1,044930	0,044930	1,044931	0,044931
$\phi_{1,1}$	0,133230	0,066770	0,133225	0,066775
$\phi_{0,2}$	3,117750	0,882250	3,117750	0,882250
$\phi_{1,2}$	0,389860	0,189860	0,389864	0,189864
<i>Threshold</i>	2,900000	0,100000	3,000248	0,000248
SSE	666,03714162		666,03714160	

Berdasarkan hasil pada Tabel 4.11 diketahui bahwa estimasi parameter data simulasi model 2 untuk n=200 memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.

Secara matematis, model terbaik dari data simulasi model 2 untuk n=200 dapat dituliskan seperti pada persamaan (4.4) berikut.

$$Z_t = \begin{cases} 1,044931 + 0,133225Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 3,000248 \\ 3,117750 + 0,389864Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 3,000248 \end{cases} \quad (2.31)$$

Model terbaik pada persamaan (4.4) menunjukkan bahwa data simulasi model 2 untuk n=200 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 3,000248$  dan

masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 3,000248$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 1,044931 dan data pada pengamatan sebelumnya sebesar 0,133225 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 3,117750 dan data pada pengamatan sebelumnya sebesar 0,389864.

Berdasarkan hasil pada Tabel 4.12 diketahui bahwa estimasi parameter data simulasi model 2 untuk  $n=500$  memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.

Tabel 4.12 Estimasi Parameter Data Simulasi Model 2 untuk  $n=500$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	0,994473	0,005527	0,994472	0,005528
$\phi_{1,1}$	0,236040	0,036040	0,236040	0,036040
$\phi_{0,2}$	3,955021	0,044979	3,955016	0,044984
$\phi_{1,2}$	0,191692	0,008308	0,191693	0,008307
<i>Threshold</i>	3,052000	0,052000	3,053453	0,053453
SSE	1881,78924254		1881,78924259	

Secara matematis, model terbaik dari data simulasi model 2 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.5) berikut.

$$Z_t = \begin{cases} 0,994473 + 0,236040Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 3,052000 \\ 3,955021 + 0,191692Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 3,052000 \end{cases} \quad (2.32)$$

Model terbaik pada persamaan (4.5) menunjukkan bahwa data simulasi model 2 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 3,052000$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 3,052000$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 0,994473 dan data pada pengamatan sebelumnya sebesar 0,236040 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 3,955021 dan data pada pengamatan sebelumnya sebesar 0,191692.

Tabel 4.13 Estimasi Parameter Data Simulasi Model 2 untuk n=1000

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	1,016022	0,016022	1,016022	0,016022
$\phi_{1,1}$	0,211835	0,011835	0,211835	0,011835
$\phi_{0,2}$	4,233147	0,233147	4,233147	0,233147
$\phi_{1,2}$	0,157656	0,042344	0,157656	0,042344
<i>Threshold</i>	3,010000	0,010000	3,013241	0,013241
SSE	3716,35003382		3716,35003394	

Berdasarkan hasil pada Tabel 4.13 diketahui bahwa estimasi parameter data simulasi model 2 untuk n=1000 memberikan beberapa hasil yang sama ketika estimasi parameter dilakukan menggunakan metode CLS dan GA. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.

Secara matematis, model terbaik dari data simulasi model 2 untuk n=1000 dapat dituliskan seperti pada persamaan (4.6) berikut.

$$Z_t = \begin{cases} 1,016022 + 0,211835Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 3,010000 \\ 4,233147 + 0,157656Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 3,010000 \end{cases} \quad (2.33)$$

Model terbaik pada persamaan (4.6) menunjukkan bahwa data simulasi model 2 untuk n=1000 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 3,010000$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 3,010000$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 1,016022 dan data pada pengamatan sebelumnya sebesar 0,211835 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 4,233147 dan data pada pengamatan sebelumnya sebesar 0,157656.

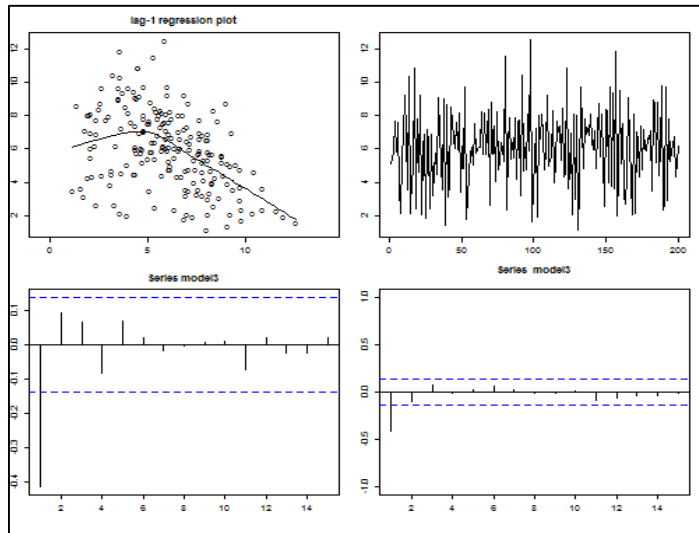
Berdasarkan hasil estimasi parameter pada data simulasi 2 untuk n=200, 500, dan 1000 dapat disimpulkan bahwa estimasi parameter dengan metode CLS memberikan hasil SSE lebih kecil dibandingkan metode GA.

#### 4.1.1.3 Model 3 SETAR(2,1,1)

Model 3 adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model AR(1). Pada model 3 koefisien masing-masing *regime* yaitu



$\phi_{1,1} = 0,5$  dan  $\phi_{1,2} = -0,5$ . Untuk nilai konstanta *regime* 1 dan *regime* 2 berturut-turut adalah  $\phi_{0,1} = 5, \phi_{0,2} = 9$  dengan *threshold* yang digunakan pada simulasi model 3 adalah 6.



Gambar 4.3 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 3

Salah satu cara untuk mengetahui pola data simulasi model 3 dengan menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$  seperti terlihat pada Gambar 4.3. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk cenderung nonlinier. Pola yang dihasilkan terlihat seperti bentuk kuadrat yang menghadap kebawah. Apabila dipotong pada garis  $Z_{t-1} = 6$ , maka diperoleh 2 buah garis linier yang berbeda arah. Garis yang terbentuk pada sebelah kiri garis  $Z_{t-1} = 6$  merupakan garis linier yang cenderung naik. Sedangkan garis yang terbentuk pada sebelah kanan garis  $Z_{t-1} = 6$  merupakan garis linier yang cenderung turun. Perbedaan arah pada kedua garis linier yang diperoleh dikarenakan adanya perbedaan tanda pada koefisien parameter pada model yang disimulasi. Berdasarkan uraian diatas, pemodelan data simulasi model 3 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier.

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji dapat dilihat pada Tabel 4.3. Hasil yang diperoleh menunjukkan bahwa data simulasi model 3 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan

pengujian nonlinieritas, data simulasi model 3 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum ditentukan berdasarkan plot PACF yang dibuat. Plot PACF pada Gambar 4.3 yang menunjukkan bahwa data simulasi model 3 *cuts off* pada lag 1 sehingga pendefinisian orde awal identifikasi menggunakan nilai  $d = 1, p_1 = 1, \text{ dan } p_2 = 1$ . Identifikasi model terbaik menggunakan metode GS dan GA. Hasil Identifikasi model yang dilakukan dengan metode GS dan GA dapat dilihat pada Tabel 4.14 dan Tabel 4.15.

Tabel 4.14 Identifikasi Data Simulasi Model 3 dengan GS

<b>N</b>	<b>d</b>	<b><math>p_1</math></b>	<b><math>p_2</math></b>	<b><i>Threshold</i></b>	<b>AIC</b>
200	1	1	1	6,048400	280,8375
500	1	1	1	5,716954	713,5152
1000	1	1	1	5,987987	1380,374

Tabel 4.15 Identifikasi Data Simulasi Model 3 dengan GA

<b>N</b>	<b>Region</b>	<b>d</b>	<b><math>p_1</math></b>	<b><math>p_2</math></b>	<b><i>Threshold</i></b>	<b>AIC</b>
200	5%	1	1	1	6,070048	280,8375
	10%	1	1	1	6,073964	280,8375
	15%	1	1	1	6,069044	280,8375
	20%	1	1	1	6,074092	280,8375
500	5%	1	1	1	5,731187	713,5152
	10%	1	1	1	5,735990	713,5152
	15%	1	1	1	5,726243	713,5152
	20%	1	1	1	5,736511	713,5152
1000	5%	1	1	1	5,999814	1380,374
	10%	1	1	1	5,997975	1380,374
	15%	1	1	1	6,000833	1380,374
	20%	1	1	1	5,997396	1380,374

Tabel 4.14 menunjukkan hasil identifikasi data simulasi model 3 dengan menggunakan metode GS. Kriteria yang digunakan untuk menentukan model terbaik yaitu menggunakan minimum AIC. Hasil yang diperoleh menunjukkan bahwa model yang terbentuk untuk  $n=200$  adalah model SETAR(2,1,1) dengan  $threshold=6,048400$ . Untuk  $n=500$  model yang terbentuk adalah model

SETAR(2,1,1) dengan  $threshold=5,716954$ . Sedangkan untuk  $n=1000$  model yang terbentuk adalah model SETAR(2,1,1) dengan  $threshold=5,987987$ . Ketiga model yang terbentuk adalah model yang meminimumkan AIC.

Tabel 4.15 menunjukkan hasil identifikasi data simulasi model 3 dengan metode GA. Pada tabel tersebut dapat dilihat bahwa pencarian nilai  $threshold$  pada metode GA dengan menggunakan 4 daerah pencarian  $threshold$  yaitu 5% – 95% , 10% – 90%, 15% – 85%, dan 20% – 80%. Kriteria untuk mendapatkan model terbaik dilakukan dengan meminimumkan AIC. Hasil identifikasi model yang diperoleh dari 4 daerah pencarian  $threshold$  memberikan hasil yang sama. Hal ini dikarenakan nilai  $threshold$  yang disimulasikan berada disekitar nilai rata-rata.

Dari Tabel 4.14 dan Tabel 4.15 diketahui bahwa hasil identifikasi data simulasi model 3 dengan metode GS dan GA memberikan hasil yang sama yaitu model SETAR(2,1,1). Perbedaan antara kedua metode ini terletak pada nilai  $threshold$  yang diperoleh. Perbedaan  $threshold$  terjadi karena metode GA menggunakan semua nilai antara selang pencarian untuk mendapatkan nilai  $threshold$ . Sedangkan pada metode GS,  $threshold$  diperoleh berdasarkan data deret waktu itu sendiri. Hasil estimasi parameter menggunakan metode CLS dan GA untuk  $n=200, 500, \text{ dan } 1000$  dapat dilihat pada Tabel 4.16, Tabel 4.17, dan Tabel 4.18.

Tabel 4.16 Estimasi Parameter Data Simulasi Model 3 untuk  $n=200$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	5,767840	0,767840	5,767836	0,767836
$\phi_{1,1}$	0,303670	0,196330	0,303673	0,196327
$\phi_{0,2}$	9,958420	0,958420	9,958412	0,958412
$\phi_{1,2}$	-0,627610	0,127610	-0,627614	0,127614
<i>Threshold</i>	6,048400	0,048400	6,069044	0,069044
SSE	774,72256650		774,72256637	

Berdasarkan hasil Tabel 4.16 diketahui bahwa estimasi parameter data simulasi model 3 untuk  $n=200$  memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada  $threshold$  serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.

Secara matematis, model terbaik dari data simulasi model 3 untuk  $n=200$  dapat dituliskan seperti pada persamaan (4.7) berikut.

$$Z_t = \begin{cases} 5,767836 + 0,303673Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 6,069044 \\ 9,958412 - 0,627614Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 6,069044 \end{cases} \quad (2.34)$$

Model terbaik pada persamaan (4.7) menunjukkan bahwa data simulasi model 3 untuk  $n=200$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 6,069044$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 6,069044$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 5,767836 dan data pada pengamatan sebelumnya sebesar 0,303673 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 9,958412 dan data pada pengamatan sebelumnya sebesar -0,627614.

Berdasarkan hasil pada Tabel 4.17 diketahui bahwa estimasi parameter data simulasi model 3 untuk  $n=500$  memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.

Tabel 4.17 Estimasi Parameter Data Simulasi Model 3 untuk  $n=500$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	5,064693	0,064693	5,064694	0,064694
$\phi_{1,1}$	0,567662	0,067662	0,567659	0,067659
$\phi_{0,2}$	8,702868	0,297132	8,702872	0,297128
$\phi_{1,2}$	-0,451879	0,048121	-0,451880	0,048120
<i>Threshold</i>	5,716954	0,283046	5,726243	0,273757
SSE	2041,9052191		2041,9052190	

Secara matematis, model terbaik dari data simulasi model 3 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.8) berikut.

$$Z_t = \begin{cases} 5,064694 + 0,567659Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 5,726243 \\ 8,702872 - 0,451880Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 5,726243 \end{cases} \quad (2.35)$$

Model terbaik pada persamaan (4.8) menunjukkan bahwa data simulasi model 3 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 5,726243$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 5,726243$ . Nilai *forecast* data pada

*regime* 1 dipengaruhi oleh konstanta sebesar 5,064694 dan data pada pengamatan sebelumnya sebesar 0,567659 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 8,702872 dan data pada pengamatan sebelumnya sebesar -0,451880.

Tabel 4.18 Estimasi Parameter Data Simulasi Model 3 untuk n=1000

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	4,706500	0,293500	4,706504	0,293496
$\phi_{1,1}$	0,585830	0,085830	0,585829	0,085829
$\phi_{0,2}$	8,725430	0,274570	8,725430	0,274570
$\phi_{1,2}$	-0,457120	0,042880	-0,457123	0,042877
<i>Threshold</i>	5,987987	0,012013	6,000833	0,000833
SSE	3936,82329240		3936,82329245	

Berdasarkan hasil pada Tabel 4.18 diketahui bahwa estimasi parameter data simulasi model 3 untuk n=1000 memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Secara matematis, model terbaik dari data simulasi model 3 untuk n=1000 dapat dituliskan seperti pada persamaan (4.9) berikut.

$$Z_t = \begin{cases} 4,706500 + 0,585830Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 5,987987 \\ 8,725430 - 0,457120Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 5,987987 \end{cases} \quad (2.36)$$

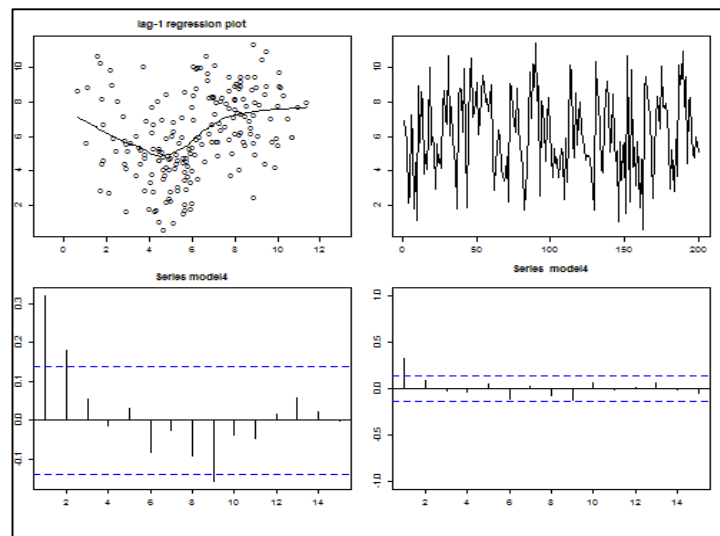
Model terbaik pada persamaan (4.9) menunjukkan bahwa data simulasi model 3 untuk n=1000 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 5,987987$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 5,987987$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 4,706500 dan data pada pengamatan sebelumnya sebesar 0,585830 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 8,725430 dan data pada pengamatan sebelumnya sebesar -0,457120.

Berdasarkan hasil estimasi parameter pada data simulasi 3 untuk n=200, 500, dan 1000 dapat disimpulkan bahwa estimasi parameter dengan metode GA memberikan hasil SSE lebih kecil dibandingkan metode CLS.

#### 4.1.1.4 Model 4 SETAR(2,1,1)

Model 4 adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model AR(1). Pada model 4 konstanta *regime* 1 adalah  $\phi_{0,1} = 8$  dan konstanta *regime* 2 adalah  $\phi_{0,2} = 9$ . *Threshold* yang digunakan pada model 4 adalah 6. Koefisien pada kedua *regime* menggunakan nilai yang negatif yaitu koefisien pada *regime* 1 adalah  $\phi_{1,1} = -0,8$  dan koefisien *regime* 2 adalah  $\phi_{1,2} = -0,2$ .

Cara yang dapat digunakan untuk mengetahui pola data simulasi model 4 yaitu menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$  seperti terlihat pada Gambar 4.4. Berdasarkan plot tersebut terlihat bahwa plot yang dihasilkan berbentuk kuadratik yang menghadap keatas. Apabila dipotong pada garis  $Z_{t-1} = 6$ , maka diperoleh 2 buah garis linier yang berbeda arah. Garis yang terbentuk pada sebelah kiri garis  $Z_{t-1} = 6$  merupakan garis linier yang cenderung turun. Sedangkan garis yang terbentuk pada sebelah kanan garis  $Z_{t-1} = 6$  merupakan garis linier yang cenderung naik. Berdasarkan uraian diatas, pemodelan data simulasi model 4 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier.



Gambar 4.4 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 4

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ , untuk melihat pola dari data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji

tersebut dapat dilihat pada Tabel 4.3. Hasil yang diperoleh menunjukkan bahwa data simulasi model 4 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan pengujian nonlinieritas, data simulasi model 4 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari plot PACF yang dibuat. Plot PACF pada Gambar 4.4 menunjukkan bahwa data simulasi model 4 *cuts off* pada lag 1 sehingga pendefinisian orde awal identifikasi menggunakan nilai  $d = 1, p_1 = 1, \text{ dan } p_2 = 1$ . Identifikasi model yang dilakukan dengan metode GS dan GA dapat dilihat pada Tabel 4.19 dan Tabel 4.20.

Tabel 4.19 Identifikasi Data Simulasi Model 4 dengan GS

N	d	$p_1$	$p_2$	Threshold	AIC
200	1	1	1	5,995423	261,6115
500	1	1	1	5,998954	675,5407
1000	1	1	1	5,977284	1381,774

Tabel 4.20 Identifikasi Data Simulasi Model 4 dengan GA

N	Region	d	$p_1$	$p_2$	Threshold	AIC
200	5%	1	1	1	6,021934	261,6115
	10%	1	1	1	6,023254	261,6115
	15%	1	1	1	6,004834	261,6115
	20%	1	1	1	6,017663	261,6115
500	5%	1	1	1	6,005689	675,5407
	10%	1	1	1	6,002331	675,5407
	15%	1	1	1	6,001313	675,5407
	20%	1	1	1	6,006257	675,5407
1000	5%	1	1	1	5,985516	1381,774
	10%	1	1	1	5,985677	1381,774
	15%	1	1	1	5,987836	1381,774
	20%	1	1	1	5,985804	1381,774

Tabel 4.19 menunjukkan hasil identifikasi data simulasi model 4 dengan menggunakan metode GS. Kriteria yang digunakan untuk menentukan model terbaik dengan meminimumkan AIC. Hasil yang diperoleh menunjukkan bahwa model yang terbentuk untuk  $n=200$  adalah model SETAR(2,1,1) dengan  $threshold=5,995423$ . Untuk  $n=500$  model yang terbentuk adalah model

SETAR(2,1,1) dengan  $threshold=5,998954$ . Sedangkan untuk  $n=1000$  model yang terbentuk adalah model SETAR(2,1,1) dengan  $threshold=5,977284$ . Ketiga model yang terbentuk adalah model yang meminimumkan AIC.

Tabel 4.20 menunjukkan hasil identifikasi data simulasi model 4 dengan metode GA. Pada tabel tersebut dapat dilihat bahwa pencarian nilai  $threshold$  pada metode GA dengan menggunakan 4 daerah pencarian  $threshold$  yaitu 5% – 95% , 10% – 90%, 15% – 85%, dan 20% – 80%. Kriteria untuk mendapatkan model terbaik dilakukan dengan meminimumkan AIC. Hasil identifikasi model yang diperoleh dari 4 daerah pencarian  $threshold$  memberikan hasil yang sama. Hal ini dikarenakan nilai  $threshold$  yang disimulasikan berada disekitar nilai rata-rata.

Dari Tabel 4.19 dan Tabel 4.20 diketahui bahwa hasil identifikasi data simulasi model 4 dengan metode GS dan GA memberikan hasil yang sama yaitu model SETAR(2,1,1). Perbedaan antara kedua metode ini terletak pada nilai  $threshold$  yang diperoleh. Perbedaan  $threshold$  terjadi karena metode GA menggunakan semua nilai antara selang pencarian untuk mendapatkan nilai  $threshold$ . Sedangkan pada metode GS,  $threshold$  diperoleh berdasarkan data deret waktu itu sendiri. Hasil estimasi parameter menggunakan metode CLS dan GA untuk  $n=200, 500$ , dan  $1000$  dapat dilihat pada Tabel 4.21, Tabel 4.22, dan Tabel 4.23.

Tabel 4.21 Estimasi Parameter Data Simulasi Model 4 untuk  $n=200$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	7,824190	0,175810	7,824189	0,175811
$\phi_{1,1}$	-0,686454	0,113546	-0,686454	0,113546
$\phi_{0,2}$	7,761859	1,238141	7,761854	1,238146
$\phi_{1,2}$	-0,039624	0,160376	-0,039623	0,160377
<i>Threshold</i>	5,995423	0,004577	6,004834	0,004834
SSE	703,71589156		703,71589155	

Berdasarkan hasil pada Tabel 4.21 diketahui bahwa estimasi parameter data simulasi model 4 untuk  $n=200$  memberikan hasil estimasi parameter yang sama. Perbedaan hasil terletak pada  $threshold$  serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.



