



TESIS - SS142501

**MODEL *GENERALIZED AUTOREGRESSIVE  
CONDITIONAL HETEROSCEDASTICITY*  
DENGAN OPTIMASI ALGORITMA *GENETIKA*  
UNTUK PERAMALAN KECEPATAN ANGIN  
RATA-RATA HARIAN DI JUANDA SURABAYA**

**INDAH FAHMIYAH  
NRP. 06211650010032**

**DOSEN PEMBIMBING  
Irhamah, M.Si., Ph.D.  
Dr.rer.pol. Heri Kuswanto, M.Si.**

**PROGRAM MAGISTER  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2018**



TESIS - SS142501

**MODEL *GENERALIZED AUTOREGRESSIVE  
CONDITIONAL HETEROSCEDASTICITY*  
DENGAN OPTIMASI ALGORITMA GENETIKA  
UNTUK PERAMALAN KECEPATAN ANGIN  
RATA-RATA HARIAN DI JUANDA SURABAYA**

**INDAH FAHMIYAH  
NRP. 06211650010032**

**DOSEN PEMBIMBING  
Irhamah, M.Si., Ph.D.  
Dr.rer.pol. Heri Kuswanto, M.Si.**

**PROGRAM MAGISTER  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2018**



THESIS - SS142501

**GENERALIZED AUTOREGRESSIVE CONDITIONAL  
HETEROSCEDASTICITY MODEL WITH GENETIC  
ALGORITHM OPTIMIZATION FOR DAILY AVERAGE  
WIND SPEED FORECASTING IN JUANDA SURABAYA**

**INDAH FAHMIYAH  
NRP. 06211650010032**

**SUPERVISORS  
Irhamah, M.Si., Ph.D.  
Dr.rer.pol. Heri Kuswanto, M.Si.**

**MASTER PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA  
2018**

**MODEL GENERALIZED AUTOREGRESSIVE CONDITIONAL  
HETEROSCEDASTICITY DENGAN OPTIMASI ALGORITMA  
GENETIKA UNTUK PERAMALAN KECEPATAN ANGIN  
RATA-RATA HARIAN DI JUANDA SURABAYA**

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

oleh :  
**INDAH FAHMIYAH  
NRP. 06211650010032**

**Tanggal Ujian : 12 Juli 2018  
Periode Wisuda : September 2018**

**Disetujui oleh:**

1.   
**Irhamah, M.Si., Ph.D.**  
**NIP: 19780406 200112 2 002**

**(Pembimbing I)**

2.   
**Dr. rer. pol. Heri Kuswanto, M.Si.**  
**NIP: 19820326 200312 1 004**

**(Pembimbing II)**

3.   
**Dr. Kartika Fithriasari, M.Si.**  
**NIP: 19691212 199303 2 002**

**(Penguji)**

4.   
**Dr. Puhadi, M.Sc.**  
**NIP: 19620204 198701 1 001**

**(Penguji)**

**Dekan**

**Fakultas Matematika, Komputasi, dan Sains Data  
Institut Teknologi Sepuluh Nopember**

  
**Prof. Dr. Basuki Widodo, M.Sc.**  
**NIP: 19650605 198903 1 002**

**FAKULTAS  
MATEMATIKA, KOMPUTASI  
DAN SAINS DATA**

# **MODEL *GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY* DENGAN OPTIMASI ALGORITMA GENETIKA UNTUK PERAMALAN KECEPATAN ANGIN RATA-RATA HARIAN DI JUANDA SURABAYA**

**Nama** : Indah Fahmiyah  
**NRP** : 06211650010032  
**Departemen** : Statistika  
**Pembimbing** : Irhamah, M.Si., Ph.D.  
**Co-Pembimbing** : Dr.rer.pol. Heri Kuswanto, M.Si.

## **ABSTRAK**

Peramalan kecepatan angin yang akurat menjadi penting untuk menghindari bencana atau kecelakaan, seperti kecelakaan transportasi udara, bangunan runtuh, pohon tumbang, dan sebagainya. Hal tersebut disebabkan karena angin yang tidak terkendali dapat membahayakan manusia dan merusak lingkungan. Metode peramalan yang paling umum digunakan adalah metode ARIMA. Kecepatan angin pada umumnya memiliki fluktuasi dan volatilitas sehingga model ARIMA tidak dapat menangkap fenomena heteroskedastisitas atau volatilitas. Oleh karena itu, kecepatan angin diprediksi dengan metode *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). Hasil pemodelan kecepatan angin menggunakan ARIMA-GARCH lebih baik daripada model ARIMA berdasarkan nilai kriteria evaluasi model. Keakuratan hasil peramalan juga dipengaruhi oleh metode optimasi parameter selain dipengaruhi oleh metode peramalan yang digunakan. Algoritma Genetika atau *Genetic Algorithm* (GA) adalah salah satu mesin pembelajaran (*machine learning*) yang dapat digunakan untuk mengoptimalkan estimasi parameter model. GA dapat menghasilkan solusi global dan optimal. Parameter model ARIMA-GARCH yang dioptimasi dengan menggunakan GA memiliki nilai evaluasi model yang lebih baik daripada menggunakan model ARIMA-GARCH. Model ini diterapkan untuk kecepatan angin rata-rata harian di stasiun BMKG Juanda, Surabaya. Simulasi dari skenario yang dibuat juga menunjukkan bahwa GARCH-GA dapat menghasilkan RMSE yang lebih kecil.

**Kata Kunci:** Algoritma Genetika, GARCH, Kecepatan Angin Rata-rata

*(Halaman ini sengaja dikosongkan)*

# **GENERALIZED AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY MODEL WITH GENETIC ALGORITHM OPTIMIZATION FOR DAILY AVERAGE WIND SPEED FORECASTING IN JUANDA SURABAYA**

**Name** : Indah Fahmiyah  
**Student Identity Number** : 06211650010032  
**Department** : Statistics  
**Supervisor** : Irhamah, M.Si., Ph.D.  
**Co-Supervisor** : Dr.rer.pol. Heri Kuswanto, M.Si.

## **ABSTRACT**

Accurate wind speed forecasting is becoming important to avoid any disaster or accident, such as air transport accidents, collapsed buildings, fallen trees, and so on. It is because uncontrolled wind speed can harm people and damage the environment. The most commonly used forecasting method is ARIMA method. Wind speed generally has fluctuations and volatility, so the ARIMA model cannot capture the phenomenon of heteroscedasticity or volatility. Therefore, wind speed is predicted by Generalized Autoregressive Conditional Heteroscedasticity (GARCH) method. The result of wind speed modeling using ARIMA-GARCH is better than with ARIMA model based on model evaluation criteria value. The accuracy of the forecasting results apart from the correct forecasting method is also influenced by the parameter optimization method. Genetic algorithm (GA) is one of the machine learning that can be used to optimize model parameters. That algorithm can produce global and optimal solutions. The ARIMA-GARCH model parameters optimized using GA have a better model evaluation value than using the ARIMA-GARCH model. The model is applied to daily average wind speed at a meteorological station on Juanda, Surabaya. The results of simulation also show that GARCH-GA models can perform RMSE lower than GARCH models.

**Keywords:** Average Wind Speed, GARCH, Genetic Algorithm

*(This page is intentionally left blank)*



## KATA PENGANTAR

Puji syukur kehadiran Allah SWT yang telah melimpahkan rahmat dan hidayah-Nya sehingga atas izin-Nya penulis dapat menyelesaikan Tesis yang berjudul “**Model Generalized Autoregressive Conditional Heteroscedasticity dengan Optimasi Algoritma Genetika untuk Peramalan Kecepatan Angin Rata-rata Harian di Juanda Surabaya**” dengan baik dan tepat waktu. Dengan demikian, penulis dapat menyelesaikan Tesis yang merupakan salah satu syarat untuk menyelesaikan Program Magister sesuai dengan kurikulum di Departemen Statistika ITS.

Penyusunan buku Tesis tidak terlepas dari bantuan, dukungan, dan bimbingan berbagai pihak. Oleh karena itu, penulis mengucapkan banyak terima kasih kepada.

1. Irhamah, M.Si., Ph.D. dan Dr.rer.pol. Heri Kuswanto, M.Si. selaku dosen pembimbing penulis yang senantiasa membimbing dengan sabar serta memberikan motivasi dalam menyelesaikan Tesis.
2. Dr. Kartika Fithriasari, M.Si. dan Dr. Purnadi, M.Sc. selaku dosen penguji yang telah memberikan saran dan kritik demi kebaikan Tesis penulis.
3. Dr. Suhartono selaku Ketua Departemen Statistika ITS.
4. Seluruh pimpinan dan tenaga kependidikan Program Studi S2 Departemen Statistika ITS.
5. Bapak, Ibu, dan keluarga penulis yang selalu mendo'akan dan memberikan dukungan agar penulis dapat menyelesaikan Program Studi S2 dengan baik.
6. Seluruh dosen Departemen Statistika ITS.
7. Nur Silviah Rahmi, Febrian Kristianda, Nita Cahyani, Ratih, Syarifah, Ratna, dan T Dwi selaku teman sebimbingan Tesis.
8. Farida Nur Hayati, Niswatul Qona'ah, Ernawati, dan Silviatul Hasanah, Bahagiati, dan Meranggi selaku teman seperjuangan *fresh graduate*.
9. Teman-teman S2 Statistika ITS angkatan 2016 yang telah membantu dalam penyelesaian Tesis.
10. Teman-teman forum diskusi dan semua pihak yang telah membantu penulis dalam penyelesaian Tesis.

Penulis berharap Tesis ini dapat memberikan manfaat bagi masyarakat dan mengembangkan ilmu pengetahuan. Penulis menyadari bahwa Tesis ini masih jauh dari kesempurnaan. Oleh karena itu, penulis meminta maaf serta mengharapkan saran dan kritik yang bersifat membangun guna perbaikan penelitian-penelitian selanjutnya.

Surabaya, Juli 2018

Penulis

## DAFTAR ISI

<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>xvii</b>
<b>DAFTAR GAMBAR.....</b>	<b>xix</b>
<b>BAB 1 PENDAHULUAN .....</b>	<b>1</b>
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah.....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	4
1.5 Batasan Masalah .....	4
<b>BAB 2 TINJAUAN PUSTAKA.....</b>	<b>5</b>
2.1 Model <i>Time Series</i> .....	5
2.1.1 Uji Stasioner dalam Rata-Rata.....	5
2.1.2 Transformasi .....	7
2.1.3 <i>Autocorrelation Function</i> dan <i>Partial Autocorrelation Function</i> ..	7
2.2 <i>Autoregressive Integrated Moving Average (ARIMA)</i> .....	9
2.2.1 Identifikasi Model ARIMA .....	11
2.2.2 Estimasi Parameter .....	12
2.2.3 Pengujian Signifikansi Parameter .....	13
2.2.4 Pemeriksaan Diagnostik .....	13
2.2.4.1 Uji Asumsi <i>White Noise</i> .....	14
2.2.4.2 Uji Asumsi Distribusi Normal .....	14
2.2.5 Seleksi Model Terbaik .....	15
2.2.5.1 Berdasarkan Data <i>In Sample</i> .....	15
2.2.5.2 Berdasarkan Data <i>Out Sample</i> .....	15
2.2.6 Peramalan Model ARIMA.....	16