Secara matematis, model terbaik dari data simulasi model 4 untuk  $n=200$  dapat dituliskan seperti pada persamaan (4.10) berikut.

$$Z_t = \begin{cases} 7,824189 - 0,686454Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 6,004834 \\ 7,761854 - 0,039623Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 6,004834 \end{cases} \quad (2.37)$$

Model terbaik pada persamaan (4.10) menunjukkan bahwa data simulasi model 4 untuk  $n=200$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 6,004834$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 6,004834$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 7,824189 dan data pada pengamatan sebelumnya sebesar -0,686454 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 7,761854 dan data pada pengamatan sebelumnya sebesar -0,039623.

Tabel 4.22 Estimasi Parameter Data Simulasi Model 4 untuk  $n=500$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	7,677050	0,322950	7,677065	0,322935
$\phi_{1,1}$	-0,671300	0,128700	-0,671304	0,128696
$\phi_{0,2}$	9,451638	0,451638	9,451621	0,451621
$\phi_{1,2}$	-0,258316	0,058316	-0,258314	0,058314
<i>Threshold</i>	5,998954	0,001046	6,001313	0,001313
SSE	1892,56702253		1892,56702243	

Berdasarkan hasil pada Tabel 4.22 diketahui bahwa estimasi parameter data simulasi model 4 untuk  $n=500$  memberikan hasil yang sama. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dibandingkan dengan CLS.

Secara matematis, model terbaik dari data simulasi model 4 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.11) berikut.

$$Z_t = \begin{cases} 7,677065 - 0,671304Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 6,001313 \\ 9,451621 - 0,258314Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 6,001313 \end{cases} \quad (2.38)$$

Model terbaik pada persamaan (4.11) menunjukkan bahwa data simulasi model 4 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 6,001313$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 6,001313$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 7,677065 dan data pada pengamatan

sebelumnya sebesar -0,671304 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 9,451621 dan data pada pengamatan sebelumnya sebesar -0,258314.

Tabel 4.23 Estimasi Parameter Data Simulasi Model 4 untuk n=1000

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	7,850974	0,149027	7,850975	0,149025
$\phi_{1,1}$	-0,777600	0,022401	-0,777600	0,022400
$\phi_{0,2}$	9,024953	0,024953	9,024940	0,024940
$\phi_{1,2}$	-0,211695	0,011695	-0,211693	0,011693
<i>Threshold</i>	5,977284	0,022716	5,987836	0,012164
SSE	3942,338634		3942,338634	

Berdasarkan hasil pada Tabel 4.23 diketahui bahwa estimasi parameter data simulasi model 4 untuk n=1000 memberikan beberapa hasil yang sama. Perbedaan hasil terletak pada *threshold*. Hasil SSE yang diperoleh pada metode GA sama dengan CLS.

Secara matematis, model terbaik dari data simulasi model 4 untuk n=1000 dapat dituliskan seperti pada persamaan (4.12) berikut.

$$Z_t = \begin{cases} 7,850975 - 0,777600Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq 5,987836 \\ 9,024940 - 0,211693Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > 5,987836 \end{cases} \quad (2.39)$$

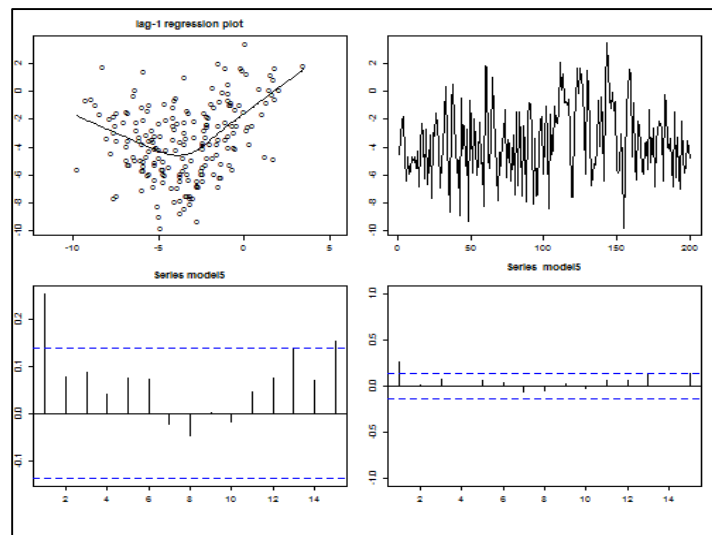
Model terbaik pada persamaan (4.12) menunjukkan bahwa data simulasi model 4 untuk n=500 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 5,987836$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 5,987836$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 7,850975 dan data pada pengamatan sebelumnya sebesar -0,777600 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 9,024940 dan data pada pengamatan sebelumnya sebesar -0,211693.

Berdasarkan hasil estimasi parameter pada data simulasi 4 untuk n=200, 500, dan 1000 dapat disimpulkan bahwa estimasi parameter dengan metode GA memberikan hasil SSE lebih kecil dibandingkan metode CLS.

#### 4.1.1.5 Model 5 SETAR(2,1,1)

Model 5 adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model AR(1). Konstanta pada *regime* 1 adalah -7, konstanta *regime* 2 adalah -2 dan *threshold* adalah -2,5. Koefisien pada kedua *regime* ditentukan nilainya berbeda serta berbeda tanda yaitu  $\phi_{1,1} = -0,5$  dan  $\phi_{1,2} = 0,8$ .

Salah satu cara untuk mengetahui pola data simulasi model 5 yaitu menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$  seperti terlihat pada Gambar 4.5. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk bukan merupakan garis linier dan cenderung nonlinier. Pola yang dihasilkan membentuk pola kuadratik yang menghadap keatas. Apabila dipotong pada garis  $Z_{t-1} = -2,5$ , maka diperoleh 2 buah garis linier yang berbeda arah. Garis yang terbentuk di sebelah kiri garis  $Z_{t-1} = -2,5$  merupakan garis linier yang cenderung turun. Sedangkan garis yang terbentuk di sebelah kanan garis  $Z_{t-1} = 2,5$  merupakan garis linier yang cenderung naik. Berdasarkan uraian diatas, pemodelan data simulasi model 5 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier.



Gambar 4.5 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model 5

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.3. Hasil yang diperoleh menunjukkan bahwa

data simulasi model 5 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan pengujian nonlinieritas, data simulasi model 5 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari plot PACF yang dibuat. Plot PACF pada Gambar 4.5 yang menunjukkan bahwa data simulasi model 5 *cuts off* pada lag 1 sehingga pendefinisian orde awal identifikasi menggunakan nilai  $d = 1$ ,  $p_1 = 1$ , dan  $p_2 = 1$ . Identifikasi model terbaik menggunakan 2 metode yaitu GS dan GA. Hasil Identifikasi model yang dilakukan dengan metode GS dan GA dapat dilihat pada Tabel 4.24 dan Tabel 4.25.

Tabel 4.24 Identifikasi Model 5 SETAR(2,1,1) dengan GS

N	d	$p_1$	$p_2$	Threshold	AIC
200	1	1	1	-2,441198	322,5045
500	1	1	1	-2,506375	709,9335
1000	1	1	1	-2,539616	1383,093

Tabel 4.25 Identifikasi Model 5 SETAR(2,1,1) dengan Algoritma Genetika

N	Region	d	$p_1$	$p_2$	Threshold	AIC
200	5%	1	1	1	-2,440793	322,5045
	10%	1	1	1	-2,440869	322,5045
	15%	1	1	1	-2,441174	322,5045
	20%	1	1	1	-2,440899	322,5045
500	5%	1	1	1	-2,470702	709,9335
	10%	1	1	1	-2,489715	709,9335
	15%	1	1	1	-2,502940	709,9335
	20%	1	1	1	-2,482473	709,9335
1000	5%	1	1	1	-2,519200	1383,093
	10%	1	1	1	-2,510527	1383,093
	15%	1	1	1	-2,510769	1383,093
	20%	1	1	1	-2,531156	1383,093

Tabel 4.24 menunjukkan hasil identifikasi data simulasi model 5 dengan menggunakan metode GS. Kriteria yang digunakan untuk menentukan model terbaik dengan meminimumkan AIC. Hasil yang diperoleh menunjukkan bahwa model yang terbentuk untuk  $n=200$  adalah model SETAR(2,1,1) dengan  $threshold=-2,441198$ . Untuk  $n=500$  model yang terbentuk adalah model

SETAR(2,1,1) dengan  $threshold=-2,506375$ . Sedangkan untuk  $n=1000$  model terbentuk model SETAR(2,1,1) dengan  $threshold=-2,5396$ . Ketiga model yang terbentuk adalah model yang meminimumkan AIC.

Tabel 4.25 menunjukkan hasil identifikasi data simulasi model 5 dengan metode GA. Pada tabel tersebut dapat dilihat bahwa pencarian nilai  $threshold$  pada metode GA dengan menggunakan 4 daerah pencarian  $threshold$  yaitu 5% – 95% , 10% – 90%, 15% – 85%, dan 20% – 80%. Kriteria untuk mendapatkan model terbaik juga dilakukan dengan meminimumkan AIC. Hasil identifikasi model yang diperoleh dari 4 daerah pencarian  $threshold$  memberikan hasil yang sama. Hal ini dikarenakan nilai  $threshold$  yang disimulasikan berada disekitar nilai rata-rata.

Dari Tabel 4.24 dan Tabel 4.25 diketahui bahwa hasil identifikasi data simulasi model 5 dengan metode GS dan GA memberikan hasil yang sama yaitu model SETAR(2,1,1). Perbedaan antara kedua metode ini terletak pada nilai  $threshold$  yang diperoleh. Perbedaan  $threshold$  terjadi karena metode GA menggunakan semua nilai yang mungkin antara selang pencarian untuk mendapatkan nilai  $threshold$ . Sedangkan pada metode GS,  $threshold$  diperoleh berdasarkan data deret waktu itu sendiri. Hasil estimasi parameter menggunakan metode CLS dan GA untuk  $n=200, 500, \text{ dan } 1000$  dapat dilihat pada Tabel 4.26, Tabel 4.27, dan Tabel 4.28.

Tabel 4.26 Estimasi Parameter Model 5 SETAR(2,1,1) untuk  $n=200$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	-7,088170	0,088170	-7,088173	0,088173
$\phi_{1,1}$	-0,553150	0,053150	-0,553146	0,053146
$\phi_{0,2}$	-1,541410	0,458590	-1,541407	0,458593
$\phi_{1,2}$	0,776480	0,023520	0,776480	0,023520
<i>Threshold</i>	-2,441198	0,058802	-2,441174	0,058826
SSE	954,168021		954,168021	

Berdasarkan hasil pada Tabel 4.26 diketahui bahwa estimasi parameter data simulasi model 5 untuk  $n=200$  memberikan beberapa hasil yang sama. Perbedaan hasil terletak pada  $threshold$  serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA sama dengan CLS.

Secara matematis, model terbaik dari data simulasi model 5 untuk  $n=200$  dapat dituliskan seperti pada persamaan (4.13) berikut.

$$Z_t = \begin{cases} -7,088170 - 0,553150Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq -2,441198 \\ -1,541410 + 0,776480Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > -2,441198 \end{cases} \quad (2.40)$$

Model terbaik pada persamaan (4.13) menunjukkan bahwa data simulasi model 5 untuk  $n=200$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq -2,441198$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > -2,441198$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar  $-7,088170$  dan data pada pengamatan sebelumnya sebesar  $-0,553150$  sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar  $-1,541410$  dan data pada pengamatan sebelumnya sebesar  $0,776480$ .

Tabel 4.27 Estimasi Parameter Model 5 SETAR(2,1,1) untuk  $n=500$

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	-6,702218	0,297782	-6,702217	0,297783
$\phi_{1,1}$	-0,473212	0,026788	-0,473211	0,026789
$\phi_{0,2}$	-1,667536	0,332464	-1,667536	0,332464
$\phi_{1,2}$	0,758681	0,041319	0,758680	0,041320
<i>Threshold</i>	-2,506375	0,006375	-2,502940	0,002940
SSE	2027,33036335		2027,33036334	

Berdasarkan hasil pada Tabel 4.27 diketahui bahwa estimasi parameter data simulasi model 5 untuk  $n=500$  memberikan beberapa hasil yang sama ketika estimasi parameter dilakukan menggunakan metode CLS dan GA. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dari pada CLS.

Secara matematis, model terbaik dari data simulasi model 5 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.14) berikut.

$$Z_t = \begin{cases} -6,702217 - 0,473211Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq -2,502940 \\ -1,667536 + 0,758680Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > -2,502940 \end{cases} \quad (2.41)$$

Model terbaik pada persamaan (4.14) menunjukkan bahwa data simulasi model 5 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq -2,502940$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > -2,502940$ . Nilai *forecast* data

pada *regime* 1 dipengaruhi oleh konstanta sebesar -6,702217 dan data pada pengamatan sebelumnya sebesar -0,473211 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar -1,667536 dan data pada pengamatan sebelumnya sebesar 0,758680.

Tabel 4.28 Estimasi Parameter Model 5 SETAR(2,1,1) untuk n=1000

Metode	CLS		GA	
	Parameter	Bias	Parameter	Bias
$\phi_{0,1}$	-7,022827	0,022827	-7,022826	0,022826
$\phi_{1,1}$	-0,487721	0,012279	-0,487721	0,012279
$\phi_{0,2}$	-2,014490	0,014490	-2,014490	0,014490
$\phi_{1,2}$	0,781885	0,018115	0,781885	0,018115
<i>Threshold</i>	-2,539616	0,039616	-2,510769	0,010769
SSE	3947,54296212		3947,54296211	

Berdasarkan hasil pada Tabel 4.28 diketahui bahwa estimasi parameter data simulasi model 5 untuk n=1000 memberikan beberapa hasil yang sama ketika estimasi parameter dilakukan menggunakan metode CLS dan GA. Perbedaan hasil terletak pada *threshold* serta SSE yang dihasilkan. Hasil SSE yang diperoleh pada metode GA lebih kecil dari pada CLS.

Secara matematis, model terbaik dari data simulasi model 5 untuk n=1000 dapat dituliskan seperti pada persamaan (4.15) berikut.

$$Z_t = \begin{cases} -7,022826 - 0,487721Z_{t-1} + a_{t,1} & \text{jika } Z_{t-1} \leq -2,510769 \\ -2,014490 + 0,781885Z_{t-1} + a_{t,2} & \text{jika } Z_{t-1} > -2,510769 \end{cases} \quad (2.42)$$

Model terbaik pada persamaan (4.15) menunjukkan bahwa data simulasi model 5 untuk n=1000 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq -2,510769$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > -2,510769$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar -7,022826 dan data pada pengamatan sebelumnya sebesar -0,487721 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar -2,014490 dan data pada pengamatan sebelumnya sebesar 0,781885.

Hasil estimasi parameter pada data simulasi 5 untuk n=200, 500, dan 1000 dapat disimpulkan bahwa estimasi parameter dengan metode GA memberikan hasil SSE lebih kecil dibandingkan metode CLS.

Berdasarkan hasil simulasi data pada model 1, model 2, model 3, model 4, dan model 5 SETAR(2,1,1) dapat diambil suatu kesimpulan bahwa hasil identifikasi model yang dilakukan menggunakan metode GA memberikan hasil yang sama dengan metode GS. Perbedaan antara kedua metode terletak pada nilai *threshold* yang dihasilkan. Untuk estimasi parameter yang dilakukan dengan menggunakan metode GA memberikan hasil SSE lebih kecil dibandingkan dengan metode CLS. Akan tetapi perbedaan SSE antara kedua metode ini sangat kecil.

#### 4.1.2 Metode Subset SETAR

Model subset SETAR adalah model SETAR dengan 2 *regime* dan setiap *regime* mengikuti model subset dari model AR(3). Nilai parameter model serta *threshold* yang digunakan dapat dilihat pada Tabel 4.29. Model 1 dan model 2 adalah model subset SETAR dengan delay=1 dan merupakan model yang simulasikan oleh Barogona, Battaglia dan Cucina (2004). Model 3 adalah model 1 dengan mengganti *delay* yang digunakan menjadi 2. Model 4 adalah model 2 dengan mengganti *delay* yang digunakan menjadi 2. Masing-masing model dibangkitkan dengan  $n=200$  dan  $n=500$ .  $a_t$  dibangkitkan mengikuti distribusi normal dengan rata-rata 0 dan standar deviasi 1.

Tabel 4.29 Parameter Simulasi Model Subset SETAR

	<b>Model 1</b>	<b>Model 2</b>	<b>Model 3</b>	<b>Model 4</b>
<b>Regime 1</b>	$\phi_{1,1} = -1,2$ $\phi_{2,1} = -0,7$	$\phi_{2,1} = -0,6$	$\phi_{1,1} = -1,2$ $\phi_{2,1} = -0,7$	$\phi_{2,1} = -0,6$
<b>Regime 2</b>	$\phi_{3,2} = 0,8$	$\phi_{1,2} = 0,75$ $\phi_{3,2} = -0,5$	$\phi_{3,2} = 0,8$	$\phi_{1,2} = 0,75$ $\phi_{3,2} = -0,5$
<b>Threshold</b>	0	0	0	0
<b>Delay</b>	1	1	2	2

Sebelum melakukan analisis hasil simulasi model subset SETAR, data simulasi model subset 1 sampai dengan model subset 4 digambarkan dalam bentuk diskriptif statistik. Tabel 4.29 menunjukkan hasil diskriptif statistik dari data simulasi model subset SETAR. Hasil yang diperoleh menunjukkan bahwa data simulasi dari model yang ditetapkan memiliki rata-rata di sekitar nilai *threshold*



yang ditetapkan pada simulasi model dengan standar deviasi berkisar antara 1 hingga 5.

Tabel 4.30 Statistik Diskriptif Data Simulasi Model Subset SETAR

Model	N	Min	Max	Mean	Sd deviasi
1	200	-5,180482	6,489227	0,3556257	1,745873
	500	-4,507372	6,092582	0,4211405	1,83561
2	200	-12,27556	10,67344	-0,326962	3,749742
	500	-16,69915	14,24953	-0,397958	4,811692
3	200	-2,99373	4,627556	0,4892581	1,406653
	500	-3,201369	6,181812	0,5395028	1,418973
4	200	-3,659974	4,789737	0,2730463	1,337818
	500	-5,826756	5,705486	0,3094578	1,503246

Pemodelan data dengan SETAR dilakukan dengan beberapa tahapan. Tahapan pertama adalah melakukan uji nonlinieritas Terasvirta dan White. Kedua uji ini dilakukan sebagai syarat suatu data dapat dianalisis dengan menggunakan model SETAR.

Tabel 4.31 Uji Nonlinieritas Terasvirta dan White Simulasi Model Subset SETAR

Model	N	Terasvirta		White	
		F	p-value	F	p-value
1	200	13,6304	2,856e-06	13,4295	3,408e-06
	500	40,3998	2,2e-16	40,8137	2,2e-16
2	200	48,8628	2,2e-16	63,752	2,2e-16
	500	145,5336	2,2e-16	216,6274	2,2e-16
3	200	6,3669	0,002092	7,2209	0,0009411
	500	10,4235	3,677e-05	13,4642	2,021e-06
4	200	5,5798	0,004394	4,8468	0,008815
	500	12,0407	7,826e-06	6,0525	0,002529

Rangkuman hasil uji nonlinieritas Terasvirta dan White untuk simulasi model subset 1 sampai model subset 4 ditunjukkan pada Tabel 4.31. Berdasarkan hasil yang ditunjukkan pada Tabel 4.31 tersebut diketahui bahwa hasil simulasi model subset 1 sampai model subset 4 memenuhi sifat nonlinieritas karena nilai  $F_{hitung}$  pada masing-masing model lebih dari 3,04175 atau karena semua nilai p value lebih kecil dari  $\alpha$  (0,05).

Pada studi simulasi model subset SETAR, identifikasi model dilakukan dengan menggunakan 2 metode yaitu metode *GS* (software R package *tsDyn*) dan

algoritma genetika (GA). Identifikasi model dengan menggunakan metode GA menggunakan daerah pencarian *threshold* 15%-85%. Daerah pencarian ini adalah daerah pencarian yang digunakan pada metode *GS*. Untuk estimasi parameter dilakukan dengan menggunakan metode *conditional least square* (CLS) dan Algoritma Genetika (GA).