2.3	<i>Generalized Autoregressive Conditional Heteroscedasticity</i> .....	16
2.3.1	Peramalan Variansi Residual .....	17
2.4	Algoritma Genetika .....	18
2.4.1	Kromosom.....	19
2.4.2	Fungsi <i>Fitness</i> .....	20
2.4.3	Seleksi dengan Roda <i>Roulette</i> .....	20
2.4.4	<i>Crossover</i> atau Pindah Silang .....	20
2.4.5	Mutasi.....	21
2.4.6	Elitisme.....	21
2.5	Kecepatan Angin .....	21
<b>BAB 3 METODOLOGI PENELITIAN .....</b>		<b>23</b>
3.1	Sumber Data dan Variabel Penelitian .....	23
3.2	Langkah Analisis.....	23
3.2.1	Data Kecepatan Angin Rata-rata.....	23
3.2.2	Data Simulasi .....	25
3.3	Diagram Alir .....	26
<b>BAB 4 ANALISIS DAN PEMBAHASAN.....</b>		<b>29</b>
4.1	Karakteristik Kecepatan Angin .....	29
4.2	Model ARIMA .....	30
4.2.1	Identifikasi Model ARIMA.....	30
4.2.2	Estimasi Parameter dan Uji Signifikansi.....	32
4.2.3	Cek Diagnostik.....	33
4.2.4	Evaluasi Model.....	35
4.2.5	Peramalan Kecepatan Angin dengan Model ARIMA.....	37
4.3	Model GARCH .....	38
4.3.1	Uji <i>Lagrange Multiplier</i> .....	39
4.3.2	Estimasi Parameter Model GARCH .....	39
4.3.3	Evaluasi Kebaikan Model GARCH .....	41
4.3.4	Peramalan Variansi Residual .....	41
4.4	Model GARCH-GA .....	42
4.5	Studi Simulasi.....	45
4.5.1	Data Simulasi Skenario 1 .....	46

4.5.2 Data Simulasi Skenario 2.....	49
<b>BAB 5 KESIMPULAN DAN SARAN .....</b>	<b>53</b>
5.1 Kesimpulan .....	53
5.2 Saran .....	53
<b>DAFTAR PUSTAKA .....</b>	<b>55</b>
<b>LAMPIRAN.....</b>	<b>57</b>
<b>BIODATA PENULIS.....</b>	<b>69</b>

*(Halaman ini sengaja dikosongkan)*

## DAFTAR TABEL

<b>Tabel 2.1</b>	Transformasi Box-Cox.....	7
<b>Tabel 2.2</b>	Karakteristik Plot ACF-PACF pada model ARIMA.....	12
<b>Tabel 3.1</b>	Struktur Data Kecepatan Angin Rata-rata.....	23
<b>Tabel 4.1</b>	Uji ADF.....	31
<b>Tabel 4.2</b>	Estimasi Parameter Model ARIMA .....	33
<b>Tabel 4.3</b>	Uji Ljung-Box Residual Model AR([1,2,3,9]).....	34
<b>Tabel 4.4</b>	Uji Ljung-Box Residual Model ARMA ([1,2,3,9],[19]), dan ARMA([3,20],3).....	35
<b>Tabel 4.5</b>	Uji KS Residual Model ARMA([1,2,3,9],[19]) dan ARMA([3,20],3).....	35
<b>Tabel 4.6</b>	Evaluasi Model ARIMA .....	36
<b>Tabel 4.7</b>	Hasil Ramalan Model AR([1,2,3,9]).....	37
<b>Tabel 4.8</b>	Hasil Uji LM .....	39
<b>Tabel 4.9</b>	Estimasi Model GARCH(1,1).....	40
<b>Tabel 4.10</b>	Estimasi Model ARCH(1).....	40
<b>Tabel 4.11</b>	Evaluasi Model AR([1,2,3,9])-ARCH(1).....	41
<b>Tabel 4.12</b>	Ramalan variansi residual .....	42
<b>Tabel 4.13</b>	Ilustrasi Seleksi <i>Rhoulette Whell</i> .....	43
<b>Tabel 4.14</b>	Solusi Estimasi Parameter ARCH(1) dengan GA.....	45
<b>Tabel 4.15</b>	Skenario Simulasi.....	45
<b>Tabel 4.16</b>	<i>Skewness</i> dan Kurtosis Data Simulasi Skenario 1 untuk $n = 300$ ....	47
<b>Tabel 4.17</b>	Hasil Estimasi dan RMSE GARCH(1,1)-GA Skenario 1 untuk $n = 300$ .....	47
<b>Tabel 4.18</b>	<i>Skewness</i> dan Kurtosis Data Simulasi Skenario 1 untuk $n = 1000$ ..	48
<b>Tabel 4.19</b>	Hasil Estimasi dan RMSE GARCH(1,1)-GA Skenario 1 untuk $n = 1000$ .....	49
<b>Tabel 4.20</b>	<i>Skewness</i> dan Kurtosis Data Simulasi Skenario 2 untuk $n = 1000$ ..	50
<b>Tabel 4.21</b>	Hasil Estimasi dan RMSE GARCH(1,1)-GA Skenario 2 untuk $n = 1000$ .....	51

*(Halaman ini sengaja dikosongkan)*



## DAFTAR GAMBAR

<b>Gambar 2.1</b>	Ilustrasi Kromosom dengan Bilangan Biner .....	19
<b>Gambar 2.2</b>	Ilustrasi Proses <i>Crossover</i> .....	21
<b>Gambar 2.3</b>	Ilustrasi Proses Mutasi.....	21
<b>Gambar 3.1</b>	Diagram Alir GA.....	26
<b>Gambar 3.2</b>	Diagram Alir Peramalan Data Kecepatan Angin Rata-rata dengan GARCH-GA .....	27
<b>Gambar 4.1</b>	Kecepatan Angin Rata-rata Juanda Surabaya.....	29
<b>Gambar 4.2</b>	Histogram Kecepatan Angin Rata-rata.....	30
<b>Gambar 4.3</b>	<i>Boxplot</i> Kecepatan Angin Rata-rata .....	30
<b>Gambar 4.4</b>	Transformasi Box-Cox Kecepatan Angin Rata-rata.....	30
<b>Gambar 4.5</b>	ACF dan PACF Kecepatan Angin Rata-rata .....	31
<b>Gambar 4.6</b>	ACF dan PACF Data Transformasi.....	32
<b>Gambar 4.7</b>	Uji KS Residual Model AR([1,2,3,9]).....	34
<b>Gambar 4.8</b>	Grafik Ramalan Model ARIMA dan Data <i>Outsample</i> .....	38
<b>Gambar 4.9</b>	Plot Residual Model AR([1,2,3,9]) .....	38
<b>Gambar 4.10</b>	Ilustrasi Solusi Metode GA .....	43
<b>Gambar 4.11</b>	Ilustrasi Kromosom ARCH(1) yang Melakukan Pindah Silang ..	44
<b>Gambar 4.12</b>	Ilustrasi Proses Mutasi.....	44
<b>Gambar 4.13</b>	Skenario 1 untuk $n = 300$ .....	46
<b>Gambar 4.14</b>	Skenario 1 untuk $n = 1000$ .....	48
<b>Gambar 4.15</b>	Skenario 2 untuk $n = 1000$ .....	50

*(Halaman ini sengaja dikosongkan)*

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Data Kecepatan Angin Rata-rata .....	57
<b>Lampiran 2</b> Uji ADF Dengan <i>Software R</i> .....	58
<b>Lampiran 3.</b> <i>Syntax</i> Model AR([1,2,3,9]) dengan SAS .....	59
<b>Lampiran 4.</b> <i>Syntax</i> Model AR([1,2,3,9])-GARCH(1,1) dengan SAS .....	60
<b>Lampiran 5.</b> <i>Syntax</i> Model AR([1,2,3,9])-ARCH(1) dengan SAS .....	61
<b>Lampiran 6.</b> <i>Syntax</i> Estimasi ARCH(1)-GA Data Kecepatan Angin Rata-rata .....	62
<b>Lampiran 7.</b> <i>Output</i> ARCH(1)-GA Data Kecepatan Angin Rata-Rata.....	63
<b>Lampiran 8.</b> Simulasi GARCH(1,1)-GA Skenario 1 untuk $n = 300$ .....	63
<b>Lampiran 9.</b> Simulasi GARCH(1,1)-GA Skenario 1 untuk $n = 1000$ .....	64
<b>Lampiran 10.</b> Simulasi GARCH(1,1)-GA Skenario 2 untuk $n = 1000$ .....	64
<b>Lampiran 11.</b> <i>Syntax</i> GARCH(1,1)-GA.....	65
<b>Lampiran 12.</b> <i>Output</i> Estimasi GARCH(1,1) Replikasi ke-1 Skenario 1 ( $n = 300$ ) .....	66
<b>Lampiran 13.</b> <i>Output</i> Estimasi GARCH(1,1) Data <i>In Sample</i> Replikasi ke-1 Skenario 1 ( $n = 300$ ).....	67

*(Halaman ini sengaja dikosongkan)*

# BAB 1

## PENDAHULUAN

### 1.1 Latar Belakang

Angin merupakan udara yang bergerak dari tekanan tinggi ke tekanan yang lebih rendah. Energi angin banyak dimanfaatkan dalam kehidupan manusia sehingga bermanfaat bagi kehidupan manusia. Angin juga salah satu sumber energi baru dan terbarukan (EBT) yang ramah lingkungan. Sumber energi angin digunakan untuk menggerakkan turbin kincir angin yang terdapat pada pembangkit listrik tenaga angin/ bayu (PLTB) sehingga menghasilkan energi listrik. Angin juga dimanfaatkan dalam sistem penerbangan, pelayaran, penyerbukan tanaman, dan lain sebagainya.