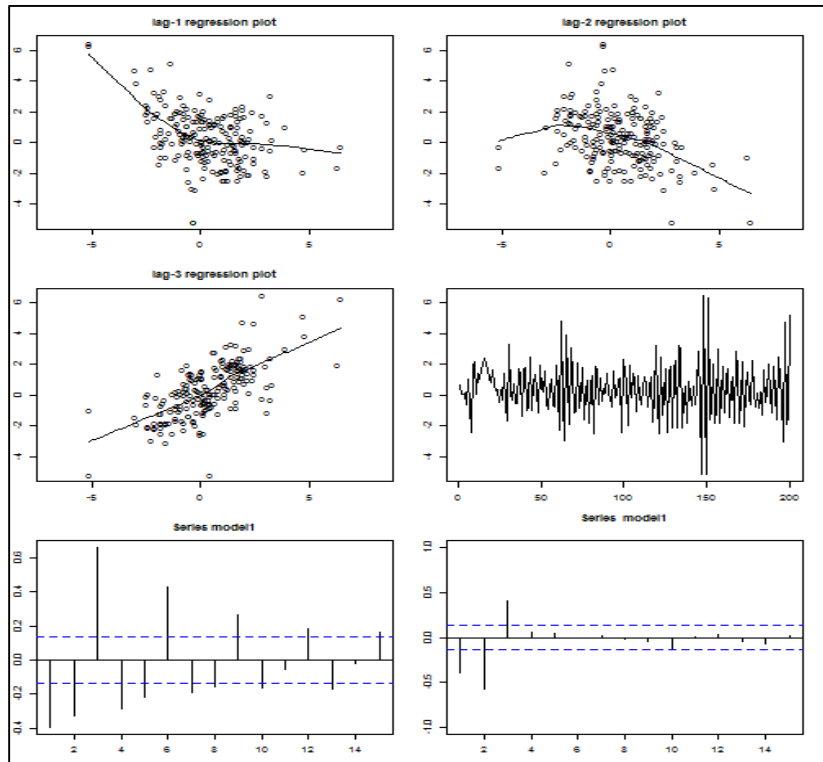
Dalam melakukan pemodelan dengan GA perlu didefinisikan *fitness* yang digunakan. *Fitness* yang digunakan sangat mempengaruhi hasil yang akan diperoleh. Selain *fitness*, parameter-parameter seperti peluang crossover, mutasi, banyak generasi, dan panjang kromosom yang digunakan juga harus ditentukan. Nilai dari masing-masing parameter yang digunakan pada pemodelan dengan GA dapat dilihat pada Tabel 3.3.

#### 4.1.2.1 Model 1 SETAR(2, 2, [3])

Model 1 adalah model SETAR 2 *regime* dimana *regime* 1 mengikuti model AR(2) dan *regime* 2 mengikuti model AR([3]). *Delay* yang digunakan pada model 1 adalah 1 dengan *threshold* adalah 0.

Salah satu cara untuk mengetahui pola dari data simulasi model 1 dengan menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$  seperti terlihat pada Gambar 4.6. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk untuk  $Z_t$  dan  $Z_{t-1}$  serta  $Z_t$  dan  $Z_{t-2}$  tidak membentuk garis linier, sedangkan  $Z_t$  dan  $Z_{t-3}$  membentuk garis linier. Sehingga pemodelan data simulasi model 1 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier dengan *delay* yang digunakan adalah 1 atau 2.

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.31. Hasil yang diperoleh menunjukkan bahwa data simulasi model 1 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan pengujian nonlinieritas, data simulasi model 1 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.



Gambar 4.6 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 1

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari plot PACF yang dibuat. Plot PACF pada Gambar 4.6 menunjukkan bahwa data *cuts off* pada lag 3. Selanjutnya, pendefinisian orde awal untuk identifikasi model 1 dilakukan dengan mengambil nilai  $d = 1, 2$ ,  $p_1 = 3$ , dan  $p_2 = 3$ .

Tabel 4.32 Identifikasi Model Subset 1 dengan Metode GS dan GA

N	Metode	d	p1	p2	Threshold	AIC
200	GS	1	2	3	0,062794647	1,361723
	GA	1	2	[3]	0,06279465	-3,976178
500	GS	1	2	3	0,01980077	40,34689
	GA	1	2	[3]	0,01980077	33,7398

Tabel 4.32 menunjukkan hasil identifikasi model dengan menggunakan metode GS dan GA. Metode GS dan GA memiliki hasil yang sama untuk orde  $d$ ,  $p_1$ , dan threshold. Tetapi terdapat perbedaan untuk orde  $p_2$  yang dihasilkan. Metode GS memberikan hasil  $p_2 = 3$  sedangkan metode GA memberikan hasil

$p_2 = [3]$ . Berdasarkan hasil tersebut, metode GA memberikan hasil identifikasi yang lebih sesuai dibandingkan metode GS.

Untuk  $n=200$ , metode GA menghasilkan nilai  $AIC=-3,976178$  sedangkan metode GS menghasilkan  $AIC=1,361723$ . Untuk  $n=500$ , metode GA menghasilkan nilai  $AIC=33,7398$  sedangkan metode GS menghasilkan  $AIC=40,34689$ . Berdasarkan hasil AIC yang diperoleh pada model 1, metode GA menghasilkan AIC lebih kecil dibandingkan metode GS. Jadi metode GA lebih baik dibandingkan dengan metode GS karena AIC yang dihasilkan lebih kecil. Dari model identifikasi yang diperoleh pada masing-masing metode, selanjutnya dilakukan estimasi dengan menggunakan metode CLS dan metode GA dan diperoleh hasil seperti pada Tabel 4.33 dan Tabel 4.34.

Tabel 4.33 Estimasi Model Identifikasi 1 dengan metode CLS untuk  $n=200$

Metode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	0,2548	0,1685	1,5117	0,1323
	$\phi_{1,1}$	-1,1773	0,1052	-11,1869	<2e-16
	$\phi_{2,1}$	-0,8225	0,0647	-12,7187	<2e-16
	$\phi_{0,2}$	-0,1148	0,1721	-0,6669	0,5056
	$\phi_{1,2}$	0,0313	0,0980	0,3188	0,7502
	$\phi_{2,2}$	0,0101	0,0721	0,1402	0,8886
	$\phi_{2,1}$	0,7529	0,0797	9,4485	<2e-16
	Threshold	0,062794647			
GA	$\phi_{0,1}$	0,2548	0,1688	1,5090	0,135
	$\phi_{1,1}$	-1,1770	0,1054	-11,1700	<2e-16
	$\phi_{2,1}$	-0,8225	0,0648	-12,7000	<2e-16
	$\phi_{3,2}$	0,7450	0,0699	10,6600	<2e-16
		Threshold	0,062794647		

Tabel 4.33 menunjukkan hasil estimasi parameter dari model hasil identifikasi dengan GS dan GA untuk  $n=200$ . Berdasarkan hasil estimasi tersebut, terlihat bahwa terdapat 4 parameter pada model GS yang tidak signifikan yaitu  $\phi_{0,1}$ ,  $\phi_{0,2}$ ,  $\phi_{1,2}$ , dan  $\phi_{2,2}$ . Hasil estimasi parameter pada model identifikasi GA memperlihatkan bahwa terdapat 1 parameter yang tidak signifikan yaitu parameter  $\phi_{0,1}$ . Berdasarkan hasil yang diperoleh, estimasi parameter dari model identifikasi

model GA memberikan hasil yang lebih baik dibandingkan dengan model hasil metode *GS* karena hanya terdapat sedikit parameter yang tidak signifikan.

Secara matematis, model terbaik dari data simulasi model 1 untuk  $n=200$  dapat dituliskan seperti pada persamaan (4.16) berikut.

$$Z_t = \begin{cases} 0,2548 - 1,17701Z_{t-1} - 0,8225Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq 0,062794647 \\ 0,7450Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > 0,062794647 \end{cases} \quad (2.43)$$

Model terbaik pada persamaan (4.16) menunjukkan bahwa data simulasi model 1 untuk  $n=200$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 0,062794647$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 0,062794647$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 0,2548, data pada 1 pengamatan sebelumnya sebesar -1,1770 dan data pada 2 pengamatan sebelumnya sebesar -0,8225 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 3 pengamatan sebelumnya sebesar 0,7450.

Tabel 4.34 Estimasi Model Identifikasi 1 dengan metode CLS untuk  $n=500$

Metode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	0,0700	0,1327	0,5275	0,5981
	$\phi_{1,1}$	-1,2639	0,0781	-16,1737	<2e-16
	$\phi_{2,1}$	-0,6939	0,0460	-15,0936	<2e-16
	$\phi_{0,2}$	-0,0808	0,1161	-0,6953	0,4872
	$\phi_{1,2}$	0,0495	0,0582	0,8502	0,3956
	$\phi_{2,2}$	0,0393	0,0408	0,9633	0,3359
	$\phi_{3,2}$	0,8083	0,0442	18,2981	<2e-16
	Threshold	0,01980077			
GA	$\phi_{1,1}$	-1,2960	0,0494	-26,2600	<2e-16
	$\phi_{2,1}$	-0,6841	0,0421	-16,2400	<2e-16
	$\phi_{3,2}$	0,7866	0,0385	20,4300	<2e-16
		Threshold	0,01980077		

Tabel 4.34 menunjukkan hasil estimasi parameter model identifikasi dengan menggunakan metode *GS* dan *GA* untuk  $n=500$ . Berdasarkan hasil tersebut, terdapat 4 parameter yang tidak signifikan pada model identifikasi *GS* yaitu  $\phi_{0,1}$ ,  $\phi_{0,2}$ ,  $\phi_{1,2}$ , dan  $\phi_{2,2}$ . Hasil estimasi parameter pada model identifikasi *GA* memperlihatkan bahwa semua parameter signifikan. Berdasarkan hasil yang diperoleh, estimasi parameter dari model identifikasi model *GA* memberikan hasil

yang lebih baik dibandingkan dengan model hasil metode *GS* karena semua parameter yang dimiliki signifikan.

Secara matematis, model terbaik dari data simulasi model 1 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.17) berikut.

$$Z_t = \begin{cases} -1,2960Z_{t-1} - 0,6841Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq 0,01980077 \\ 0,7866Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > 0,01980077 \end{cases} \quad (2.44)$$

Model terbaik pada persamaan (4.17) menunjukkan bahwa data simulasi model 1 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 0,01980077$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 0,01980077$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh data pada 1 pengamatan sebelumnya sebesar -1,2960 dan data pada 2 pengamatan sebelumnya sebesar -0,6841 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 3 pengamatan sebelumnya sebesar 0,7866.

Hasil analisis yang diperoleh dari data simulasi model 1 adalah metode GA memberikan hasil identifikasi dan estimasi yang lebih baik dibandingkan dengan metode *GS*. Metode *GS* akan menghasilkan estimasi yang lebih baik ketika model hasil identifikasi selanjutnya dianalisis dengan menggunakan metode *stepwise* untuk mendapatkan model terbaik. Sehingga hasil yang diperoleh menghasilkan parameter model yang signifikan.

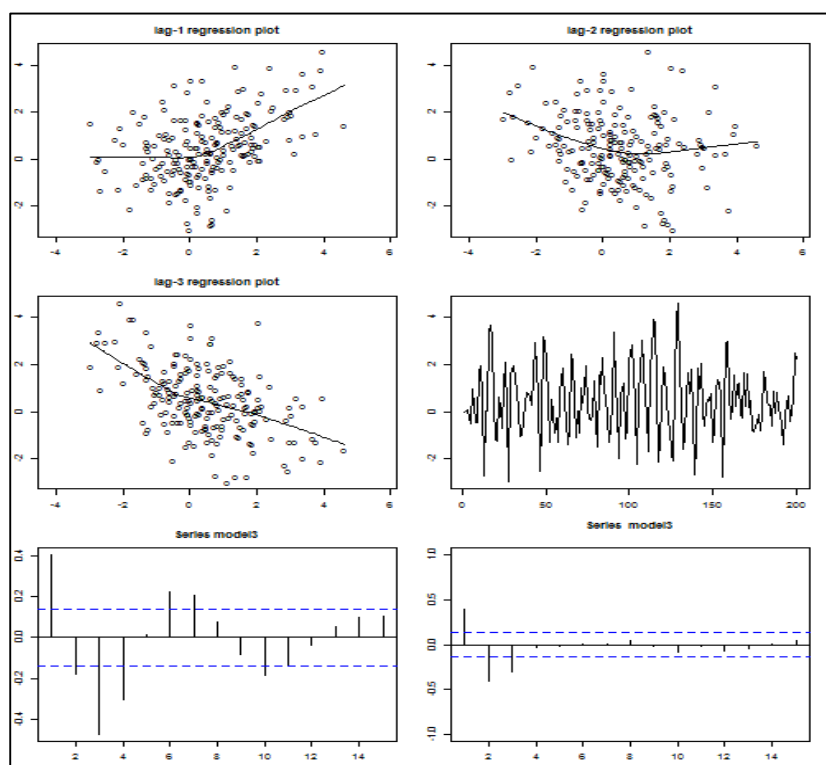
#### 4.1.2.2 Model 2 Subset SETAR(2, [2], [1, 3])

Model 2 adalah model SETAR 2 *regime* dimana *regime* 1 mengikuti model model AR([2]) dan *regime* 2 mengikuti model AR([1,3]). *Delay* yang digunakan pada model 2 adalah 1 dengan *threshold* adalah 0.

Salah satu cara untuk mengetahui pola dari data simulasi model 2 dengan menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$  seperti terlihat pada Gambar 4.7. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk untuk  $Z_t$  dan  $Z_{t-1}$  serta  $Z_t$  dan  $Z_{t-2}$  tidak membentuk garis linier, sedangkan  $Z_t$  dan  $Z_{t-3}$  membentuk garis linier. Sehingga pemodelan data simulasi model 2

dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier dengan *delay* yang digunakan adalah 1 atau 2.

Selain menggunakan lag plot antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.31. Hasil yang diperoleh menunjukkan bahwa data simulasi model 2 memenuhi asumsi nonlinier. Berdasarkan hasil lag plot dan pengujian nonlinieritas, data simulasi model 2 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.



Gambar 4.7 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 2

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari plot PACF yang dibuat. Plot PACF pada Gambar 4.7 menunjukkan bahwa data *cuts off* pada lag 3. Selanjutnya, pendefinisian orde awal untuk identifikasi model 2 dilakukan dengan mengambil nilai  $d = 1, 2$ ,  $p_1 = 3$ , dan  $p_2 = 3$ .

Tabel 4.35 Identifikasi Model 2 SETAR(2,[2],[1,3]) dengan GS dan GA

N	Metode	d	p1	p2	Threshold	AIC
200	GS	1	2	3	0,01262748	9,043961
	GA	1	[2]	[1,3]	0,01534868	4,130931
500	GS	1	2	3	0,003743762	43,19298
	GA	1	[2]	[1,3]	0,003743762	37,14112

Tabel 4.35 menunjukkan hasil identifikasi model dengan menggunakan metode GS dan GA. Untuk  $n=200$ , metode GS dan GA memberikan hasil yang sama untuk orde  $d$  tetapi terdapat perbedaan untuk orde  $p_1, p_2$  dan threshold yang dihasilkan. Metode GS memberikan hasil  $p_1 = 2, p_2 = 3$  dan  $threshold=0,01262748$  sedangkan metode GA memberikan hasil  $p_1 = [2], p_2 = [1,3]$  dan  $threshold=0,01534868$ . Metode GA menghasilkan nilai  $AIC=4,130931$  sedangkan metode GS menghasilkan  $AIC=9,043961$ . Berdasarkan hasil tersebut, metode GA memberikan hasil identifikasi yang lebih sesuai dibandingkan metode GS dan menghasilkan nilai AIC yang lebih kecil.

Untuk  $n=500$ , metode GS dan GA memiliki hasil yang sama untuk orde  $d$  dan  $threshold$  yang dihasilkan tetapi terdapat perbedaan untuk orde  $p_1$  dan  $p_2$ . Metode GS memberikan hasil  $p_1 = 2, p_2 = 3$  sedangkan metode GA memberikan hasil  $p_1 = [2], p_2 = [1,3]$ . Metode GA menghasilkan nilai  $AIC=37,14112$  sedangkan metode GS menghasilkan  $AIC=43,19298$ . Berdasarkan hasil tersebut, metode GA memberikan hasil identifikasi yang lebih sesuai dibandingkan metode GS dan menghasilkan nilai AIC yang lebih kecil.

Hasil identifikasi model dari masing-masing metode, selanjutnya diestimasi dengan menggunakan metode CLS dan metode GA dan diperoleh hasil seperti pada Tabel 4.36 dan Tabel 4.37.

Tabel 4.36 menunjukkan hasil estimasi parameter dari model hasil identifikasi dengan GS dan GA untuk  $n=200$ . Berdasarkan hasil estimasi tersebut, terlihat bahwa terdapat 4 parameter pada model GS yang tidak signifikan yaitu  $\phi_{0,1}, \phi_{1,1}, \phi_{0,2}$ , dan  $\phi_{2,2}$ . Hasil estimasi parameter pada model identifikasi GA memperlihatkan bahwa semua parameter signifikan. Berdasarkan hasil yang



diperoleh, estimasi parameter dari model identifikasi model GA memberikan hasil yang lebih baik dibandingkan dengan model hasil metode GS karena semua parameter yang dimiliki signifikan.

Tabel 4.36 Estimasi Model Identifikasi 2 dengan metode CLS untuk n=200

Metode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	-0,14244	0,180229	-0,7903	0,4303
	$\phi_{1,1}$	-0,08371	0,154395	-0,5422	0,5883
	$\phi_{2,1}$	-0,73005	0,116545	-6,2641	2,39E-09
	$\phi_{0,2}$	0,16994	0,15276	1,1125	0,2673
	$\phi_{1,2}$	0,642485	0,099585	6,4516	8,71E-10
	$\phi_{2,2}$	-0,07943	0,077143	-1,0296	0,3045
	$\phi_{3,2}$	-0,45335	0,069792	-6,4957	6,85E-10
	Threshold	0,01262748			
GA	$\phi_{2,1}$	-0,723	0,1164	-6,211	3,17E-08
	$\phi_{1,2}$	0,6819	0,05533	12,32	< 2,2e-16
	$\phi_{3,2}$	-0,478	0,05661	-8,444	8,70E-14
	Threshold	0,01534868			

Secara matematis, model terbaik dari data simulasi model 2 untuk n=200 dapat dituliskan seperti pada persamaan (4.18) berikut.

$$Z_t = \begin{cases} -0,723Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq 0,01534868 \\ 0,6819Z_{t-1} - 0,478Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > 0,01534868 \end{cases} \quad (2.45)$$

Model terbaik pada persamaan (4.18) menunjukkan bahwa data simulasi model 2 untuk n=200 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 0,01534868$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 0,01534868$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh data pada 2 pengamatan sebelumnya sebesar -0,723 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 1 pengamatan sebelumnya sebesar 0,6819 dan data pada 3 pengamatan sebelumnya sebesar -0,478.

Tabel 4.37 adalah hasil estimasi model identifikasi dengan metode GS dan GA untuk n=500. Berdasarkan tabel tersebut terlihat bahwa hasil estimasi parameter untuk model identifikasi GS memiliki 4 parameter yang tidak signifikan yaitu  $\phi_{0,1}$ ,  $\phi_{1,1}$ ,  $\phi_{0,2}$ , dan  $\phi_{2,2}$ . Untuk model identifikasi GA, hasil estimasi yang diperoleh menunjukkan bahwa semua parameter yang masuk dalam model adalah signifikan. Berdasarkan hasil yang diperoleh, estimasi parameter dari model

identifikasi model GA memberikan hasil yang lebih baik dibandingkan dengan model hasil metode GS karena semua parameter yang dimiliki signifikan.

Tabel 4.37 Estimasi Model Identifikasi 2 dengan metode CLS untuk n=500

Metode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	0,095015	0,122097	0,7782	0,4368
	$\phi_{1,1}$	0,040501	0,112335	0,3605	0,7186
	$\phi_{2,1}$	-0,57522	0,067736	-8,4920	< 2,2e-16
	$\phi_{0,2}$	0,044425	0,097838	0,4541	0,65
	$\phi_{1,2}$	0,729282	0,063516	11,4819	< 2,2e-16
	$\phi_{2,2}$	-0,05056	0,053137	-0,9515	0,3418
	$\phi_{3,2}$	-0,46536	0,047866	-9,7221	< 2,2e-16
	Threshold	0,003743762			
GA	$\phi_{2,1}$	-0,5756	0,06786	-8,4820	7,29E-15
	$\phi_{1,2}$	0,7236	0,03522	20,5400	< 2,2e-16
	$\phi_{3,2}$	-0,4885	0,039	-12,520	< 2,2e-16
	Threshold	0,003743762			

Secara matematis, model terbaik dari data simulasi model 2 untuk n=500 dapat dituliskan seperti pada persamaan (4.19) berikut.

$$Z_t = \begin{cases} -0,5756Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq 0,003743762 \\ 0,7236Z_{t-1} - 0,4885Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > 0,003743762 \end{cases} \quad (2.46)$$

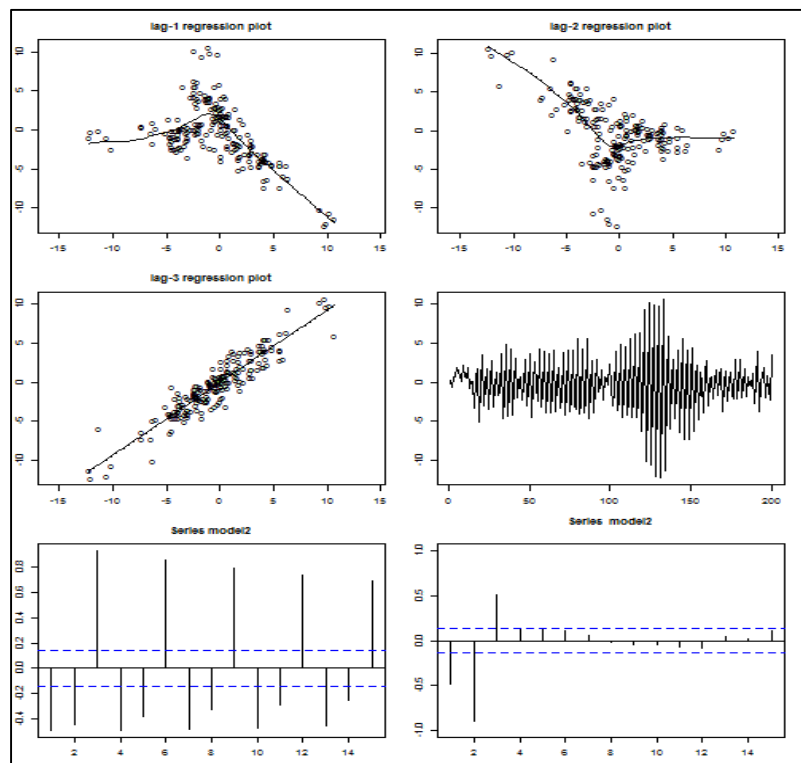
Model terbaik pada persamaan (4.19) menunjukkan bahwa data simulasi model 2 untuk n=500 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 0,003743762$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 0,003743762$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh data pada 2 pengamatan sebelumnya sebesar -0,5756 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 1 pengamatan sebelumnya sebesar 0,7236 dan data pada 3 pengamatan sebelumnya sebesar -0,4885.