Kecepatan angin yang tidak terkendali seringkali dapat menimbulkan bencana, seperti bangunan roboh, pohon tumbang, kebakaran hutan, kecelakaan transportasi udara, dan lain sebagainya. Beberapa bencana tersebut dapat menyebabkan kerugian bagi masyarakat bahkan korban jiwa, jika tidak ditangani dengan tanggap dan serius. Dengan demikian, angin tidak hanya memberikan dampak positif bagi kehidupan manusia, tetapi angin yang sangat kencang dapat berpotensi merusak lingkungan sekitar. Dengan demikian, diperlukan peramalan kecepatan angin rata-rata dalam satu hari sehingga energi angin tersebut dapat dimanfaatkan dengan maksimal.

Metode peramalan yang sering digunakan adalah model *time series*, yaitu memodelkan kecepatan angin berdasarkan histori kecepatan angin sebelumnya. Model *time series* yang umum digunakan adalah model *Autoregressive Integrated Moving Average* (ARIMA). Kecepatan angin per jam lima lokasi topografi berbeda di Navarre, Spanyol diramalkan oleh Torres, Garcia, Blas, dan Francisco (2005) dengan model ARMA. Pengembangan metode ARIMA dilakukan agar diperoleh hasil ramalan yang lebih akurat. Wahyuningtyas (2015) meramalkan kecepatan angin dan ketinggian gelombang menggunakan ARIMA-Kalman *filter*. Peramalan kecepatan angin 24 jam dan 48 jam ke depan dilakukan oleh Kavasseri dan Seetharaman (2009) menggunakan *fractional-ARIMA* atau *f-ARIMA* dengan mengasumsikan parameter *differencing d* bersifat kontinu.

Selain menggunakan model ARIMA, metode peramalan lainnya adalah pemodelan dengan *intelligent algorithm*. Peramalan dilakukan menggunakan algoritma pembelajaran. Peramalan kecepatan angin di ladang angin dilakukan oleh Liu, Niu, Wang, dan Fan (2014) menggunakan *Support Vector Machines* (SVM) dengan optimasi algoritma genetika dan data sebelumnya ditransformasi dengan transformasi Wavelet. Chen dan Yu (2014) memprediksi kecepatan angin jangka pendek di ladang angin menggunakan *Support Vector Regression* dengan *unscented Kalman filter* (UKF).

Kecepatan angin rata-rata juga telah dimodelkan dengan menggunakan Metode *Bayesian Model Averaging* dengan pendekatan *Expectation Maximization* (Kusumawati, 2011). Pada tahun 2012, Anityaloka memodelkan kecepatan angin di Bandara Juanda Surabaya dengan metode ANFIS, lalu setahun kemudian Permai, Iriawan, dan Irhamah (2013) memodelkan kecepatan angin dengan *Mixture of ANFIS*. Metode ANFIS merupakan gabungan dari metode *artificial neural network* dan sistem inferensi yang baik digunakan untuk meramalkan kecepatan angin yang memiliki pola non-linier.

Metode peramalan dengan ARIMA, *intelligent algorithm*, Bayesian, serta ANFIS tersebut menghasilkan hasil peramalan yang akurat jika kecepatan angin tidak memiliki turbulensi atau volatilitas (heteroskedastisitas). Namun, kecepatan angin seringkali memiliki pola variansi tidak konstan sehingga kecepatan angin bersifat tidak homogen atau terjadi heteroskedastisitas. Heteroskedastisitas dapat menimbulkan beberapa permasalahan, misalnya hasil ramalan dari model menjadi kurang akurat karena tidak mengakomodasi efek heteroskedastisitas dari fenomena kecepatan angin sesungguhnya.

Model yang mengakomodasi adanya efek heteroskedastisitas untuk meramalkan kecepatan angin adalah *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). GARCH diusulkan Bollerslev (1986) dengan memodifikasi model *Autoregressive Conditional Heteroscedasticity* (ARCH) yang dikembangkan oleh Engle (1982). Dalam model GARCH, variansi residual tidak hanya berkorelasi dengan variansi residual periode lalu tetapi juga dengan residual kuadrat periode lalu sehingga hasil peramalan cenderung lebih baik daripada dengan menggunakan ARIMA.

Data kecepatan angin yang diramalkan Gao, He, dan Chen (2009) dimodelkan dengan ARMA-ARCH karena memiliki pola pengelompokan volatilitas (*volatility clustering*) sehingga variansi berubah dari waktu ke waktu. Wang, Liu, Zhen, Ceng, dan Zhou (2009) meneliti distribusi frekuensi dari kecepatan angin dan memodelkannya dengan AR-GARCH. Dengan memodelkan volatilitas kecepatan angin, ketepatan hasil peramalan kecepatan angin akan cenderung lebih tinggi. Rata-rata dan volatilitas dari kecepatan angin di Colorado, USA telah dimodelkan oleh Liu, Erdem, dan Shi (2011) dengan menggunakan ARMA-GARCH(-M) yang memiliki tipe *variant* yang tidak linier, yaitu dengan SGARCH, QGARCH, GJRGARCH, EGARCH, dan NGARCH.

Kebaikan dalam hasil ramalan selain dari metode peramalan yang tepat juga didapatkan dari penggunaan parameter yang tepat. Oleh karena itu, seringkali digunakan metode optimasi parameter agar didapatkan hasil peramalan kecepatan angin yang lebih baik, salah satunya adalah algoritma genetika. Algoritma genetika atau *Genetic Algorithm* (GA) merupakan algoritma pembelajaran (*machine learning*) yang diperkenalkan John Holland pada tahun 1975 yang mengadopsi dari sistem genetika, seperti reproduksi, *crossover*, dan mutasi. GA menggunakan konsep seleksi alam, yaitu individu yang kuat akan tetap bertahan/ *survive*.