Hasil analisis yang diperoleh dari data simulasi model 2 adalah metode GA memberikan hasil identifikasi dan estimasi yang lebih baik dibandingkan dengan metode GS. Metode GS akan menghasilkan estimasi yang lebih baik ketika model hasil identifikasi selanjutnya dianalisis dengan menggunakan metode stepwise untuk mendapatkan model terbaik. Sehingga hasil yang diperoleh menghasilkan parameter model yang signifikan.

#### 4.1.2.3 Model 3 Subset SETAR(2, 2, [3])

Model 3 adalah model SETAR 2 *regime* dimana *regime* 1 mengikuti model AR(2) dan *regime* 2 mengikuti model AR([3]). *Delay* yang digunakan pada model 3 adalah 2 dengan *threshold* adalah 0.

Salah satu cara untuk mengetahui pola dari data simulasi model 3 dengan menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$  seperti terlihat pada Gambar 4.8. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk untuk  $Z_t$  dan  $Z_{t-1}$  serta  $Z_t$  dan  $Z_{t-2}$  tidak membentuk garis linier, sedangkan  $Z_t$  dan  $Z_{t-3}$  membentuk garis linier. Sehingga pemodelan data simulasi model 3 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier dengan *delay* yang digunakan adalah 1 atau 2.



Gambar 4.8 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 3

Selain menggunakan plot lag antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.31. Hasil

yang diperoleh menunjukkan bahwa data simulasi model 3 memenuhi asumsi nonlinier. Berdasarkan hasil plot lag dan pengujian nonlinieritas, data simulasi model 3 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Selanjutnya, untuk menentukan orde maksimum yang digunakan pada identifikasi model 3 dapat dilakukan dengan melihat hasil plot PACF pada Gambar 4.8. Hasil Plot PACF pada gambar tersebut menunjukkan bahwa data *cuts off* pada lag 3 sehingga pendefinisian orde awal untuk identifikasi model 3 dilakukan dengan mengambil orde maksimum 3 dan diperoleh identifikasi model awal adalah  $d = 1, 2, 3, p_1 = 3, \text{ dan } p_2 = 3$ .

Tabel 4.38 Identifikasi Model 3 SETAR(2,2,[3]) dengan GS dan GA

N	Metode	d	p1	p2	Threshold	AIC
200	GS	2	2	3	0,006884976	-6,794098
	GA	2	2	[3]	0,01975018	-12,928997
500	GS	2	2	3	-0,02614313	42,31569
	GA	2	2	[3]	-0,02350497	34,10763

Tabel 4.38 menunjukkan hasil identifikasi model dengan menggunakan metode GS dan GA. Untuk  $n=200$ , metode GS dan GA memiliki hasil yang sama untuk orde  $d$  dan  $p_1$  tetapi terdapat perbedaan untuk orde  $p_2$  dan *threshold* yang dihasilkan. Metode GS memberikan hasil  $p_2 = 3$  dengan *threshold*=0,006884976 sedangkan metode GA memberikan hasil  $p_2 = [3]$  dengan *threshold*=0,01975018. Metode GA menghasilkan nilai  $AIC=-12,928997$  sedangkan metode GS menghasilkan  $AIC=-6,794098$ . Berdasarkan hasil tersebut, metode GA memberikan hasil identifikasi yang lebih sesuai dibandingkan metode GS dan menghasilkan nilai AIC yang lebih kecil.

Untuk  $n=500$ , metode GS dan GA memiliki hasil yang sama untuk  $d$  dan  $p_1$  tetapi terdapat perbedaan untuk  $p_2$  dan *threshold* yang dihasilkan. Metode GS memberikan hasil  $p_2 = 3$  dengan *threshold*=-0,02614313 sedangkan metode GA memberikan hasil  $p_2 = [3]$  dengan *threshold*=-0,02350497. Metode GA menghasilkan nilai  $AIC=34,10763$  sedangkan metode GS menghasilkan

AIC=42,31569. Berdasarkan hasil tersebut, metode GA memberikan hasil identifikasi yang lebih sesuai dibandingkan metode GS dan menghasilkan nilai AIC yang lebih kecil.

Dari model identifikasi yang diperoleh pada masing-masing metode, selanjutnya dilakukan estimasi dengan metode CLS dan metode GA dan diperoleh hasil seperti pada Tabel 4.39 dan Tabel 4.40.

Tabel 4.39 Estimasi Model Identifikasi 3 dengan metode CLS untuk n=200

Motode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	-0,12802	0,170017	-0,753	0,4524
	$\phi_{1,1}$	-1,18276	0,033517	-35,2882	<2e-16
	$\phi_{2,1}$	-0,70244	0,040443	-17,3688	<2e-16
	$\phi_{0,2}$	-0,21694	0,168029	-1,2911	0,1982
	$\phi_{1,2}$	-0,06472	0,073088	-0,8855	0,377
	$\phi_{2,2}$	-0,0295	0,098054	-0,3009	0,7638
	$\phi_{3,2}$	0,795308	0,068332	11,6388	<2e-16
	Threshold	0,006884976			
GA	$\phi_{1,1}$	-1,193	0,02782	-42,9	<2e-16
	$\phi_{2,1}$	-0,6778	0,0244	-27,78	<2e-16
	$\phi_{3,2}$	0,847	0,04771	17,75	<2e-16
	Threshold	0,01975018			

Tabel 4.39 menunjukkan hasil estimasi parameter model hasil identifikasi dengan GS dan GA untuk n=200. Berdasarkan hasil tersebut, terdapat 4 parameter yang tidak signifikan pada model identifikasi GS yaitu  $\phi_{0,1}$ ,  $\phi_{0,2}$ ,  $\phi_{1,2}$ , dan  $\phi_{2,2}$ . Hasil estimasi parameter pada model identifikasi GA memperlihatkan bahwa semua parameter signifikan. Berdasarkan hasil yang diperoleh, estimasi parameter dari model identifikasi model GA memberikan hasil yang lebih baik dibandingkan dengan model hasil metode GS karena semua parameter yang dimiliki signifikan.

Secara matematis, model terbaik dari data simulasi model 3 untuk n=200 dapat dituliskan seperti pada persamaan (4.20) berikut.

$$Z_t = \begin{cases} -1,193Z_{t-1} - 0,6778Z_{t-2} + a_{t,1} & \text{jika } Z_{t-2} \leq 0,01975018 \\ 0,847Z_{t-3} + a_{t,2} & \text{jika } Z_{t-2} > 0,01975018 \end{cases} \quad (2.47)$$

Model terbaik pada persamaan (4.20) menunjukkan bahwa data simulasi model 3 untuk n=200 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-1} \leq 0,01975018$

dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-1} > 0,01975018$ . Nilai *forecast* data pada *regime* 1 dipengaruhi data pada 1 pengamatan sebelumnya sebesar -1,193 dan data pada 2 pengamatan sebelumnya sebesar -0,6778 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 3 pengamatan sebelumnya sebesar 0,847.

Tabel 4.40 Estimasi Model Identifikasi 3 dengan metode CLS untuk n=500

Metode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	-0,0672	0,109249	-0,6151	0,5388
	$\phi_{1,1}$	-1,18502	0,016991	-69,7457	<2e-16
	$\phi_{2,1}$	-0,71235	0,020973	-33,9654	<2e-16
	$\phi_{0,2}$	0,021772	0,112225	0,194	0,8463
	$\phi_{1,2}$	0,031018	0,055848	0,5554	0,5789
	$\phi_{2,2}$	0,04044	0,068138	0,5935	0,5531
	$\phi_{3,2}$	0,804285	0,048017	16,75	<2e-16
	Threshold	-0,02614313			
GA	$\phi_{1,1}$	-1,19	0,01416	-84,05	<2e-16
	$\phi_{2,1}$	-0,7022	0,0131	-53,61	<2e-16
	$\phi_{3,2}$	0,7805	0,02625	29,74	<2e-16
	Threshold	-0,02350497			

Tabel 4.40 menunjukkan hasil estimasi parameter model hasil identifikasi dengan GS dan GA untuk n=500. Berdasarkan tabel tersebut terlihat bahwa hasil estimasi untuk model identifikasi GS memiliki 4 parameter yang belum signifikan yaitu  $\phi_{0,1}$ ,  $\phi_{0,2}$ ,  $\phi_{1,2}$ , dan  $\phi_{2,2}$ . Untuk model identifikasi GA, hasil estimasi yang diperoleh menunjukkan bahwa semua parameter yang masuk dalam model adalah signifikan. Berdasarkan hasil yang diperoleh, estimasi parameter dari model identifikasi model GA memberikan hasil yang lebih baik dibandingkan dengan model hasil metode GS karena semua parameter yang dimiliki signifikan.

Secara matematis, model terbaik dari data simulasi model 3 untuk n=500 dapat dituliskan seperti pada persamaan (4.21) berikut.

$$Z_t = \begin{cases} -1,19Z_{t-1} - 0,7022Z_{t-2} + a_{t,1} & \text{jika } Z_{t-2} \leq -0,02350497 \\ 0,8705Z_{t-3} + a_{t,2} & \text{jika } Z_{t-2} > -0,02350497 \end{cases} \quad (2.48)$$

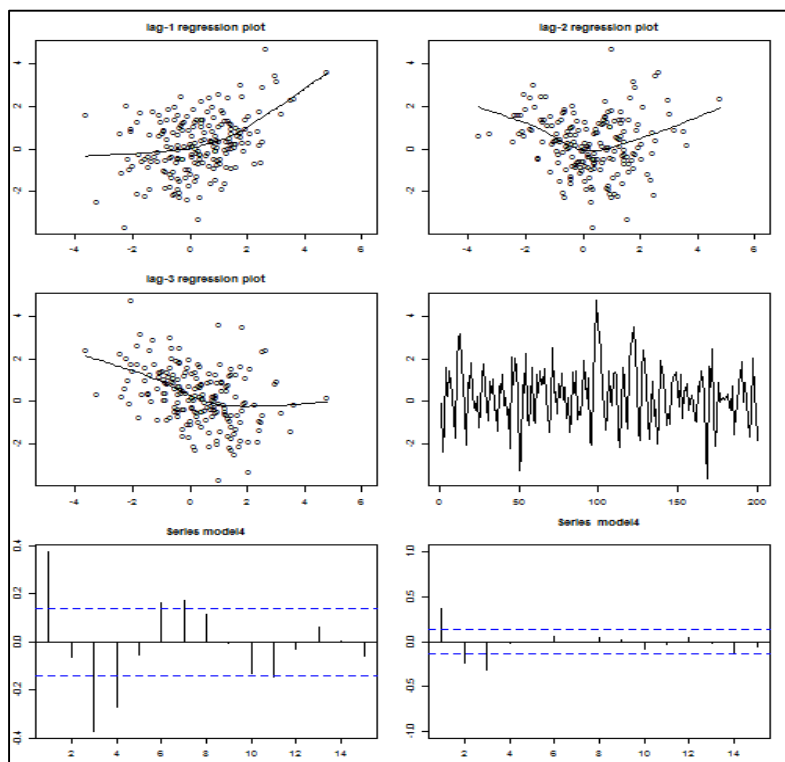
Model terbaik pada persamaan (4.21) menunjukkan bahwa data simulasi model 3 untuk n=500 masuk *regime* 1 ketika nilai pengamatan  $Z_{t-2} \leq -0,02350497$

dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-2} > -0,02350497$ . Nilai *forecast* data pada *regime* 1 dipengaruhi data pada 1 pengamatan sebelumnya sebesar -1,19 dan data pada 2 pengamatan sebelumnya sebesar -0,7022 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 3 pengamatan sebelumnya sebesar 0,78405.

Hasil analisis yang diperoleh dari data simulasi model 3 adalah metode GA memberikan hasil identifikasi dan estimasi yang lebih baik dibandingkan dengan metode GS. Metode GS akan menghasilkan estimasi yang lebih baik ketika model hasil identifikasi selanjutnya dianalisis dengan menggunakan metode stepwise untuk mendapatkan model terbaik. Sehingga hasil yang diperoleh menghasilkan parameter model yang signifikan.

#### 4.1.2.4 Model 4 Subset SETAR(2, [2], [1, 3])

Model 4 adalah model SETAR 2 *regime* dimana *regime* 1 mengikuti model AR([2]) sedangkan *regime* 2 mengikuti model AR([1,3]). *Delay* yang digunakan pada model 4 adalah 2 dengan *threshold* adalah 0.



Gambar 4.9 Lag Plot, Plot Deret Waktu, ACF, dan PACF Data Simulasi Model Subset 4

Salah satu cara untuk mengetahui pola dari data simulasi model 4 dengan menggunakan plot antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$  seperti terlihat pada Gambar 4.9. Berdasarkan plot tersebut terlihat bahwa plot yang terbentuk untuk  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$  serta  $Z_t$  dan  $Z_{t-3}$  tidak membentuk garis linier. Sehingga pemodelan data simulasi model 4 dapat dilakukan dengan menggunakan pendekatan analisis deret waktu nonlinier dengan *delay* yang digunakan adalah 1, 2, atau 3.

Selain menggunakan plot lag antara  $Z_t$  dan  $Z_{t-1}$ ,  $Z_t$  dan  $Z_{t-2}$ , serta  $Z_t$  dan  $Z_{t-3}$ , untuk melihat pola suatu data dapat dilakukan dengan uji nonlinieritas terasvirta dan white. Hasil kedua uji tersebut dapat dilihat pada Tabel 4.31. Hasil yang diperoleh menunjukkan bahwa data simulasi model 4 memenuhi asumsi nonlinier. Berdasarkan hasil plot lag dan pengujian nonlinieritas, data simulasi model 4 selanjutnya dianalisis menggunakan analisis deret waktu nonlinier model SETAR.

Langkah analisis model SETAR dilanjutkan dengan menentukan orde maksimum yang digunakan dalam identifikasi model. Orde maksimum dilihat dari plot PACF yang dibuat. Plot PACF pada Gambar 4.9 menunjukkan bahwa data *cuts off* pada lag 3. Selanjutnya, pendefinisian orde awal untuk identifikasi model 4 dilakukan dengan mengambil nilai  $d = 1, 2, 3$ ,  $p_1 = 3$ , dan  $p_2 = 3$ .

Tabel 4.41 Identifikasi Model 4 SETAR(2,[2],[1,3]) dengan GS dan GA

N	Metode	d	p1	p2	Threshold	AIC
200	GS	2	3	3	0,10775167	-15,25968
	GA	2	[2,3]	[1,3]	0,1603969	-19,50702
500	GS	2	2	3	-0,004528023	61,80553
	GA	2	[2]	[1,3]	-0,004528023	55,97946

Tabel 4.41 menunjukkan hasil identifikasi model dengan menggunakan metode GS dan GA. Untuk  $n=200$ , metode GS dan GA memiliki hasil yang sama untuk orde  $d$  tetapi terdapat perbedaan untuk orde  $p_1$ ,  $p_2$ , dan *threshold* yang dihasilkan. Metode GS memberikan hasil  $p_1 = 3$ ,  $p_2 = 3$  dengan *threshold*=0,10775167 sedangkan metode GA memberikan hasil  $p_1 = [2,3]$



$p_2 = [1,3]$  dengan  $threshold=0,1603969$ . Berdasarkan hasil tersebut, metode GA dan metode GS memberikan hasil identifikasi berbeda dengan model yang disimulasi. Untuk AIC dari metode GA lebih kecil dibandingkan metode GS. Metode GA memberikan hasil  $AIC=-19,50702$  sedangkan metode GS menghasilkan  $AIC=-15,25968$ .

Untuk  $n=500$ , metode GS dan GA memiliki hasil yang sama untuk orde  $d$  dan  $threshold$  yang dihasilkan tetapi terdapat perbedaan untuk orde  $p_1$  dan  $p_2$ . Metode GS memberikan hasil  $p_1 = 2, p_2 = 3$  sedangkan metode GA memberikan hasil  $p_1 = [2], p_2 = [1,3]$ . Nilai AIC yang dihasilkan metode GA lebih kecil dibandingkan metode GS. Metode GA menghasilkan nilai  $AIC=55,97946$  sedangkan metode GS menghasilkan  $AIC=61,80553$ . Pada model 4, metode GA memberikan hasil identifikasi model yang lebih sesuai dibandingkan dengan metode GS untuk  $n=500$ . Hasil identifikasi model dari masing-masing metode, selanjutnya diestimasi dengan metode CLS dan metode GA dan diperoleh hasil seperti pada Tabel 4.42 dan Tabel 4.43.

Tabel 4.42 Estimasi Model Identifikasi 4 dengan metode CLS untuk  $n=200$

Metode	parameter	Coef	SE Coef	T	p-val
GS	$\phi_{0,1}$	-0,02576	0,15339	-0,168	0,866793
	$\phi_{1,1}$	-0,07547	0,096336	-0,7834	0,434367
	$\phi_{2,1}$	-0,50613	0,138776	-3,6471	0,000342
	$\phi_{3,1}$	-0,25251	0,112481	-2,2449	0,025915
	$\phi_{0,2}$	-0,22436	0,158866	-1,4123	0,159491
	$\phi_{1,2}$	0,716542	0,077951	9,1922	< 2,2e-16
	$\phi_{2,2}$	0,213774	0,12205	1,7515	0,081452
	$\phi_{3,2}$	-0,49742	0,066493	-7,4808	2,58E-12
	<i>Threshold</i>	0,10775167			
GA	$\phi_{2,1}$	-0,5065	0,08905	-5,687	1,74E-07
	$\phi_{3,1}$	-0,2109	0,09226	-2,286	0,02472
	$\phi_{1,2}$	0,7856	0,06631	11,85	< 2,2e-16
	$\phi_{3,2}$	-0,4738	0,06142	-7,714	8,06E-12
	<i>Threshold</i>	0,1603969			

Tabel 4.42 menunjukkan hasil estimasi parameter model hasil identifikasi dengan GS dan GA untuk  $n=200$ . Berdasarkan tabel tersebut terlihat bahwa hasil

estimasi untuk model identifikasi *GS* memiliki 4 parameter yang belum signifikan. Untuk model identifikasi *GA*, hasil estimasi yang diperoleh menunjukkan bahwa semua parameter yang masuk dalam model adalah signifikan. Berdasarkan hasil yang diperoleh, estimasi parameter dari model identifikasi model *GA* memberikan hasil yang lebih baik dibandingkan dengan model hasil metode *GS* karena semua parameter yang dimiliki signifikan.

Secara matematis, model terbaik dari data simulasi model 4 untuk  $n=200$  dapat dituliskan seperti pada persamaan (4.22) berikut.

$$Z_t = \begin{cases} -0,5065Z_{t-2} - 0,2109Z_{t-3} + a_{t,1} & \text{jika } Z_{t-2} \leq 0,1603969 \\ 0,7856Z_{t-1} - 0,4738Z_{t-3} + a_{t,2} & \text{jika } Z_{t-2} > 0,1603969 \end{cases} \quad (2.49)$$

Model terbaik pada persamaan (4.22) menunjukkan bahwa data simulasi model 2 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-2} \leq 0,1603969$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-2} > 0,1603969$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh data pada 2 pengamatan sebelumnya sebesar -0,5065 dan data pada 3 pengamatan sebelumnya sebesar -0,2109 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 1 pengamatan sebelumnya sebesar 0,7856 dan data pada 3 pengamatan sebelumnya sebesar -0,4885.

Tabel 4.43 Estimasi Model Identifikasi 4 dengan metode CLS untuk  $n=500$

Metode	parameter	Coef	SE Coef	T	p-val
<b>GS</b>	$\phi_{0,1}$	-0,03758	0,108218	-0,3473	0,7285
	$\phi_{1,1}$	0,006082	0,047846	0,1271	0,8989
	$\phi_{2,1}$	-0,55303	0,073366	-7,538	2,31E-13
	$\phi_{0,2}$	-0,0272	0,100971	-0,2694	0,7878
	$\phi_{1,2}$	0,771014	0,049364	15,6189	< 2,2e-16
	$\phi_{2,2}$	-0,0586	0,073227	-0,8003	0,4239
	$\phi_{3,2}$	-0,41922	0,040042	-10,4695	< 2,2e-16
	Threshold	-0,004528023			
<b>GA</b>	$\phi_{2,1}$	-0,5338	0,05046	-10,58	< 2,2e-16
	$\phi_{1,2}$	0,7323	0,04062	18,03	< 2,2e-16
	$\phi_{3,2}$	-0,4445	0,03583	-12,41	< 2,2e-16
	Threshold	-0,004528023			

Tabel 4.43 menunjukkan hasil estimasi parameter model hasil identifikasi dengan *GS* dan *GA* untuk  $n=500$ . Berdasarkan tabel tersebut terlihat bahwa hasil

estimasi untuk model identifikasi *GS* memiliki 4 parameter yang belum signifikan. Untuk model identifikasi *GA*-subset, hasil estimasi yang diperoleh menunjukkan bahwa semua parameter yang masuk dalam model adalah signifikan. Berdasarkan hasil yang diperoleh, estimasi parameter dari model identifikasi model *GA* memberikan hasil yang lebih baik dibandingkan dengan model hasil metode *GS* karena semua parameter yang dimiliki signifikan.