Pada setiap generasi, algoritma mengeksplorasi banyak area dari ruang parameter, kemudian memilih daerah yang memiliki probabilitas terbesar. GA mampu mengevaluasi banyak titik dalam ruang parameter secara simultan sehingga akan konvergen pada solusi yang optimal dan global (Hung, 2009). Dengan demikian, GA sesuai untuk fungsi non-linier yang bervariasi dalam waktu atau *time-varying* (Zhou dan Khotanzad, 2007 dalam Hung, 2009). Hung (2009) memodelkan skenario pasar modal menggunakan *fuzzy* GARCH-GA, kemudian Hung (2011) memodelkan skenario pasar modal dengan menambahkan menjadi *adaptive fuzzy* GARCH-GA untuk kasus yang berbeda.

Dalam penelitian ini, kecepatan angin di Stasiun BMKG Juanda Surabaya akan diramalkan menggunakan metode GARCH dengan optimasi GA sehingga volatilitas kecepatan angin dapat diakomodasi oleh model GARCH dan mendapatkan hasil estimasi global dan optimal dengan GA. Selain itu, metode GARCH-GA akan diterapkan untuk data simulasi dengan skenario tertentu.

## 1.2 Rumusan Masalah

Angin dapat memberikan manfaat dalam kehidupan manusia, seperti dalam pelayaran, penyerbukan tanaman, dll. Selain itu, kecepatan yang tidak terkendali dapat memberikan dampak negatif sehingga dilakukan peramalan kecepatan angin yang mempertimbangkan karakteristiknya yang tidak stabil atau terjadi heteroskedastisitas dengan menggunakan metode GARCH dengan optimasi parameter model menggunakan GA. Selain itu, dilakukan peramalan dengan GARCH-GA dengan menggunakan data simulasi agar mengetahui kebaikan dari model GARCH-GA untuk beberapa skenario.

## 1.3 Tujuan Penelitian

Tujuan melakukan penelitian ini adalah.

1. Mendeskripsikan karakteristik kecepatan angin rata-rata harian di Juanda Surabaya.
2. Meramalkan kecepatan angin rata-rata harian dengan menggunakan metode GARCH dengan optimasi GA.
3. Membandingkan hasil ramalan antara model GARCH dan GARCH-GA untuk data riil dan simulasi guna mengetahui kinerja GA dalam optimasi parameter.

## 1.4 Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini, antara lain meningkatkan optimasi hasil peramalan dengan menggunakan GA, khususnya peramalan kecepatan angin rata-rata harian serta menerapkan metode GARCH-GA dengan menggunakan data simulasi agar dapat diterapkan untuk data *time series* lain yang memiliki kasus heteroskedastisitas.

## 1.5 Batasan Masalah

Data yang digunakan adalah data kecepatan angin rata-rata harian di Stasiun pengamatan Juanda Surabaya pada tahun 2008-2009. Data simulasi yang digunakan merupakan data yang dibangkitkan dari model GARCH(1,1) dengan laju *crossover* sebesar 0,8 dan laju mutasi sebesar 0,1.



## BAB 2

### TINJAUAN PUSTAKA

#### 2.1 Model *Time Series*

*Time series* adalah deretan observasi atau pengamatan yang diukur/ diamati secara berurutan/ berkala dalam periode waktu yang sama dan biasanya dinotasikan dengan  $Z_t$ , yaitu pengamatan pada waktu  $t$ . Salah satu model *time series* adalah model *Autoregressive* seperti dalam persamaan (2.1).

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t \quad (2.1)$$

yaitu pengamatan pada saat ke- $t$  yang dimodelkan dari pengamatan-pengamatan pada waktu sebelumnya dan  $a_t$  adalah *error* dari model.

Analisis menggunakan data *time series* secara umum bertujuan untuk meramalkan pengamatan pada waktu mendatang dari pengamatan-pengamatan sebelumnya dengan membentuk model *time series*. Aspek yang berhubungan erat dalam pemodelan data *time series* adalah konsep stasioneritas. Stasioneritas dapat dibagi menjadi dua, yaitu stasioner dalam rata-rata dan stasioner dalam variansi.

##### 2.1.1 Uji Stasioner dalam Rata-Rata

Data *time series* merupakan suatu proses stokastik. Suatu proses stokastik dikatakan stasioner, jika memiliki nilai rata-rata dan variansi yang konstan seiring waktu, yaitu berfluktuasi di sekitar rata-rata dan akan kembali ke nilai rata-rata (*mean reversion*). Stasioneritas data *time series* sangat diperlukan, khususnya untuk peramalan dengan model *time series*. Data yang tidak stasioner tidak baik digunakan untuk meramalkan pengamatan pada waktu berikutnya karena hanya baik untuk menganalisis periode waktu yang sudah atau sedang berlangsung (Gujarati, 2004).

Pengujian stasioneritas data *time series* yang sangat luas dipergunakan adalah *unit root test*, dimulai dengan persamaan (2.2)

$$Z_t = \rho Z_{t-1} + u_t, \quad -1 \leq \rho \leq 1 \quad (2.2)$$

dengan  $u_t$  adalah *error* yang *white noise*. Kedua sisi dari persamaan (2.2) dikurangi dengan  $Z_{t-1}$  sehingga didapatkan

$$Z_t - Z_{t-1} = \rho Z_{t-1} - Z_{t-1} + u_t$$

$$\Delta Z_t = (\rho - 1)Z_{t-1} + u_t \quad (2.3)$$

atau persamaan (2.3) dapat dinyatakan sebagai

$$\Delta Z_t = \delta Z_{t-1} + u_t \quad (2.4)$$

dengan  $\delta = (\rho - 1)$  dan  $\Delta$  sebagai operator selisih pertama (*first-difference*). Jika  $\rho = 1$  (ekivalen dengan  $\delta = 0$ ) data tidak stasioner sehingga persamaan (2.4) menjadi  $\Delta Z_t = u_t$  atau disebut *random walk*.