Secara matematis, model terbaik dari data simulasi model 4 untuk  $n=500$  dapat dituliskan seperti pada persamaan (4.23) berikut.

$$Z_t = \begin{cases} -0,5338Z_{t-2} + a_{t,1} & \text{jika } Z_{t-1} \leq -0,004528023 \\ 0,7323Z_{t-1} - 0,4445Z_{t-3} + a_{t,2} & \text{jika } Z_{t-1} > -0,004528023 \end{cases} \quad (2.50)$$

Model terbaik pada persamaan (4.23) menunjukkan bahwa data simulasi model 4 untuk  $n=500$  masuk *regime* 1 ketika nilai pengamatan  $Z_{t-2} \leq -0,004528023$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-2} > -0,004528023$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh data pada 2 pengamatan sebelumnya sebesar  $-0,5338$  sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh data pada 1 pengamatan sebelumnya sebesar  $0,7323$  dan data pada 3 pengamatan sebelumnya sebesar  $-0,4445$ .

Hasil analisis yang diperoleh dari data simulasi model 4 adalah metode *GA* memberikan hasil identifikasi dan estimasi yang lebih baik dibandingkan dengan metode *GS*. Metode *GS* akan menghasilkan estimasi yang lebih baik ketika model hasil identifikasi selanjutnya dianalisis dengan menggunakan metode *stepwise* untuk mendapatkan model terbaik. Sehingga hasil yang diperoleh menghasilkan parameter model yang signifikan.

Berdasarkan hasil dari simulasi model 1 hingga 4 dapat disimpulkan bahwa metode *GA* memberikan hasil identifikasi yang lebih baik dibandingkan metode *GS* karena menghasilkan nilai *AIC* lebih kecil.

## 4.2 Aplikasi Metode SETAR dengan Algoritma Genetika Pada Data Return Saham

Data yang digunakan dalam penelitian ini adalah data return saham Bank Tabungan Negara (Persero) Tbk (BBTN) dan data return saham Wijaya Karya

(Persero) Tbk (WIKA). Untuk data return saham BBTN dimulai dari 4 Januari 2010 sampai 31 Desember 2014 sedangkan untuk data return saham WIKA dimulai dari 1 November 2007 sampai 31 Desember 2014 data tersebut digunakan sebagai data *in sample*. Untuk data dari 1 Januari sampai 28 Februari 2015 akan digunakan sebagai validasi (data *out sample*).

Sebelum melakukan analisis lebih lanjut, data return saham dapat digambarkan dalam bentuk diskripsi data seperti terlihat pada Tabel 4.44. Hasil pada tabel tersebut menunjukkan bahwa data return saham BBTN memiliki mean sebesar 0,0003926327 dengan *variance* 0,0006247415. sedangkan untuk data return saham WIKA memiliki mean sebesar 0,001089609 dengan nilai *variance* sebesar 0,0009339951. berdasarkan hasil *variance* dari kedua data return saham tersebut diketahui bahwa return saham WIKA lebih menyebar dibandingkan dengan return saham BBTN. Hal ini juga ditunjukkan dari nilai maksimum dan minimum pada data return saham yang rentangnya lebih besar dibandingkan dengan return saham BBTN.

Tabel 4.44 Statistik Diskriptif Data Return Saham

Data	N	Mean	Stdev	Variance	Minimum	Maximum
<b>BBTN</b>	1273	0,0003926327	0,02499483	0,0006247415	-0,196367	0,1356645
<b>WIKA</b>	1792	0,001089609	0,03056133	0,0009339951	-0,232620	0,2231349

Untuk dapat melakukan identifikasi model dengan menggunakan model SETAR perlu dilakukan pengujian nonlinieritas dari data. Pada penelitian ini digunakan 2 uji nonlinieritas yaitu uji nonlinieritas terasvirta dan uji white yang merupakan uji nonlinieritas *unspecified model*. Hasil dari kedua uji tersebut ditunjukkan pada Tabel 4.45.

Tabel 4.45 Uji Nonlinieritas Data Return Saham

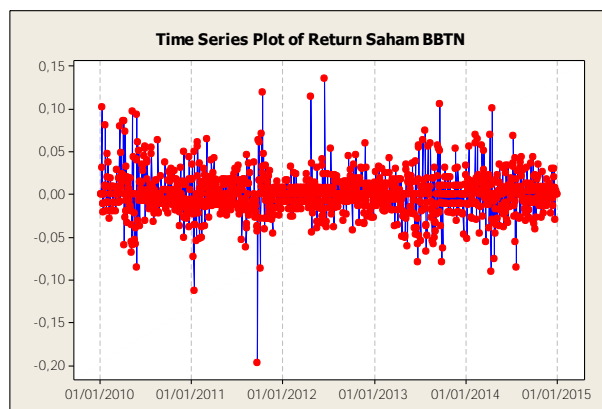
Data	Uji Terasvirta		Uji White	
	F	Pvalue	F	Pvalue
<b>BBTN</b>	4,6784	0,009455	3,2554	0,03889
<b>WIKA</b>	4,4149	0,01223	7,127	0,0008261

Dari hasil yang ditunjukkan pada Tabel 4.45 tampak bahwa data return saham BBTN dan WIKA memenuhi sifat nonlinieritas karena nilai p-value lebih kecil dari  $\alpha$  (5%) untuk kedua uji yang dilakukan. Selain uji nonlinieritas, perlu

dilakukan uji stasioneritas dari data untuk melihat apakah data stasioner dalam mean. Salah satu uji yang dapat dilakukan adalah uji ADF (*Augmented Dickey-Fuller*). Dengan menggunakan uji ADF diperoleh  $p\text{-value}=0,01$ . Berdasarkan hasil yang diperoleh tersebut dapat diambil kesimpulan bahwa data return saham BBTN dan WIKA stasioner dalam mean.

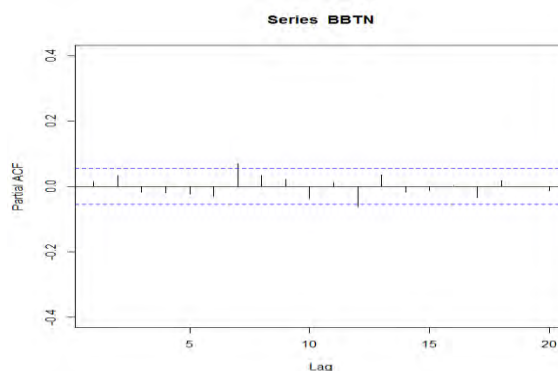
#### 4.2.1 Retun Saham BBTN

Gambar 4.10 menunjukkan plot deret waktu data return saham BBTN. Berdasarkan plot tersebut diketahui bahwa return saham BBTN nilainya disekitar 0. Plot deret waktu tersebut juga menunjukkan terdapat beberapa data yang menyebar dan nilainya jauh dari mean yang dimiliki.



Gambar 4.10 Time Series Plot Return Saham BBTN

Langkah identifikasi model SETAR dilakukan dengan menentukan orde tertinggi yang digunakan. Salah satu cara yaitu menggunakan plot PACF. Plot PACF dapat memberikan informasi orde lag tertinggi yang keluar. Berdasarkan Gambar 4.11, orde tertinggi yang keluar dari batas adalah 12 sehingga pada pemodelan dengan menggunakan SETAR akan digunakan orde tertinggi adalah 12.



Gambar 4.11 Plot PACF Data Return Saham BBTN

Identifikasi model SETAR dilakukan dengan menggunakan metode *GS* dan *GA* untuk mendapatkan hasil estimasi yang paling baik. Kriteria yang digunakan untuk mendapatkan nilai *delay*, *threshold*, orde *AR regime 1*, dan orde *AR regime 2* dengan menggunakan kriteria meminimumkan AIC.

Tabel 4.46 Identifikasi Model SETAR pada Data BBTN

Metode	Region	$d$	$p_1$	$p_2$	<i>Threshold</i>	AIC
<i>GS</i>	15%	4	1	12	0,0092593	-9436,24
<i>GA</i>	5%	3	12	1	-0,03584448	-9450,111
	10%	4	1	12	0,009259325	-9436,236
	15%	4	1	12	0,009259325	-9436,236
	20%	4	1	12	0,009259325	-9436,236

Tabel 4.46 memberikan hasil identifikasi model SETAR pada data return saham BBTN. Berdasarkan hasil kedua metode tersebut diketahui bahwa metode *GA* memberikan hasil AIC lebih kecil dibandingkan dengan hasil AIC metode *GS* ketika region pencarian pada metode *GA* adalah 5%. Untuk region pencarian 10%, 15%, dan 20%, metode *GA* memberikan hasil model yang sama dengan metode *GS*. Estimasi yang dilakukan untuk mendapatkan model pada Tabel 4.47 dilakukan tanpa mempertimbangkan apakah parameter hasil estimasi signifikan atau tidak. Jadi hasil model identifikasi dengan metode *GS* dilanjutkan dengan metode *stepwise* untuk mendapatkan model terbaik dengan parameter yang signifikan.

Selain identifikasi model tanpa memperhatikan parameter yang signifikan, metode *GA* dapat digunakan untuk identifikasi model dengan menghasilkan parameter yang signifikan. Sehingga langsung diperoleh hasil identifikasi model dengan parameter yang merupakan model terbaik. Pada metode ini, daerah

pencarian *threshold* yang digunakan adalah 15%-85%. Ini dilakukan untuk mendapatkan perbandingan hasil antara GA dengan metode *GS-subset*. Hasil identifikasi kedua metode tersebut ditunjukkan pada Tabel 4.47.

Tabel 4.47 Identifikasi *GS* subset dan GA Data return Saham BBTN

Metode	d	$p_1$	$p_2$	Threshold	AIC
<b>GS-subset</b>	4	1	[1,4,7,11,12]	0,0092593	-9426,76
<b>GA</b>	4	[1,3,7,12]	[1,3,4]	-0,0078999	-9444,41

Berdasarkan hasil yang diperoleh pada tahap identifikasi, metode *GA-subset* memberikan hasil AIC lebih kecil dibandingkan dengan menggunakan metode *GS-subset*. Hasil estimasi kedua metode tersebut dapat dilihat pada Tabel 4.48.

Tabel 4.48 Estimasi parameter model subset Data Retun Saham BBTN

Metode	Parameter	Coef	SE Coef	T	P
<b>GS-subset</b>	$\phi_{0,1}$	-0,00033	0,00080952	-0,410	0,682
	$\phi_{1,1}$	-0,04176	0,03291534	-1,269	0,205
	$\phi_{0,2}$	0,006615	0,002433	2,720	0,007
	$\phi_{1,2}$	0,14112	0,05523	2,560	0,011
	$\phi_{4,2}$	-0,15387	0,07181	-2,140	0,033
	$\phi_{7,2}$	0,14004	0,04904	2,860	0,005
	$\phi_{11,2}$	0,09318	0,05845	1,590	0,112
	$\phi_{12,2}$	-0,21471	0,05292	-4,060	0,000
<b>GA</b>	$\phi_{0,1}$	0,0006458	0,0008418	0,767	0,443
	$\phi_{1,1}$	0,1168	0,03639	3,210	0,001
	$\phi_{3,1}$	0,05977	0,03666	1,631	0,103
	$\phi_{7,1}$	0,1045	0,03343	3,124	0,002
	$\phi_{12,1}$	-0,1058	0,03527	-3,000	0,003
	$\phi_{0,2}$	-0,004219	0,001992	-2,118	0,035
	$\phi_{1,2}$	-0,1178	0,04293	-2,745	0,006
	$\phi_{3,2}$	-0,1048	0,0425	-2,465	0,014
	$\phi_{4,2}$	-0,1611	0,06724	-2,396	0,017

Berdasarkan Tabel 4.48 terdapat 3 parameter pada metode *GS* subset yang nilai parameternya tidak signifikan yaitu pada  $\phi_{0,1}$ ,  $\phi_{1,1}$ , dan  $\phi_{11,2}$ . Sedangkan untuk metode *GA*, terdapat 2 parameter yang tidak signifikan yaitu pada  $\phi_{0,1}$  dan  $\phi_{3,1}$ .

Secara matematis, model terbaik yang terbentuk dari data return saham BBTN dengan metode GS-subset dituliskan seperti pada persamaan (4.24) berikut.

$$Z_t = \begin{cases} -0,00033 - 0,04176Z_{t-1} + a_{t,1} & \text{jika } Z_{t-4} \leq 0,0092593 \\ 0,006615 + 0,14112Z_{t-1} - 0,15387Z_{t-4} \\ + 0,14004Z_{t-7} + 0,09318Z_{t-11} - 0,21471Z_{t-12} + a_{t,2} & \text{jika } Z_{t-4} > 0,0092593 \end{cases} \quad (2.51)$$

Model terbaik pada persamaan (4.24) menunjukkan bahwa data masuk *regime* 1 ketika nilai pengamatan  $Z_{t-4} \leq 0,0092593$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-4} > 0,0092593$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar -0,00033 dan data pengamatan sebelumnya sebesar -0,04176 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 0,006615; data pada 1 pengamatan sebelumnya sebesar 0,14112; data pada 3 pengamatan sebelumnya sebesar -0,15387; data pada 4 pengamatan sebelumnya sebesar 0,14004; data pada 11 pengamatan sebelumnya sebesar 0,09318 serta data pada 12 pengamatan sebelumnya sebesar -0,21471.

Untuk model terbaik yang terbentuk dari data return saham BBTN dengan metode GA secara matematis dituliskan seperti pada persamaan (4.25) berikut.

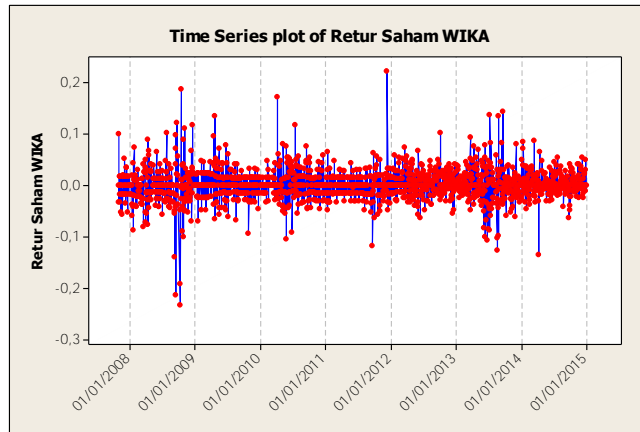
$$Z_t = \begin{cases} 0,0006458 + 0,1168Z_{t-1} + 0,05977Z_{t-3} + 0,1045Z_{t-7} - 0,1058Z_{t-12} + a_{t,1} & \text{jika } Z_{t-4} \leq -0,00789994 \\ -0,004219 - 0,1178Z_{t-1} - 0,1048Z_{t-3} - 0,1611Z_{t-4} + a_{t,2} & \text{jika } Z_{t-4} > -0,00789994 \end{cases} \quad (2.52)$$

Model terbaik pada persamaan (4.25) menunjukkan bahwa data masuk *regime* 1 ketika nilai pengamatan  $Z_{t-4} \leq -0,00789994$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-4} > -0,00789994$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 0,0006458; dan data pada 1 pengamatan sebelumnya sebesar 0,1168; data pada 3 pengamatan sebelumnya sebesar 0,05977; data pada 7 pengamatan sebelumnya sebesar 0,1045; dan data pada 12 pengamatan sebelumnya sebesar -0,1058 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar -0,004219; data pada 1 pengamatan sebelumnya sebesar -0,1178; data pada 3 pengamatan sebelumnya sebesar -0,1048; data pada 4 pengamatan sebelumnya sebesar -0,1048.



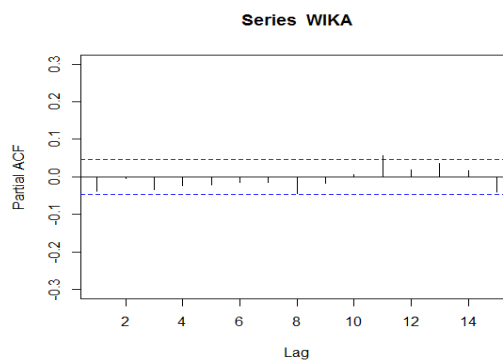
#### 4.2.2 Return Saham WIKA

Gambar 4.12 menunjukkan plot deret waktu data return saham WIKA. Berdasarkan plot tersebut diketahui bahwa return saham WIKA nilainya disekitar 0. Plot deret waktu tersebut juga menunjukkan terdapat beberapa data yang menyebar dan nilainya jauh dari mean yang dimiliki.



Gambar 4.12 Time Series Plot Return Saham WIKA

Langkah identifikasi model SETAR dilakukan dengan menentukan orde tertinggi yang digunakan dengan menggunakan plot PACF. Plot PACF dapat memberikan informasi orde lag tertinggi yang keluar. Berdasarkan Gambar 4.13, orde tertinggi yang keluar dari batas adalah 11 sehingga pada pemodelan dengan menggunakan SETAR digunakan orde tertinggi adalah 11.



Gambar 4.13 Plot PACF Data Return Saham WIKA

Identifikasi model SETAR dilakukan dengan menggunakan metode *GS* dan *GA* untuk mendapatkan hasil estimasi yang paling baik. Kriteria yang digunakan

untuk mendapatkan nilai *delay*, *threshold*, orde AR *regime* 1, dan orde AR *regime* 2 dengan menggunakan kriteria meminimumkan AIC.

Tabel 4.49 Identifikasi Model SETAR pada Data Return Saham WIKA

Metode	Region	D	$p_1$	$p_2$	Threshold	AIC
GS	15%	11	11	3	-0,0201987	-12551,37
GA	5%	2	9	1	-0,0408138	-12569,56
	10%	11	11	3	-0,0201987	-12551,37
	15%	11	11	3	-0,0201987	-12551,37
	20%	11	11	3	-0,0201987	-12551,37

Tabel 4.51 memberikan hasil identifikasi model SETAR pada data return saham WIKA. Berdasarkan hasil kedua metode tersebut diketahui bahwa metode GA memberikan hasil AIC lebih kecil dibandingkan dengan hasil AIC metode GS ketika region pencarian pada metode GA adalah 5%. Untuk region pencarian 10%, 15%, dan 20%, metode GA memberikan hasil model yang sama dengan metode GS. Estimasi yang dilakukan untuk mendapatkan model pada Tabel 4.51 dilakukan tanpa mempertimbangkan apakah parameter hasil estimasi signifikan atau tidak. Jadi hasil model identifikasi dengan metode GS dilanjutkan dengan metode stepwise untuk mendapatkan model terbaik dengan parameter yang signifikan.

Selain identifikasi model tanpa memperhatikan parameter yang signifikan, metode GA dapat digunakan untuk identifikasi model dengan menghasilkan parameter yang signifikan. Sehingga langsung diperoleh hasil identifikasi model dengan parameter yang merupakan model terbaik. Pada metode ini, daerah pencarian *threshold* yang digunakan adalah 15%-85%. Ini dilakukan untuk mendapatkan perbandingan hasil antara GA dengan metode GS-subset. Hasil identifikasi kedua metode tersebut ditunjukkan pada Tabel 4.50.

Tabel 4.50 Identifikasi GS subset dan GA Data return Saham WIKA

Metode	d	$p_1$	$p_2$	Threshold	AIC
GS-subset	11	[4,6,7,8,11]	[1,3]	-0,02009280	-12556,6
GA	8	[1,2,3,4,5,7]	[11]	-0,01780124	-12559,2

Berdasarkan hasil yang diperoleh pada tahap identifikasi, metode GA memberikan hasil AIC lebih kecil dibandingkan dengan menggunakan metode GS-subset. Berdasarkan model yang diperoleh, selanjutnya dilakukan estimasi

parameter dengan menggunakan metode CLS. Hasil estimasi kedua metode tersebut dapat dilihat pada Tabel 4.51.

Berdasarkan Tabel 4.51 terdapat 1 parameter pada metode GS subset yang nilai parameterinya tidak signifikan yaitu pada  $\phi_{0,1}$ . Sedangkan untuk metode GA, terdapat 1 parameter yang tidak signifikan yaitu pada  $\phi_{0,2}$ .

Tabel 4.51 Estimasi parameter model subset Data Retun Saham WIKA

Metode	Parameter	Coef	SE Coef	T	P
GS-Subset	$\phi_{0,1}$	0,006679	0,003576	1,87	0,063
	$\phi_{4,1}$	-0,19617	0,05997	-3,27	0,001
	$\phi_{6,1}$	-0,15096	0,06001	-2,52	0,012
	$\phi_{7,1}$	-0,10681	0,05022	-2,13	0,034
	$\phi_{8,1}$	-0,14362	0,04967	-2,89	0,004
	$\phi_{11,1}$	0,16229	0,07477	2,17	0,031
	$\phi_{0,2}$	0,001704	0,000763	2,23	0,026
	$\phi_{1,2}$	-0,06238	0,02521	-2,47	0,013
	$\phi_{3,2}$	-0,05637	0,02517	-2,24	0,025
GA	$\phi_{0,1}$	0,005637	0,001675	3,366	0,000855
	$\phi_{1,1}$	-0,1623	0,04891	-3,318	0,001011
	$\phi_{2,1}$	-0,1493	0,04952	-3,015	0,002774
	$\phi_{3,1}$	-0,1247	0,04726	-2,64	0,008703
	$\phi_{4,1}$	-0,125	0,04373	-2,86	0,004519
	$\phi_{5,1}$	-0,1087	0,04253	-2,555	0,01109
	$\phi_{7,1}$	-0,1311	0,04027	-3,257	0,001246
	$\phi_{0,2}$	0,000367	0,000791	0,4638	0,6428
	$\phi_{11,2}$	0,07571	0,02711	2,792	0,005302

Secara matematis, model terbaik yang terbentuk dari data return saham WIKA dengan metode GS-subset dituliskan seperti pada persamaan (4.26) berikut.

$$Z_t = \begin{cases} 0,006679 - 0,19617Z_{t-4} - 0,15096Z_{t-6} - 0,10681Z_{t-7} \\ -0,14362Z_{t-8} + 0,16229Z_{t-11} + a_{t,1} & \text{jika } Z_{t-11} \leq -0,02009280 \\ 0,001704 - 0,06238Z_{t-2} - 0,05637Z_{t-3} + a_{t,2} & \text{jika } Z_{t-11} > -0,02009280 \end{cases} \quad (2.53)$$

Model terbaik pada persamaan (4.26) menunjukkan bahwa data masuk *regime* 1 ketika nilai pengamatan  $Z_{t-11} \leq -0,02009280$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-11} > -0,02009280$ . Nilai *forecast* data pada *regime* 1

dipengaruhi oleh konstanta sebesar 0,006679; data pada 4 pengamatan sebelumnya sebesar -0,19617; data pada 6 pengamatan sebelumnya sebesar -0,15096; data pada 7 pengamatan sebelumnya sebesar -0,10681; data pada 8 pengamatan sebelumnya sebesar -0,14362; dan data pada 11 pengamatan sebelumnya sebesar 0,16229 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 0,001704; data pada 2 pengamatan sebelumnya sebesar -0,06238; dan data pada 3 pengamatan sebelumnya sebesar -0,05637.

Secara matematis, model terbaik yang terbentuk dari data return saham WIKA dengan metode GA dituliskan seperti pada persamaan (4.27) berikut.

$$Z_t = \begin{cases} 0,005637 - 0,1623Z_{t-1} - 0,1493Z_{t-2} - 0,12471Z_{t-3} \\ -0,125Z_{t-4} - 0,1087Z_{t-5} - 0,1311Z_{t-7} + a_{t,1} & \text{jika } Z_{t-8} \leq -0,01780124 \\ 0,000367 + 0,07571Z_{t-11} + a_{t,2} & \text{jika } Z_{t-8} > -0,01780124 \end{cases} \quad (2.54)$$

Model terbaik pada persamaan (4.27) menunjukkan bahwa data masuk *regime* 1 ketika nilai pengamatan  $Z_{t-8} \leq -0,01780124$  dan masuk *regime* 2 ketika nilai pengamatan  $Z_{t-8} > -0,01780124$ . Nilai *forecast* data pada *regime* 1 dipengaruhi oleh konstanta sebesar 0,005637; data pada 1 pengamatan sebelumnya -0,1623; data pada 2 pengamatan sebelumnya -0,1493; data pada 3 pengamatan sebelumnya sebesar -0,1247; data pada 5 pengamatan sebelumnya sebesar -0,125; data pada 5 pengamatan sebelumnya sebesar -0,1087 dan data pengamatan sebelumnya sebesar -0,1311 sedangkan untuk *forecast* data pada *regime* 2 dipengaruhi oleh konstanta sebesar 0,000367 dan data pada 12 pengamatan sebelumnya sebesar 0,07571.

### 4.3 Perbandingan Metode pada Data Out Sample

Hasil estimasi dari masing-masing model selanjutnya dilakukan untuk peramalan pada data *out sample* pada data return saham BBTN dan return saham WIKA. Hasil peramalan dari masing-masing model ditunjukkan pada Tabel 4.52 dan Tabel 4.53.

Tabel 4.52 Perbandingan kebaikan model Data Retun Saham BBTN

<b>Metode</b>	<b>AIC (in sample)</b>	<b>MSE (out sample)</b>
<b>GS subset</b>	-9,4228E+03	5,7261E-04
<b>GA</b>	-9,4444E+03	5,5312E-04

Berdasarkan hasil pada Tabel 4.52 diketahui bahwa untuk data return saham BBTN, metode GA memberikan hasil AIC lebih kecil dibandingkan dengan metode GS subset pada data in sample. Untuk data out sample, metode GA memberikan hasil peramalan dengan MSE lebih kecil dibandingkan dengan metode GS subset.

Tabel 4.53 Perbandingan kebaikan model Data Retun Saham WIKA

<b>Metode</b>	<b>AIC (in samlpe)</b>	<b>MSE (out sample)</b>
<b>GS-subset</b>	-12556,6	2,209E-04
<b>GA</b>	-12559,2	2,118E-04

Berdasarkan Tabel 4.53 diketahui bahwa pada data in sample, metode GA memberikan hasil AIC lebih kecil dibandingkan metode GS subset. Untuk data out sample, metode GA memberikan hasil MSE yang lebih kecil dibandingkan dengan metode GS subset.

## **BAB 5**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Berdasarkan hasil penelitian ini, diperoleh kesimpulan sebagai berikut:

1. Identifikasi data simulasi model SETAR(2,1,1) dengan menggunakan metode *GS* dan *GA* memberikan hasil yang sama. Hal ini dikarenakan metode *GS* maupun metode *GA* menggunakan pendekatan yang sama untuk mendapatkan nilai *AIC* terkecil. Untuk estimasi parameter yang dilakukan dengan menggunakan metode *CLS* dan *GA* juga memberikan hasil *SSE* yang hampir sama. Perbedaan hasil estimasi pada metode *CLS* dan *GA* terletak pada *threshold* yang diperoleh. Untuk model subset SETAR, identifikasi model yang dilakukan dengan metode *GA* memberikan hasil identifikasi yang lebih sesuai dibandingkan dengan metode *GS*. Selain itu, hasil identifikasi dari metode *GA* menghasilkan nilai *AIC* yang lebih kecil dibandingkan hasil identifikasi dari metode *GS*.
2. Hasil yang diperoleh dari identifikasi model SETAR pada data return saham BBTN dan WIKA dengan menggunakan metode *GA* memberikan hasil *AIC* yang lebih kecil dibandingkan metode *GS*.
3. Penerapan hasil identifikasi model SETAR pada data *out sample* return saham BBTN dan WIKA dengan menggunakan metode *GA* memberikan hasil *MSE* lebih kecil dibandingkan dengan menggunakan metode *GS*.

#### **5.2 Saran**

Saran yang dapat diberikan berdasarkan penelitian ini antara lain sebagai berikut:

1. Membangun model SETAR-GA dengan *regime* lebih dari 2 agar diperoleh pendekatan yang lebih baik dengan *MSE* yang lebih kecil.
2. Melakukan simulasi dengan model subset SETAR untuk menunjukkan kelebihan dari subset SETAR-GA.
3. Melakukan identifikasi model Subset SETAR-GA dengan *regime* lebih dari 2.

4. Melakukan identifikasi model SETARMA-GA untuk dapat menangkap fenomena yang mungkin terjadi pada model MA.

## DAFTAR PUSTAKA

- Baragona, R., Battaglia, F. & Cucina, D., 2004. Estimating Threshold Subset Autoregressive Moving-average Models by Genetic Algorithms. *METRON-International Journal of Statistics*, Volume LXII, pp. 39-61.
- Baragona, R., Battaglia, F. & Cucina, D., 2004. Fitting piecewise linear threshold autoregressive models by means of genetic algorithms. *Computational Statistics & Data Analysis* 47, p. 277–295.
- Dewi, S. K. & Purnomo, H., 2005. *Penyelesaian Masalah Optimasi dengan Teknik-teknik Heuristik*. Yogyakarta: Graha Ilmu.
- Franses, P. H. & Dijk, D. V., 2003. *Non-Linear Time Series In Empirical Finance*. New York: Cambridge University Press.
- Gujarati, D. N. & Porter, D. C., 2012. *Dasar-dasar ekonometrika*. Jakarta: Salemba Empat.
- Lee, T.-H., White, H. & Granger, C. W., 1993. Testing for Neglected Nonlinearity in Time Series Model. *Journal of Econometrics*, pp. 269-290.
- Santosa, B. & Willy, P., 2011. *Metode Metaheuristik Konsep dan Implementasi*. Surabaya: Guna Widya.
- Sawaka, M., 2002. *Genetic Algorithms and Fuzzy Multiobjective Optimization*. Boston: Kluwer Academic Publishers.
- Strikholm, B. & Terasvirta, T., 2006. A Sequential Procedure For Determining The Number of Regimes. *Econometrics journal*, Volume 9, pp. 472-491.
- Tandelilin, E., 2010. *Portofolio dan Investasi*. Yogyakarta: Universitas Gadjah Mada.
- Terasvirta, T., Lin, C.-F. & Granger, C. W., 1993. Power of The Neural Network Linearity Test. *Journal of Time Series Analysis*, Volume 14, pp. 209-220.
- Tong, H. & Lim, K. S., 1980. Threshold Autoregression, Limit Cycles and Cyclical Data. *Journal of the Royal Statistical Society Series B*, Volume 43, pp. 245-292.



- Tsay, R. S., 1989. Testing and Modeling Threshold Autoregressive Process. *Journal of American Statistical Association*, Volume 84 NO 405, pp. 231-241.
- Wei, W. W. S., 2006. *Time Series Analysis Univariate and Multivariate Methods*. Second Edition penyunt. New York: Pearson.
- Wu, B. & Chang, C. L., 2002. Using genetik algoritms to parameter (d,r) estimation for threshold autoregressive model. *Computational statistics and data analysis*, Volume 38, pp. 315-330.

## DAFTAR LAMPIRAN

	halaman
<b>Lampiran 1</b> Data Saham dan Return Saham BBTN.....	101
<b>Lampiran 2</b> Data Return Saham dan Return Saham WIKA.....	102
<b>Lampiran 3</b> Sintax Simulasi Data Model SETAR (2,1,1).....	103
<b>Lampiran 4</b> Sintax Simulasi Data Model Subset SETAR.....	104
<b>Lampiran 5</b> Sintax Uji Nonlinier, Identifikasi Model, dan Estimasi Parameter Model SETAR dengan metode <i>Grid search</i> .....	105
<b>Lampiran 6</b> Identifikasi Model Simulasi SETAR dengan Algoritma Genetika (Halaman 1 dari 2 halaman).....	106
<b>Lampiran 7</b> Identifikasi Model Subset SETAR (Halaman 1 dari 3 halaman)....	108
<b>Lampiran 8</b> Identifikasi Model SETAR Data Return Saham dengan GA (Halaman 1 dari 2 halaman) .....	111
<b>Lampiran 9</b> Estimasi Parameter Data Simulasi SETAR dengan Algoritma Genetika.....	113
<b>Lampiran 10</b> Sintax Estimasi Parameter Data Return Saham dengan Algoritma Genetika .....	114
<b>Lampiran 11</b> Hasil Uji Terasvirta dan White pada Data Return Saham WIKA .....	115
<b>Lampiran 12</b> Hasil Identifikasi dan Estimasi Parameter Model SETAR pada Data Simulasi Model 1 untuk n=200 dengan Metode <i>Grid Search</i> .....	116
<b>Lampiran 13</b> Hasil Identifikasi dan Estimasi Parameter Model SETAR pada Data Simulasi Model Subset SETAR 1 untuk n=500 dengan Metode GA.....	117
<b>Lampiran 14</b> Hasil Identifikasi Data Return Saham BBTN dengan <i>GS</i> .....	118
<b>Lampiran 15</b> Hasil Identifikasi dan Estimasi Data Return Saham BBTN dengan Metode GA .....	119
<b>Lampiran 16</b> Contoh Identifikasi Model Dengan Metode GA.....	120
<b>Lampiran 17</b> Contoh Estimasi Parameter Model Dengan Metode GA.....	126

## Lampiran 1 Data Saham dan Return Saham BBTN

### Data In Sample

No	Tanggal	Adj Close	Return
1	04/01/2010	731	0
2	05/01/2010	731	0
3	06/01/2010	809.32	0.101781
4	07/01/2010	809.32	0
5	08/01/2010	835.43	0.031752
6	11/01/2010	818.02	-0.02106
7	12/01/2010	809.32	-0.01069
.	.	.	.
.	.	.	.
.	.	.	.
1269	25/12/2014	1200	0
1270	26/12/2014	1200	0
1271	29/12/2014	1205	0.004158
1272	30/12/2014	1205	0
1273	31/12/2014	1205	0

### Data Out Sample

No	Tanggal	Adj Close	Return
1	01/01/2015	1205	0
2	02/01/2015	1225	0.016461
3	05/01/2015	1220	-0.00409
4	06/01/2015	1195	-0.0207
5	07/01/2015	1205	0.008333
6	08/01/2015	1210	0.004141
.	.	.	.
.	.	.	.
.	.	.	.
37	20/02/2015	1055	0.009524
38	23/02/2015	1045	-0.00952
39	24/02/2015	1050	0.004773
40	25/02/2015	1050	0
41	26/02/2015	1060	0.009479
42	27/02/2015	1070	0.00939

## Lampiran 2 Data Return Saham dan Return Saham WIKA

### Data In sample

No	Tanggal	Adj Close	Return
1	31/10/2007	577.2	0.1001
2	01/11/2007	558.87	-0.03227
3	02/11/2007	558.87	0
4	05/11/2007	549.71	-0.01653
5	06/11/2007	558.87	0.016526
6	07/11/2007	540.55	-0.03333
7	08/11/2007	540.55	0
.	.	.	.
.	.	.	.
.	.	.	.
1788	25/12/2014	3535	0
1789	26/12/2014	3535	0
1790	29/12/2014	3720	0.05101
1791	30/12/2014	3680	-0.01081
1792	31/12/2014	3680	0

### Data Out Sample

No	Tanggal	Adj Close	Return
1	01/01/2015	3659.38	0
2	02/01/2015	3654.41	-0.00136
3	05/01/2015	3554.97	-0.02759
4	06/01/2015	3535.08	-0.00561
5	07/01/2015	3629.55	0.026373
6	08/01/2015	3659.38	0.008185
.	.	.	.
.	.	.	.
.	.	.	.
38	23/02/2015	3659.38	0.012301
39	24/02/2015	3629.55	-0.00819
40	25/02/2015	3649.44	0.005465
41	26/02/2015	3654.41	0.001361
42	27/02/2015	3639.5	-0.00409

### Lampiran 3 Sintax Simulasi Data Model SETAR (2,1,1)

#### Model 1

```
set.seed(132357966)
TvarMat<-c(8,-0.2,8,0.2)
sim<-setar.sim(B=TvarMat,lag=1,
n=400,type="simul",nthresh=1,Thresh=8,sd=2)$serie
x1=as.ts(sim[201:400])
```

#### Model 2

```
set.seed(1234579)
TvarMat<-c(1,0.2,4,0.2)
sim<-setar.sim(B=TvarMat,lag=1,
n=400,type="simul",nthresh=1,Thresh=3,sd=2)$serie
x1=as.ts(sim[201:400])
```

#### Model 3

```
set.seed(12346)
TvarMat <- c(5,0.5,9,-0.5)
sim<-setar.sim(B=TvarMat,lag=1,
n=400,type="simul",nthresh=1,Thresh=6,sd=2)$serie
x1=as.ts(sim[201:400])
```

#### Model 4

```
set.seed(123489)
TvarMat <- c(8,-0.8,9,-0.2)
sim<-setar.sim(B=TvarMat,lag=1,
n=400,type="simul",nthresh=1,Thresh=6,sd=2)$serie
x1=as.ts(sim[201:400])
```

#### Model 5

```
set.seed(1234189)
TvarMat <- c(-7,-0.5,-2,0.8)
sim<-setar.sim(B=TvarMat,lag=1,
n=400,type="simul",nthresh=1,Thresh=-2.5,sd=2)$serie
x1=as.ts(sim[201:400])
```

## Lampiran 4 Sintax Simulasi Data Model Subset SETAR

### Model 1

```
set.seed(123457)
TvarMat <- c(0,-1.2,-0.7,0,0,0,0,0.8)
sim<-setar.sim(B=TvarMat,lag=3,
n=400,type="simul",nthresh=1,Thresh=0,sd=1)$serie
x1=as.ts(sim[201:400])
```

### Model 2

```
set.seed(12346978)
TvarMat <- c(0,-1.2,-0.7,0,0,0,0,0.8)
sim<-setar.sim(B=TvarMat,lag=3,
n=400,type="simul",nthresh=1,Thresh=0,sd=1,thDelay=1)
x1=as.ts(sim[201:400])
```

### Model 3

```
set.seed(123456474)
TvarMat <- c(0,0,-0.6,0,0,0.75,0,-0.5)
sim<-setar.sim(B=TvarMat,lag=3,
n=400,type="simul",nthresh=1,Thresh=0,sd=1)$serie
x1=as.ts(sim[201:400])
```

### Model 4

```
set.seed(12347477)
TvarMat <- c(0,0,-0.6,0,0,0.75,0,-0.5)
sim<-setar.sim(B=TvarMat,lag=3,
n=400,type="simul",nthresh=1,Thresh=0,sd=1,thDelay=1)
x1=as.ts(sim[201:400])
```

## Lampiran 5 Syntax Uji Nonlinier, Identifikasi Model, dan Estimasi Parameter Model SETAR dengan metode *Grid search*

### **Uji non linieritas Terasvirta dan White**

```
terasvirta.test(x, type ="F")  
white.test(x, type ="F")
```

### **Identifikasi Model dengan AIC terkecil**

```
selectSETAR(x,m=1,thDelay=0,criterion="AIC")  
%m menunjukkan orde maksimum dari model SETAR
```

### **Estimasi Parameter**

```
hasil=setar(x,mL=1,mH=1)  
summary(hasil)  
% mL menyatakan orde pada regime lower dan mH menyatakan orde  
pada regime Upper. Kedua orde ini diperoleh dari hasil  
identifikasi yang dilakukan sebelumnya.
```

**Lampiran 6** Identifikasi Model Simulasi SETAR dengan Algoritma Genetika  
(Halaman 1 dari 2 halaman)

```
esimul<-function(b)
{
  d=1
  p1=1
  p2=1
  r=b[1]*2^(7*b[1])+b[2]*2^(6*b[2])+b[3]*2^(5*b[3])+b[4]*2^(4*
b[4])+b[5]*2^(3*b[5])+b[6]*2^(2*b[6])+b[7]*2^(1*b[7])+2^b[8]
  data=sort(x)
  n=length(x)
  bb=round(n*0.2)
  bagi=256/(n*0.6)
  ambil=bb+round(r/bagi)
  thr=data[ambil]

  d.klasu=rep(0,n)
  d.klasl=rep(0,n)
  m=n-d
  for (i in 1:m)
  {
    if(x[i]>thr){
      d.klasu[i+d]=1
      d.klasl[i+d]=0
    }
    else{
      d.klasu[i+d]=0
      d.klasl[i+d]=1
    }
  }
  t1=x*d.klasu
  t2=x*d.klasl
  u.lag=matrix(0,n-p2,p2+1)
  for (j in 1:p2)
  {
    a=(p2+1)-j
    b=n-j
    u.lag[,j]=x[a:b]
  }
  l.lag=matrix(0,n-p1,p1+1)
  for (j in 1:p1)
  {
    a=(p1+1)-j
    b=n-j
    l.lag[,j]=x[a:b]
  }
}
```



**Lampiran 6** Identifikasi Model Simulasi SETAR dengan Algoritma Genetika  
(Halaman 2 dari 2 halaman)

```
a1=p2+1
b1=p1+1
u.lag[,p2+1]=d.klasu[a1:n]
l.lag[,p1+1]=d.klasl[b1:n]
y1=t1[a1:n]
y2=t2[b1:n]
klasub=u.lag[,p2+1]
klaslb=l.lag[,p1+1]
model1=u.lag*klasub
model2=l.lag*klaslb
  if ((d>p1) && (d>p2)) {
    m1=length(klasub)
    m2=length(klaslb)
    a1=abs(p2-d)+1
    a2=abs(p1-d)+1
    model1=model1[a1:m1,]
    y1=y1[a1:m1]
    model2=model2[a2:m2,]
    y2=y2[a2:m2]
  }else{
    if (p1>p2) {
      m1=length(klasub)
      a1=abs(p1-p2)+1
      model1=model1[a1:m1,]
      y1=y1[a1:m1]
    }else{
      m2=length(klaslb)
      a2=abs(p1-p2)+1
      model2=model2[a2:m2,]
      y2=y2[a2:m2]
    }
  }
OLS1=lm.fit(model1,y1)
OLS2=lm.fit(model2,y2)
class(OLS1)="lm"
class(OLS2)="lm"
beta1=coefficients(OLS1)
beta2=coefficients(OLS2)
resi1=y1-(model1%%beta1)
resi2=y2-(model2%%beta2)
n1=sum(model1[,p2+1])
n2=sum(model2[,p1+1])
sse1=sum(resi1^2)
sse2=sum(resi2^2)
AIC=n*log((sse1+sse2)/n)+2*(p1+p2+3)
-AIC
}
```