Uji *unit root Dickey-Fuller* (DF) dikembangkan dengan menambahkan *lag* dari  $\Delta Z_t$  menjadi *Augmented Dickey-Fuller* seperti dalam persamaan (2.5)

$$\Delta Z_t = \beta_1 + \beta_2 t + \delta Z_{t-1} + \sum_{i=1}^m \eta_i \Delta Z_{t-i} + \varepsilon_t \quad (2.5)$$

dengan  $\beta_1$  sebagai *drift* atau intersep,  $\beta_2$  disebut koefisien tren,  $\delta$  adalah koefisien dari  $Z_{t-1}$ ,  $\eta_i$  adalah koefisien dari  $\Delta Z_{t-i}$ , dan  $\varepsilon_t$  adalah *error* yang *whitenoise*, dan  $m$  banyaknya *lag*. Nilai  $\Delta Z_{t-1} = Z_{t-1} - Z_{t-2}$ ,  $\Delta Z_{t-2} = Z_{t-2} - Z_{t-3}$ , dan seterusnya (Gujarati, 2004). Hipotesis dari uji *Augmented Dickey-Fuller* (ADF) adalah.

$$H_0: \delta = 0 \text{ (data tidak stasioner atau non-stasioner)}$$

$$H_1: \delta < 0 \text{ (data stasioner)}$$

Statistik uji yang digunakan adalah statistik uji  $\tau$  dalam persamaan (2.6). Nilai  $\hat{\delta}$  adalah estimasi dari  $\delta$  dengan  $\hat{\delta} = (\hat{\rho} - 1)$  dan  $SE(\hat{\delta})$  adalah *standard error* dari  $\hat{\delta}$ .

$$\tau = \frac{\hat{\delta}}{SE(\hat{\delta})} \quad (2.6)$$

Nilai  $\tau$  mengikuti distribusi asimtotik (tidak mengikuti distribusi  $t$ ) dan telah ditabulasikan nilainya oleh *Dickey-Fuller*. Dengan demikian, hipotesis nol ditolak jika nilai  $|\tau|$  lebih besar dari nilai  $t$  dengan derajat bebas  $n$  banyaknya pengamatan dan taraf signifikansi  $\alpha$  dalam Tabel *Dickey Fuller*. Jadi, disimpulkan bahwa data stasioner dalam *mean*.

Apabila data *time series* berupa deret  $\{Z_t\}$  tidak stasioner dalam rata-rata, data tersebut di-*differencing* menjadi deret  $\{(1-B)^d Z_t\}$  sehingga menjadi deret yang stasioner dengan  $d$  adalah orde *differencing* dan  $B$  adalah operator *backshift*,

misalnya  $B^d Z_t = Z_{t-d}$ . Dalam beberapa kasus *differencing*, orde  $d$  sebesar 0, 1, atau 2 (Wei, 2006).

### 2.1.2 Transformasi

Box dan Cox dalam Wei (2006) memperkenalkan transformasi pangkat untuk mengatasi data *time series* yang tidak stasioner dalam variansi. Persamaan (2.7) merupakan transformasi pangkat yang digunakan untuk mentransformasi data agar menjadi stasioner. Transformasi pangkat disebut juga transformasi *Box-Cox* dan disajikan dalam Tabel 2.1 untuk beberapa nilai estimasi lambda.

$$T(Z_t) = \begin{cases} \frac{Z_t^\lambda - 1}{\lambda}, & \text{jika } \lambda \neq 0 \\ \lim_{\lambda \rightarrow 0} \frac{Z_t^\lambda - 1}{\lambda} = \ln(Z_t), & \text{jika } \lambda = 0 \end{cases} \quad (2.7)$$

**Tabel 2.1** Transformasi Box-Cox

Nilai Lambda ( $\lambda$ )	Transformasi
-1,0	$\frac{1}{Z_t}$
-0,5	$\frac{1}{\sqrt{Z_t}}$
0,0	$\ln(Z_t)$
0,5	$\sqrt{Z_t}$
1,0	$Z_t$

### 2.1.3 Autocorrelation Function dan Partial Autocorrelation Function

*Autocorrelation Function* (ACF) adalah korelasi atau keeratan hubungan secara linier antara pengamatan  $Z_t$  dan  $Z_{t+k}$ . Plot ACF dapat digunakan untuk mendeteksi stasioneritas data *time series*, khususnya stasioner dalam rata-rata dan mengidentifikasi model *time series*. *Autocorrelation function* pada lag ke- $k$  dengan  $k = 0, 1, 2, \dots$  dihitung menggunakan persamaan (2.8) (Wei, 2006),

$$\rho_k = \frac{\text{Cov}(Z_t, Z_{t+k})}{\sqrt{\text{Var}(Z_t)}\sqrt{\text{Var}(Z_{t+k})}} = \frac{\gamma_k}{\gamma_0} \quad (2.8)$$