## Lampiran 7 Identifikasi Model Subset SETAR (Halaman 1 dari 3 halaman)

```
subset2<-function(b)
{
  d=b[9]+b[10]+1
  p1=b[11]+b[12]+1
  p2=b[13]+b[14]+1
  r=b[1]*2^(7*b[1])+b[2]*2^(6*b[2])+b[3]*2^(5*b[3])+b[4]*2^(4*
b[4])+b[5]*2^(3*b[5])+b[6]*2^(2*b[6])+b[7]*2^(1*b[7])+2^b[8]
+b[23]*2^(8*b[23])
  sub1=c(b[21],b[15],b[16],b[17])
  sub2=c(b[22],b[18],b[19],b[20])
  bsub1=sub1[1:(p1+1)]
  bsub2=sub2[1:(p2+1)]

  data=sort(x)
  n=length(x)
  bb=round(n*0.05)
  bagi=512/(n*0.9)
  ambil=bb+round(r/bagi)
  thr=data[ambil]

  d.klasu=rep(0,n)
  d.klasl=rep(0,n)
  m=n-d
  for (i in 1:m)
  {
    if(x[i]>thr){
      d.klasu[i+d]=1
      d.klasl[i+d]=0
    }
    else{
      d.klasu[i+d]=0
      d.klasl[i+d]=1
    }
  }
  t1=x*d.klasu
  t2=x*d.klasl
  u.lag=matrix(0,n-p2,p2+1)
  for (j in 1:p2)
  {
    a=(p2+1)-j
    b1=n-j
    u.lag[,j]=x[a:b1]
  }
  l.lag=matrix(0,n-p1,p1+1)
```

**Lampiran 7** Identifikasi Model Subset SETAR (Halaman 2 dari 3 halaman)

```
for (j in 1:p1)
{
  a=(p1+1)-j
  b1=n-j
  l.lag[,j]=x[a:b1]
}
a1=p2+1
b1=p1+1
u.lag[,2:a1]=u.lag[,1:p2]
l.lag[,2:b1]=l.lag[,1:p1]
u.lag[,1]=d.klasu[a1:n]
l.lag[,1]=d.klasl[b1:n]
y1=t1[a1:n]
y2=t2[b1:n]

klasub=u.lag[,1]
klaslb=l.lag[,1]
model1=u.lag*klasub
model2=l.lag*klaslb

if((d>p1)&&(d>p2)){
  m1=length(klasub)
  m2=length(klaslb)
  a1=abs(p2-d)+1
  a2=abs(p1-d)+1
  model1=model1[a1:m1,]
  y1=y1[a1:m1]
  model2=model2[a2:m2,]
  y2=y2[a2:m2]
}else{
  if (p1>p2){
    m1=length(klasub)
    a1=abs(p1-p2)+1
    model1=model1[a1:m1,]
    y1=y1[a1:m1]
  }else{
    m2=length(klaslb)
    a2=abs(p1-p2)+1
    model2=model2[a2:m2,]
    y2=y2[a2:m2]
  }
}

n1=sum(model1[,1])
n2=sum(model2[,1])
model1=t(t(model1)*bsub2)
model2=t(t(model2)*bsub1)
```

## Lampiran 7 Identifikasi Model Subset SETAR (Halaman 3 dari 3 halaman)

```
OLS1=lm.fit(modell1,y1)
OLS2=lm.fit(model2,y2)

class(OLS1)="lm"
class(OLS2)="lm"
beta1=rep(0,(p2+1))
beta2=rep(0,(p1+1))
bta1=coefficients(OLS1)
bta2=coefficients(OLS2)
kk1=p2+1
kk2=p1+1
for (i in 1:kk1)
{
  if(bsub2[i]==0){
    beta1[i]=0
  }
  else{
    beta1[i]=bta1[i]
  }
}
for (i in 1:kk2)
{
  if(bsub1[i]==0){
    beta2[i]=0
  }
  else{
    beta2[i]=bta2[i]
  }
}
resi1=y1-(modell1%%beta1)
resi2=y2-(model2%%beta2)
jmlh=n1+n2
sse1=sum(resi1^2)
sse2=sum(resi2^2)
AIC=n*log((sse1+sse2)/n)+2*(sum(bsub1)+sum(bsub2)+3)
-AIC
}
```

**Lampiran 8** Identifikasi Model SETAR Data Return Saham dengan GA (Halaman 1 dari 2 halaman)

```
FSETAR1<-function(b)
{
d=b[12]*2^(2*b[12])+b[13]*2^(1*b[13])+2^b[14]+b[15]+b[16]+b[17]+b[18]
p1=b[19]*2^(2*b[19])+b[20]*2^(1*b[20])+2^b[21]+b[22]+b[23]+b[24]+b[25]
p2=b[26]*2^(2*b[26])+b[27]*2^(1*b[27])+2^b[28]+b[29]+b[30]+b[31]+b[32]
r=b[1]*2^(10*b[1])+b[2]*2^(9*b[2])+b[3]*2^(8*b[3])+b[4]*2^(7*b[4])+b[5]*2
^(6*b[5])+b[6]*2^(5*b[6])+b[7]*2^(4*b[7])+b[8]*2^(3*b[8])+b[9]*2^(2*b[9])
+b[10]*2^(1*b[10])+2^b[11]
data=sort(x)
n=length(x)
bb=round(n*0.07)
bagi=2048/(n*0.9)
ambil=bb+round(r/bagi)
thr=data[ambil]

d.klasu=rep(0,n)
d.klasl=rep(0,n)
m=n-d
for (i in 1:m)
{
  if(x[i]>thr){
    d.klasu[i+d]=1
    d.klasl[i+d]=0
  }
  else{
    d.klasu[i+d]=0
    d.klasl[i+d]=1
  }
}
t1=x*d.klasu
t2=x*d.klasl
u.lag=matrix(0,n-p2,p2+1)
for (j in 1:p2)
{
  a=(p2+1)-j
  b=n-j
  u.lag[,j]=x[a:b]
}
l.lag=matrix(0,n-p1,p1+1)
for (j in 1:p1)
{
  a=(p1+1)-j
  b=n-j
  l.lag[,j]=x[a:b]
}
```

**Lampiran 8** Identifikasi Model SETAR Data Return Saham dengan GA (Halaman 2 dari 2 halaman)

```
a1=p2+1
b1=p1+1
u.lag[,p2+1]=d.klasu[a1:n]
l.lag[,p1+1]=d.klasl[b1:n]
y1=t1[a1:n]
y2=t2[b1:n]
klasub=u.lag[,p2+1]
klaslb=l.lag[,p1+1]
model1=u.lag*klasub
model2=l.lag*klaslb
if((d>p1)&&(d>p2)){
  m1=length(klasub)
  m2=length(klaslb)
  a1=abs(p2-d)+1
  a2=abs(p1-d)+1
  model1=model1[a1:m1,]
  y1=y1[a1:m1]
  model2=model2[a2:m2,]
  y2=y2[a2:m2]
}else{
  if(p1>p2){
    m1=length(klasub)
    a1=abs(p1-p2)+1
    model1=model1[a1:m1,]
    y1=y1[a1:m1]
  }else{
    m2=length(klaslb)
    a2=abs(p1-p2)+1
    model2=model2[a2:m2,]
    y2=y2[a2:m2]
  }
}
OLS1=lm.fit(model1,y1)
OLS2=lm.fit(model2,y2)
class(OLS1)="lm"
class(OLS2)="lm"
beta1=coefficients(OLS1)
beta2=coefficients(OLS2)
resi1=y1-(model1%*%beta1)
resi2=y2-(model2%*%beta2)
n1=sum(model1[,p2+1])
n2=sum(model2[,p1+1])
sse1=sum(resi1^2)
sse2=sum(resi2^2)
mse1=sse1/(n1-p2-1)
mse2=sse2/(n2-p1-1)
AIC=n*log((sse1+sse2)/n)+2*(p1+p2+3)
-AIC
```

## Lampiran 9 Estimasi Parameter Data Simulasi SETAR dengan Algoritma Genetika

```
function sd=estimasi(z,x)
p1=1;p2=1;d=1;
a11=max(d,p1);
a1=max(a11,p2);
a2=max(p1,p2);
n=length(x);
y=x;
xt=zeros(n-a1,a2);
yp=zeros(n,1);
for j=1:a2
    a3=(a1+1)-j;
    a4=n-j;
    xt(:,j)=x(a3:a4);
end
for i=1:a2
    yp(i)=0;
end
a5=a2+1;
for i=a5:n
    if y(i)>z(1)
        yp(i)=z(2)+z(3)*xt(i-1);
    else
        yp(i)=z(4)+z(5)*xt(i-1);
    end
end
erro=y-yp;
sd=sum((erro).^2);
end
```

### Running Program Algoritma Genetika

```
FitnessFunction = @(z) estimasi(z,d,p1,p2,x);
numberOfVariables = 5;
fminuncOptions = optimset('Display','iter', 'LargeScale',
'off')
options=gaoptimset('PopulationSize',3000,'CrossoverFcn',
@crossoverarithmetic, 'SelectionFcn',@selectionroulette,
'HybridFcn',{@fminunc, fminuncOptions})
[z,fval]=ga(FitnessFunction,numberOfVariables,[],[],[],[],[],[]
,[],options)
```

## Lampiran 10 Sintax Estimasi Parameter Data Return Saham dengan Algoritma Genetika

```
function sd=gawika(z,d,x)
format long
p1=9;p2=3;
a11=max(d,p1);
a1=max(a11,p2);
a2=max(p1,p2);
n=length(x);
y=x;
xt=zeros(n-a1,a2);
yp=zeros(n,1);
for j=1:a2
    a3=(a1+1)-j;
    a4=n-j;
    xt(:,j)=x(a3:a4);
end
for i=1:a2
    yp(i)=y(i);
end
a5=a2;
m=n-a5;
k=0;
l=0;
for i=1:m
    if y(i+7)<=z(1)
        yp(i+a5)=z(2)+z(3)*xt(i,1)+z(4)*xt(i,2)+z(5)*xt(i,3)
            +z(6)*xt(i,4)+z(7)*xt(i,5)+z(8)*xt(i,6)+z(9)*xt(i,7)
            +z(10)*xt(i,8)+z(11)*xt(i,9);
    else
        yp(i+a5)=z(12)+z(13)*xt(i,1);
    end
end
ero=y-yp;
sd=sum(ero.^2);
end
```

```
d=2;k1=sort(x);
n=length(x)
k2=round(0.05*n);
k3=round(0.95*n);
FitnessFunction = @(z) gawika(z,d,x);
numberOfVariables = 13;
LB=[k1(k2)-100-100-100-100-100-100-100-100-100 -100 -100 -100];
UB=[k1(k3) 100 100 100 100 100 100 100 100 100 100 100 100];
fminuncOptions = optimset('Display','iter',
    'LargeScale','off');
options=gaoptimset('Generation',2000,'PopulationSize',100,'CrossoverFcn',@crossoverarithmetic,
    'SelectionFcn',@selectionroulette,'TolFun',1e-
    10,'PlotFcns',{@gaplotbestf,
    @gaplotbestindiv},'HybridFcn',{fminunc, fminuncOptions});
[z,fval]=ga(FitnessFunction,numberOfVariables,[],[],[],[],LB,UB,
    [],options)
```



## Lampiran 11 Hasil Uji Terasvirta dan White pada Data Return Saham WIKA

```
> terasvirta.test(x1,type="F");white.test(x1,type="F")
Teraesvirta Neural Network Test
data:  x1
F = 4.4149, df1 = 2, df2 = 1789,p-value = 0.01223

White Neural Network Test
data:  x1
F = 7.127, df1 = 2, df2 = 1789, p-value= 0.0008261

> terasvirta.test(x2,type="F");white.test(x2,type="F")
Teraesvirta Neural Network Test
data:  x2
F = 4.6784, df1 = 2, df2 = 1270,p-value = 0.009455

White Neural Network Test
data:  x2
F = 3.2554, df1 = 2, df2 = 1270,p-value = 0.03889

> terasvirta.test(x3,type="F");white.test(x3,type="F")
Teraesvirta Neural Network Test
data:  x3
F = 25.463, df1 = 2, df2 = 1616,p-value = 1.295e-11

White Neural Network Test
data:  x3
F = 20.9659, df1 = 2, df2 = 1616,p-value = 1.025e-09
```

```
> terasvirta.test(x1);white.test(x1)
      Teraesvirta Neural Network Test
data:  x1
X-squared = 8.8229, df = 2, p-value =0.01214
      White Neural Network Test
data:  x1
X-squared = 13.5365, df = 2, p-value =
0.00115
```

**Lampiran 12 Hasil Identifikasi dan Estimasi Parameter Model SETAR pada Data Simulasi Model 1 untuk n=200 dengan Metode *Grid Search***

```

> selectSETAR(x1,m=1,thDelay=0,criterion="AIC")
Using maximum autoregressive order for low regime: mL = 1
Using maximum autoregressive order for high regime: mH = 1
Searching on 349 possible threshold values within regimes with
sufficient ( 15% ) number of observations
Searching on 349 combinations of thresholds ( 349 ), thDelay
( 1 ), mL ( 1 ) and MM ( 1 )
Results of the grid search for 1 threshold
      thDelay mL mH      th      AIC
1         0  1  1 3.051782 672.6851
2         0  1  1 3.038420 674.5683
3         0  1  1 3.013915 675.7950
4         0  1  1 2.993519 676.8504
5         0  1  1 3.058609 676.9480
6         0  1  1 2.897315 678.4290
7         0  1  1 3.115771 679.2352
8         0  1  1 2.890254 679.2722
9         0  1  1 2.893701 679.3536
10        0  1  1 2.869115 680.5731
> hasil=setar(x1,mL=1,mH=1)
> summary(hasil)
Non linear autoregressive model
SETAR model ( 2 regimes)
Coefficients:
Low regime:
  const.L   phiL.1
0.9944730 0.2360398

High regime:
  const.H   phiH.1
3.9550215 0.1916919

Threshold:
-Variable: Z(t) = + (1) X(t)
-Value: 3.052
Proportion of points in low regime: 47.9%           High regime:
52.1%
Residuals:
      Min       1Q   Median       3Q      Max
-6.40772 -1.28427  0.13586  1.23634  5.24606
Fit:
residuals variance = 3.764,  AIC = 673,  MAPE = 112.8%
Coefficient(s):
      Estimate Std. Error  t value  Pr(>|t|)
const.L  0.994473   0.141035   7.0513 5.981e-12 ***
phiL.1   0.236040   0.076379   3.0904 0.002111 **
const.H  3.955021   0.448615   8.8161 < 2.2e-16 ***
phiH.1   0.191692   0.081279   2.3584 0.018740 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Threshold
Variable: Z(t) = + (1) X(t)

Value: 3.052

```

**Lampiran 13** Hasil Identifikasi dan Estimasi Parameter Model SETAR pada Data Simulasi Model Subset SETAR 1 untuk  $n=500$  dengan Metode GA

```

source("E:\\TESIS AQ\\dikirim\\subset2.r")
source("E:\\TESIS AQ\\dikirim\\cobasubset2.r")

# Algoritma genetika
GA<-ga("binary",maxiter=100,fitness
=subset2,nBits=23,popSize=100, monitor = plot)
plot(GA)
summary(GA)
+-----+
|           Genetic Algorithm           |
+-----+

GA settings:
Type                = binary
Population size     = 100
Number of generations = 100
Elitism              = 5
Crossover probability = 0.8
Mutation probability = 0.1

GA results:
Iterations           = 100
Fitness function value = 33,7394
Solutions            =
x1  x2  x3  x4  x5  x6  x7  x8  x9  x10 x11 x12 x13 x14 x15 x16
 1   1   0   1   0   1   1   1   0   0   1   0   1   1   1   1
x17 x18 x19 x20 x21 x22 x23
 0   0   0   1   0   0   0
> cobasubset2(b)
=====
model SETAR( 2 , 2 , 3 ) dengan nilai threshold adalah 0.01980077
delay 1
=====
hasil estimasi parameter regime atas
=====
          coefficient      SE coef      t-hitung      p-value
lag-3 0.7866              0.03849      20.43         1.66E-57

banyak data pada regime atas adalah 285

=====
hasil estimasi parameter regime bawah
=====
          coefficient      SE coef      t-hitung      p-value
lag-1  -1.296            0.04935     -26.26        3.97E-68
lag-2  -0.6841           0.04213     -16.24        6.71E-39
banyak data pada regime bawah adalah 212

```

## Lampiran 14 Hasil Identifikasi Data Return Saham BBTN dengan Metode GS

```
> selectSETAR(x,m=12,thDelay=0:11,criterion="AIC")
Using maximum autoregressive order for low regime: mL = 12
Using maximum autoregressive order for high regime: mH = 12
Searching on 512 possible threshold values within regimes with
sufficient ( 15% ) number of observations
Searching on 884736 combinations of thresholds ( 512 ), thDelay
( 12 ), mL ( 12 ) and MM ( 12 )
Results of the grid search for 1 threshold
  thDelay mL mH          th          AIC
1         4  1 12  0.009259325 -9436.236
2         3 12  1 -0.016949187 -9436.029
3         3 12  1 -0.016899165 -9435.807
4         3 12  1 -0.014818300 -9435.757
5         4  1 12  0.009218664 -9435.706
6         4  1 12  0.009054327 -9435.505
7         4  1 12  0.009090972 -9435.450
8         3 12  1 -0.016810777 -9435.430
9         3 12  1 -0.014184323 -9435.369
10        4  1 12  0.009129992 -9435.290
```

## Lampiran 15 Hasil Identifikasi dan Estimasi Data Return Saham BBTN dengan Metode GA

```

GA <- ga("binary",maxiter=500,fitness =subset,nBits=56,
popSize=100,monitor = plot)
> plot(GA)
> summary(GA)
+-----+
|           Genetic Algorithm           |
+-----+

GA settings:
Type                = binary
Population size     = 100
Number of generations = 500
Elitism             = 5
Crossover probability = 0.8
Mutation probability = 0.1
GA results:
Iterations          = 500
Fitness function value = -9444.409
Solutions           =
x1  x2  x3  x4  x5  x6  x7  x8  x9  x10 x11 x12 x13 x14 x15 x16
  0   1   0   0   0   0   0   0   0   0   0   0   0   0   0   1
x17 x18 x19 x20 x21 x22 x23 x24 x25 x26 x27 x28 x29 x30 x31 x32
  1   1   0   1   0   0   0   1   1   1   1   1   1   1   1   1
x33 x34 x35 x36 x37 x38 x39 x40 x41 x42 x43 x44x45 x46 x47 x48
  1   0   1   1   0   0   1   0   0   1   0   1   1   0   1   0
x49 x50 x51 x52 x53 x54 x55 x56
  0   0   1   0   0   0   0   1

> esubset(b)
=====
model SETAR( 2 , 5 , 12 ) dengan nilai threshold adalah -
0.007899935 delay 4
=====
hasil estimasi parameter regime atas
=====
      coeficient      SE coef      t-hitung      p-value
lag- 0      0.0006458      0.0008418      0.7672      0.4432
lag- 1      0.1168      0.03639      3.21      0.001376
lag- 3      0.05977      0.03666      1.631      0.1034
lag- 7      0.1045      0.03343      3.124      0.001844
lag- 12     -0.1058      0.03527      -3      0.002783

banyak data pada regime atas adalah 849
=====
hasil estimasi parameter regime bawah
=====
      coeficient      SE coef      t-hitung      p-value
lag- 0     -0.004219      0.001992     -2.118      0.03479
lag- 1     -0.1178      0.04293     -2.745      0.006329
lag- 3     -0.1048      0.0425      -2.465      0.01411
lag- 4     -0.1611      0.06724     -2.396      0.01701
banyak data pada regime bawah adalah 412

```

## Lampiran 16 Contoh Identifikasi Model Dengan Metode GA

Pada tahapan ini akan dilakukan identifikasi model SETAR dengan menggunakan metode GA. Kromosom yang digunakan mewakili nilai dari  $p_1, p_2, d, \text{threshold}, \text{subset}$  pada regime 1, dan subset pada regime 2 pada model. Kromosom yang digunakan adalah kromosom dengan nilai biner. Berikut adalah kromosom yang digunakan serta pembagian untuk masing-masing parameter yang diwakili.

d	P1		P2		Threshold							Subset R1				Subset R2				

Banyaknya gen yang dibutuhkan adalah 22.

Dekode kromosom:

```

d=b[1]+b[2]+1      p1=b[3]+b[4]+1      p2=b[5]+b[6]+1
r=b[7]*2^(7*b[7])+b[8]*2^(6*b[8])+b[9]*2^(5*b[9])+b[10]*2^(4*b[10])
+b[11]*2^(3*b[11])+b[12]*2^(2*b[12])+b[13]*2^(1*b[13])+2^b[14]
sub1=c(b[15],b[16],b[17],b[18])
sub2=c(b[19],b[20],b[21],b[22])
bsub1=sub1[1:(p1+1)]
bsub2=sub2[1:(p2+1)]
data=sort(x)
n=length(x)
bb=round(n*0.15)
bagi=256/(n*0.8)
ambil=bb+round(r/bagi)
thr=data[ambil]

```

Penentuan parameter

Popsi	=6
Pc(peluang crossover)	=0.8
Pm	=0.1
Banyak Generasi	=2
Elitism	=1
Fitness	=AIC

Generasi 1

### 1. Penentuan populasi

Membentuk populasi awal yang terdiri dari 6 kromosom. Data yang digunakan pada contoh ini adalah data model 2 subset SETAR untuk  $n=200$ . Pada software R, fungsi yang

digunakan adalah memaksimalkan fitness. Sehingga nilai fitness yang digunakan adalah  $1/AIC$ .

kromosom ke	Struktur gen																					
v1	1	0	1	1	0	0	1	1	0	1	0	1	1	1	1	0	0	0	1	1	1	0
v2	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	1	1	0	0
v3	0	1	1	1	0	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0
v4	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	0
v5	0	0	0	1	1	1	1	1	0	1	1	1	1	1	0	0	0	1	1	1	1	0
v6	0	0	1	0	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0	1	0	1

kromosom ke	decode	AIC	Fitness $\frac{1}{AIC}$
v1	d=2;p1=3;p2=1;threshold=1.693078;	110.0426	0.009087
v2	d=3;p1=2;p2=2; threshold =-0.4371832	90.7359	0.011021
v3	d=2;p1=3;p2=2;threshold =2.276606	120.9874	0.008265
v4	d=3;p1=3;p2=3;threshold =2.962414	71.02282	0.01408
v5	d=1;p1=2;p2=3;threshold=1.856403	106.5126	0.009389
v6	d=1; p1=2; p2=3;threshold=2.106127	93.68528	0.010674

## 2. Seleksi kromosom dengan roda roulette

kromosom ke	Fitness relative (pk)	Fitness komulatif(qk)
v1	0.145355	0.14536
v2	0.176291	0.32165
v3	0.132206	0.453861
v4	0.225222	0.679082
v5	0.150186	0.82926
v6	0.17074	1

Pembangkitan bilangan random untuk seleksi

r	Bilangan random
1	0.1732
2	0.3983
3	0.9043
4	0.5989
5	0.3867
6	0.1850

Bilangan random pertama  $r_1 = 0.1732$  bisa dilihat bahwa  $r_1 < q_2$ . Dengan demikian, maka  $v_2$  akan terpilih menjadi kromosom baru yang pertama. Bilangan random kedua  $r_2 = 0.3983$  terlihat bahwa  $r_2 < q_3$ . Dengan demikian, maka  $v_3$  akan terpilih menjadi kromosom baru yang kedua.  $r_3 < q_6 \rightarrow v_6$ ,  $r_4 < q_4 \rightarrow v_4$ ,  $r_5 < q_3 \rightarrow v_3$  dan  $r_6 < q_2 \rightarrow v_2$ . Diperoleh kromosom hasil seleksi sebagai berikut

Kromosom ke	Struktur gen																					
v1'	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	1	1	0	0
v2'	0	1	1	1	0	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0
v3'	0	0	1	0	1	1	1	1	1	1	0	0	0	0	0	0	1	0	0	1	0	1
v4'	1	1	1	1	1	1	1	1	1	1	1	1	0	1	1	1	1	0	1	1	1	0
v5'	0	1	1	1	0	1	1	1	1	1	0	0	1	1	0	0	0	1	1	1	1	0
v6'	1	1	0	1	1	0	0	0	1	1	1	0	0	0	1	1	0	0	1	1	0	0

Kromosom ke	decode	AIC	Fitness $\frac{1}{AIC}$	Asal
v1'	d=3;p1=2;p2=2;threshold =-0.4371832	90.7359	0.011021	v2
v2'	d=2;p1=3;p2=2;threshold =2.276606	120.9874	0.008265	v3
v3'	d=1; p1=2; p2=3;threshold=2.106127	93.68528	0.010674	v6
v4'	d=3;p1=3;p2=3;threshold =2.962414	71.02282	0.01408	v4
v5'	d=2;p1=3;p2=2;threshold =2.276606	120.9874	0.008265	v3
v6'	d=3;p1=2;p2=2;threshold =-0.4371832	90.7359	0.011021	v2

### 3. Crossover

Karena peluang crossover ( $P_c$ ) adalah 0.8 maka diharapkan 80% dari total kromosom akan mengalami crossover (5 dari 6 kromosom). Untuk memilih kromosom-kromosom mana saja yang akan dilakukan crossover, bangkitkan bilangan acak [0,1] sebanyak 6 buah.

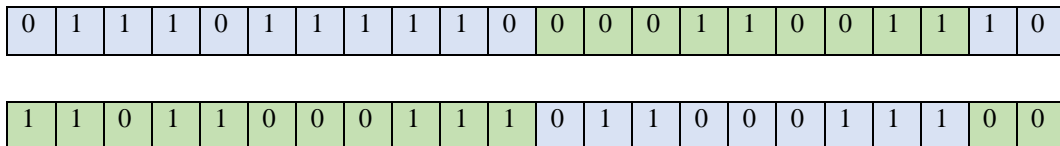
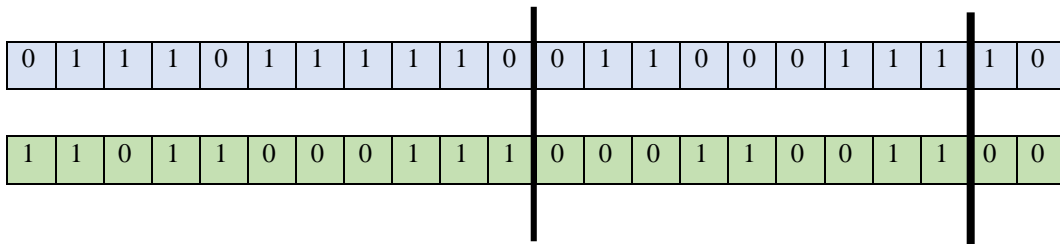
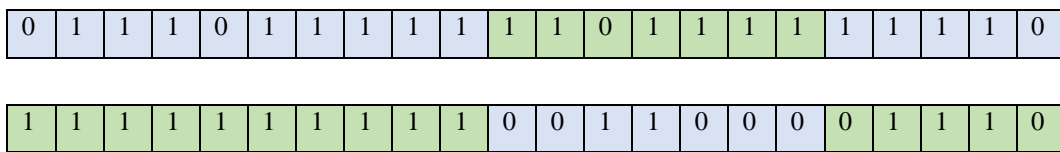
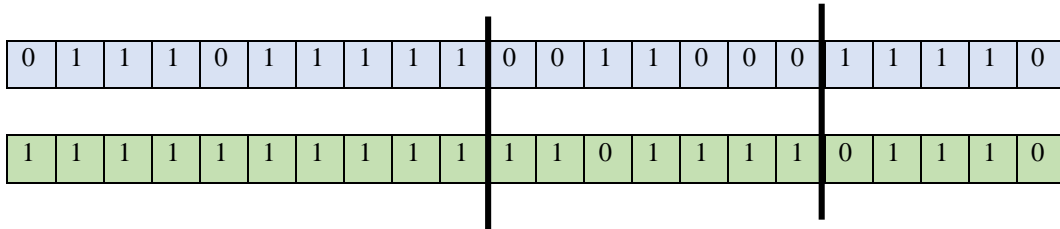
No	Bilangan random
1	0.957
2	0.169
3	0.911
4	0.175
5	0.247
6	0.168

Pilih bilangan random yang kurang dari  $P_c$  (0.80). Bilangan random  $r_1 = 0.957$  bilangan random ini lebih besar jika dibandingkan dengan  $P_c$  (0.80). Dengan demikian kromosom v1' tidak akan mengalami crossover. Bilangan random yang kurang dari  $P_c$  adalah bilangan ke-2, 4, 5, dan 6. Berarti kromosom yang berhak mengalami crossover adalah v2', v4', v5', dan v6'.

Silangkan antara v2' dan v4' serta v5' dan v6'. Pada crossover v2' dan v4' serta v5' dan v6' pilih 2 bilangan acak antara 1 hingga 21 yang akan menentukan posisi crossover 2 titik.



Crossover	Bilangan random
V2' dan v4'	10 dan 17
V5' dan v6'	11 dan 20



Diperoleh kromosom hasil crossover sebagai berikut

Kromosom ke	Struktur gen
v1'	1 1 0 1 1 0 0 0 1 1 1 0 0 0 1 1 0 0 1 1 0 0
v2''	0 1 1 1 0 1 1 1 1 1 1 1 0 1 1 1 1 1 1 1 1 0
v3'	0 0 1 0 1 1 1 1 1 1 1 0 0 0 0 0 0 1 0 0 1 0 1
v4''	1 1 1 1 1 1 1 1 1 1 0 0 1 1 0 0 0 0 1 1 1 0
v5''	0 1 1 1 0 1 1 1 1 1 0 0 0 0 1 1 0 0 1 1 1 0
v6''	1 1 0 1 1 0 0 0 1 1 1 0 1 1 0 0 0 1 1 1 0 0

Kromosom ke	decode	AIC	Fitness $\frac{1}{AIC}$	Asal
v1'	d=3;p1=2;p2=2;threshold =-0.4371832	90.7359	0.011021	v2
v2''	d=2;p1=3;p2=2;threshold =2.962414	57.85715	0.017284	c(v2,v3)
v3'	d=1; p1=2; p2=3;threshold=2.106127	93.68528	0.010674	v6
v4''	d=3;p1=3;p2=3;threshold =2.276606	158.0233	0.006328	c(v2,v4)
v5''	d=2;p1=3;p2=2;threshold =2.106127	104.315	0.009586	c(v5,v6)
v6''	d=3;p1=2;p2=2;threshold =-0.3927361	151.0623	0.00662	c(v5,v6)

#### 4. Mutasi

Karena peluang Mutasi (Pm) adalah 0.1 maka diharapkan 10% dari total gen yang akan mengalami mutasi. Untuk memilih gen mana saja yang akan dilakukan mutasi, bangkitkan bilangan acak [0,1] sebanyak  $6 \times 22 = 132$  buah .

##### kromosom ke-1

gen 1-11	0.477	0.010	0.336	0.213	0.204	0.976	0.875	0.476	0.499	0.756	0.828
gen 12-22	0.056	0.474	0.586	0.362	0.308	0.162	0.498	0.674	0.545	0.186	0.150

##### kromosom ke-2

gen 1-11	0.965	0.772	0.467	0.517	0.551	0.398	0.119	0.764	0.072	0.628	0.406
gen 12-22	0.765	0.729	0.227	0.472	0.949	0.425	0.190	0.024	0.425	0.397	0.921

##### kromosom ke-3

gen 1-11	0.819	0.150	0.865	0.861	0.760	0.176	0.588	0.752	0.028	0.158	0.746
gen 12-22	0.555	0.637	0.666	0.393	0.306	0.829	0.385	0.450	0.862	0.127	0.141

##### kromosom ke-4

gen 1-11	0.329	0.463	0.261	0.657	0.150	0.805	0.840	0.602	0.270	0.898	0.949
gen 12-22	0.615	0.740	0.535	0.241	0.074	0.793	0.325	0.284	0.114	0.975	0.121

##### kromosom ke-5

gen 1-11	0.143	0.521	0.173	0.545	0.611	0.943	0.241	0.625	0.219	0.390	0.997
gen 12-22	0.365	0.307	0.797	0.845	0.918	0.658	0.787	0.001	0.985	0.533	0.642

##### kromosom ke-6

gen 1-11	0.054	0.224	0.934	0.533	0.502	0.824	0.766	0.433	0.887	0.268	0.179
gen 12-22	0.887	0.035	0.782	0.189	0.152	0.169	0.823	0.521	0.276	0.574	0.128

Pilih bilangan random yang kurang dari Pm (0.10) dan diperoleh hasil berikut

Kromosom ke-	Bit ke-
1	2 dan 12
2	9 dan 19
3	9
4	16
5	19
6	1 dan 13

Diperoleh kromosom hasil mutasi sebagai berikut

Kromosom ke	Struktur gen
v1	1 0 0 1 1 0 0 0 1 1 1 1 0 0 1 1 0 0 1 1 0 0
v2	0 1 1 1 0 1 1 1 0 1 1 1 0 1 1 1 1 1 0 1 1 0
v3	0 0 1 0 1 1 1 1 0 1 0 0 0 0 0 0 1 0 0 1 0 1
v4	1 1 1 1 1 1 1 1 1 1 0 0 1 1 0 1 0 0 1 1 1 0
v5	0 1 1 1 0 1 1 1 1 1 0 0 0 0 1 1 0 0 0 1 1 0
v6	0 1 0 1 1 0 0 0 1 1 1 0 0 1 0 0 0 1 1 1 1 0 0

Kromosom ke	decode	AIC	Fitness $\frac{1}{AIC}$
v1	d=2;p1=2;p2=2;threshold = -0.3863525	91.50426	0.010928
v2	d=2;p1=3;p2=2;threshold = 1.788222	49.68386	0.020127
v3	d=1; p1=2; p2=3;threshold= 1.556758	72.73601	0.013748
v4	d=3;p1=3;p2=3;threshold = 2.276606	104.0341	0.009612
v5	d=2;p1=3;p2=2;threshold =2.106127	102.7946	0.009728
v6	d=2;p1=2;p2=2;threshold = -0.421802	127.1443	0.007865

Menyimpan kromosom ke-2 yang memiliki nilai fitness paling tinggi sebagai solusi dari generasi pertama. Selanjutnya, populasi akhir dari proses mutasi ini akan dijadikan sebagai populasi awal untuk generasi ke-2.

## Lampiran 17 Contoh Estimasi Parameter Model Dengan Metode GA

Pada tahapan ini akan dilakukan estimasi model SETAR dengan menggunakan metode GA. Kromosom yang digunakan mewakili nilai dari parameter pada regime 1, parameter pada regime 2, dan threshold pada model. Kromosom yang digunakan adalah kromosom dengan nilai real value. Berikut adalah kromosom yang digunakan serta pembagian untuk masing-masing parameter yang diwakili. Data yang digunakan pada tahap ini adalah data SETAR(2,1,1) untuk model 1 dengan n=200.

Threshold	$\varphi_{0,2}$	$\varphi_{1,2}$	$\varphi_{0,1}$	$\varphi_{1,1}$
	Parameter R2		Parameter R1	

### Penentuan parameter

Popsi	=6
Pc(peluang crossover)	=0.8
Pm	=0.1
Banyak Generasi	=2
Elitism	=1
Fitness	=SSE

### Generasi 1

#### 1. Penentuan populasi

Membentuk populasi awal yang terdiri dari 6 kromosom. Data yang digunakan pada contoh ini adalah data model 1 SETAR(2,1,1) untuk n=200. Pada software matlab, fungsi yang digunakan adalah meminimumkan fitness. Sehingga nilai fitness yang digunakan adalah SSE.

Kromosom ke	Struktur gen					Fitness
v1	8.2317	6.4137	0.3579	5.7693	0.1902	812.3074
v2	8.3742	6.4552	0.3542	5.1345	0.3128	834.8957
v3	8.0098	6.9209	0.3146	7.3419	-0.1204	725.6512
v4	7.6764	4.6941	0.5077	6.7194	0.0079	826.69
v5	7.7297	5.3298	0.4523	7.0707	-0.0658	796.7228
v6	7.6121	4.3686	0.5362	6.7403	0.0034	834.6865

2. Seleksi kromosom dengan roda roulette

kromosom ke	Fitness relative (pk)	Fitness komulatif(qk)
v1	0.164807	0.164807
v2	0.160348	0.325154
v3	0.184488	0.509642
v4	0.161939	0.671582
v5	0.16803	0.839612
v6	0.160388	1

Pembangkitan bilangan random untuk seleksi

r	Bilangan random
1	0.7388
2	0.3203
3	0.4975
4	0.9755
5	0.2386
6	0.5998

Bilangan random pertama  $r_1 = 0.7388$  bisa dilihat bahwa  $r_1 < q_5$ . Dengan demikian, maka  $v_5$  akan terpilih menjadi kromosom baru yang pertama. Bilangan random kedua  $r_2 = 0.3203$  terlihat bahwa  $r_2 < q_2$ . Dengan demikian, maka  $v_2$  akan terpilih menjadi kromosom baru yang kedua.  $r_3 < q_3 \rightarrow v_3$ ,  $r_4 < q_6 \rightarrow v_6$ ,  $r_5 < q_2 \rightarrow v_2$  dan  $r_6 < q_4 \rightarrow v_4$ . Diperoleh kromosom hasil seleksi sebagai berikut

Kromosom ke	Struktur gen					Fitness	Asal
v1'	7.7297	5.3298	0.4523	7.0707	-0.0658	796.7228	V5
v2'	8.3742	6.4552	0.3542	5.1345	0.3128	834.8957	V2
v3'	8.0098	6.9209	0.3146	7.3419	-0.1204	725.6512	V3
v4'	7.6121	4.3686	0.5362	6.7403	0.0034	834.6865	V6
v5'	8.3742	6.4552	0.3542	5.1345	0.3128	834.8957	V2
v6'	7.6764	4.6941	0.5077	6.7194	0.0079	826.69	V4

3. Crossover

Karena peluang crossover (Pc) adalah 0.8 maka diharapkan 80% dari total kromosom akan mengalami crossover (5 dari 6 kromosom). Untuk memilih kromosom-kromosom mana saja yang akan dilakukan crossover, bangkitkan bilangan acak [0,1] sebanyak 6 buah.

No	Bilangan random
1	0.7403
2	0.9120
3	0.5788
4	0.0260
5	0.3111
6	0.8415

Pilih bilangan random yang kurang dari  $P_c$  (0.80). Bilangan random  $r_1 = 0.7403$  bilangan random ini lebih kecil jika dibandingkan dengan  $P_c$  (0.80). Dengan demikian kromosom  $v_1$  akan mengalami crossover. Bilangan random yang kurang dari  $P_c$  adalah bilangan ke-3, 4, dan 5. Berarti kromosom yang berhak mengalami crossover adalah  $v_1'$ ,  $v_3'$ ,  $v_4'$ , dan  $v_5'$ .

Silangkan antara  $v_1'$  dan  $v_3'$  serta  $v_4'$  dan  $v_5'$ . Pada crossover  $v_1'$  dan  $v_3'$  serta  $v_4'$  dan  $v_5'$  pilih 2 bilangan acak antara 1 hingga 4 yang akan menentukan posisi crossover 2 titik.

Crossover	Bilangan random
$v_1'$ dan $v_3'$	1 dan 3
$v_4'$ dan $v_5'$	1 dan 4

7.7297	5.3298	0.4523	7.0707	-0.0658
8.0098	6.9209	0.3146	7.3419	-0.1204
7.7297	6.9209	0.3146	7.0707	-0.0658
8.0098	5.3298	0.4523	7.3419	-0.1204

7.6121	4.3686	0.5362	6.7403	0.0034
8.3742	6.4552	0.3542	5.1345	0.3128
7.6121	6.4552	0.3542	5.1345	0.0034
8.3742	4.3686	0.5362	6.7403	0.3128

Diperoleh kromosom hasil crossover sebagai berikut

Kromosom ke	Struktur gen					Fitness
$v_1''$	7.7297	6.9209	0.3146	7.0707	-0.0658	809.0416
$v_2'$	8.3742	6.4552	0.3542	5.1345	0.3128	834.8957
$v_3''$	8.0098	5.3298	0.4523	7.3419	-0.1204	736.1470
$v_4''$	7.6121	6.4552	0.3542	5.1345	0.0034	1004.6
$v_5''$	8.3742	4.3686	0.5362	6.7403	0.3128	1063.7
$v_6'$	7.6764	4.6941	0.5077	6.7194	0.0079	826.69

#### 4. Mutasi

Karena peluang Mutasi ( $P_m$ ) adalah 0.1 maka diharapkan 10% dari total gen yang akan mengalami mutasi. Untuk memilih gen mana saja yang akan dilakukan mutasi, bangkitkan bilangan acak  $[0,1]$  sebanyak  $6 \times 5 = 30$  buah .

Kromosom 1	0.357107	0.127173	0.950869	0.727676	0.967161
Kromosom 2	0.493148	0.761432	0.352283	0.047044	0.939825
Kromosom 3	0.751716	0.817442	0.608608	0.058419	0.110178
Kromosom 4	0.863392	0.499828	0.689048	0.913166	0.947924
Kromosom 5	0.871377	0.978339	0.697661	0.204111	0.522274
Kromosom 6	0.117947	0.790657	0.317802	0.850621	0.06662

Pilih bilangan random yang kurang dari  $P_m$  (0.10) dan diperoleh hasil berikut

Kromosom ke-	Bit ke-
2	4
3	4
6	5

Diperoleh kromosom hasil mutasi sebagai berikut

Kromosom ke	Struktur gen					Fitness
v1	7.7297	6.9209	0.3146	7.0707	-0.0658	809.0416
v2	8.3742	6.4552	0.3542	4.54905	0.3128	863.6899
v3	8.0098	5.3298	0.4523	6.21563	-0.1204	823.6963
v4	7.6121	6.4552	0.3542	5.1345	0.0034	1004.6
v5	8.3742	4.3686	0.5362	6.7403	0.3128	1063.7
v6	7.6764	4.6941	0.5077	6.7194	-0.1892	913.0554

Menyimpan kromosom ke-1 yang memiliki nilai fitness paling tinggi sebagai solusi dari generasi pertama. Selanjutnya, populasi akhir dari proses mutasi ini akan dijadikan sebagai populasi awal untuk generasi ke-2.





## BIOGRAFI PENULIS



Maulida Nurhidayati dilahirkan di Kediri pada 22 Oktober 1989 yang merupakan anak kedua dari tiga bersaudara pasangan Bapak Winarto dan Ibu Komsatun. Pendidikan dasar diselesaikan pada tahun 2001 di SDN Banaran I, Kemudian melanjutkan ke SMP Negeri I Kandangan dan lulus pada tahun 2004. Pada tahun yang sama melanjutkan ke SMA Negeri 1 Kandangan, Kediri dan selesai pada tahun 2007. Selanjutnya melanjutkan studi di Universitas Negeri Malang Fakultas MIPA Jurusan Matematika Prodi S1 Matematika melalui jalur SPMB dan menyelesaikan studi pada tahun 2012. Pada Tahun 2013 berkesempatan mendapatkan Beasiswa Program Pascasarjana Dalam Negeri (BPPDN) dan melanjutkan pendidikan di Institut Teknologi Sepuluh Nopember Surabaya Fakultas MIPA Jurusan Statistika Prodi S2 Statistika. Selama kuliah pernah mengikuti seminar internasional yang diselenggarakan oleh Universitas Brawijaya Malang.

Saran dan masukan terkait tesis yang saya susun dapat dikirimkan di [maulida\\_nurhidayati@yahoo.co.id](mailto:maulida_nurhidayati@yahoo.co.id) atau 08563245777. Terima kasih.