dengan  $\gamma_k$  adalah *covariance* antara  $Z_t$  dan  $Z_{t+k}$  (*autocovariance function* pada lag ke- $k$ ) serta  $\text{Var}(Z_t)$  dan  $\text{Var}(Z_{t+k})$  seperti dalam persamaan (2.9) dan (2.10),

$$\gamma_k = \text{Cov}(Z_t, Z_{t+k}) = E[(Z_t - \mu)(Z_{t+k} - \mu)] \quad (2.9)$$

dan

$$\text{Var}(Z_t) = \text{Var}(Z_{t+k}) = \gamma_0. \quad (2.10)$$

Nilai ACF sampel dari deret  $Z_1, Z_2, \dots, Z_n$  suatu data *time series* dapat didefinisikan dalam persamaan (2.11)

$$\hat{\rho}_k = \frac{\sum_{t=1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2} = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} \quad (2.11)$$

dengan  $\bar{Z} = \frac{\sum_{t=1}^n Z_t}{n}$  adalah rata-rata sampel dari deret  $Z_1, Z_2, \dots, Z_n$ . Plot dari nilai  $\hat{\rho}_k$  berdasarkan lag  $k$  disebut plot ACF sampel atau korelogram sampel.

*Partial Autocorrelation Function* (PACF) adalah korelasi atau keeratan hubungan antara pengamatan  $Z_t$  dan  $Z_{t+k}$  setelah menghilangkan dependensi linier dari pengamatan  $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$  sehingga PACF merupakan hubungan atau korelasi parsial antara  $Z_t$  dan  $Z_{t+k}$ . *Partial autocorrelation function* pada lag ke- $k$  dengan  $k = 0, 1, 2, \dots$  dihitung menggunakan persamaan (2.12) (Wei, 2006),

$$P_k = \frac{\text{Cov}[(Z_t - \hat{Z}_t), (Z_{t+k} - \hat{Z}_{t+k})]}{\sqrt{\text{Var}(Z_t - \hat{Z}_t)} \sqrt{\text{Var}(Z_{t+k} - \hat{Z}_{t+k})}} \quad (2.12)$$

dengan

$$\text{Var}(Z_{t+k} - \hat{Z}_{t+k}) = \text{Var}(Z_t - \hat{Z}_t) = \gamma_0 - \alpha_1 \gamma_1 - \dots - \alpha_{k-1} \gamma_{k-1} \quad (2.13)$$

dan

$$\text{Cov}[(Z_t - \hat{Z}_t), (Z_{t+k} - \hat{Z}_{t+k})] = \gamma_k - \alpha_1 \gamma_{k-1} - \dots - \alpha_{k-1} \gamma_1. \quad (2.14)$$

Nilai  $\hat{Z}_{t+k}$  sebagai estimasi dari  $Z_{t+k}$  dan  $\hat{Z}_{t+k} = \alpha_1 Z_{t+k-1} + \alpha_2 Z_{t+k-2} + \dots + \alpha_{k-1} Z_{t+1}$ .

Persamaan (2.12) dapat dinyatakan sebagai persamaan (2.15) dengan cara mensubstitusikan persamaan (2.13) dan (2.14) ke dalam persamaan (2.12), kemudian membagi pembilang dan penyebut masing-masing dengan  $\gamma_0$ .

$$P_k = \frac{\gamma_k - \alpha_1 \gamma_{k-1} - \dots - \alpha_{k-1} \gamma_1}{\gamma_0 - \alpha_1 \gamma_1 - \dots - \alpha_{k-1} \gamma_{k-1}} = \frac{\rho_k - \alpha_1 \rho_{k-1} - \dots - \alpha_{k-1} \rho_1}{1 - \alpha_1 \rho_1 - \dots - \alpha_{k-1} \rho_{k-1}} \quad (2.15)$$

Dengan menggunakan metode Cramer, persamaan (2.15) dapat dinyatakan sebagai rasio determinan seperti dalam persamaan (2.16).

$$P_k = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \quad (2.16)$$

Persamaan (2.16) dapat dinotasikan sebagai  $\phi_{kk}$  yang merupakan PACF antara  $Z_t$  dan  $Z_{t+k}$ .

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_1 & 1 \end{vmatrix}} \quad (2.17)$$

Nilai PACF sampel merupakan estimasi dari  $\phi_{kk}$  dalam persamaan (2.17) dan dinotasikan  $\hat{\phi}_{kk}$ . PACF sampel  $\hat{\phi}_{kk}$  diperoleh dengan cara mensubstitusikan  $\rho_i$  dengan  $\hat{\rho}_i$  yang diperoleh dengan metode rekursif, dimulai dengan  $\hat{\phi}_{11} = \hat{\rho}_1$ ,

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^k \hat{\phi}_{kj} \hat{\rho}_j} \quad (2.18)$$

dan

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1} \hat{\phi}_{k,k+1-j}, j = 1, 2, \dots, k \quad (2.19)$$

## 2.2 Autoregressive Integrated Moving Average (ARIMA)

Model *time series* secara umum dapat dibagi menjadi dua, yaitu model stasioner dan non-stasioner (dalam rata-rata). Model stasioner, yaitu model

*Autoregressive, Moving Average, dan Autoregressive Moving Average*. Model non-stasioner adalah model *Autoregressive Integrated Moving Average*.

**a. Model Autoregressive**

Model *Autoregressive* (AR) dapat dinyatakan sebagai model yang meregresikan antara pengamatan ke- $t$  dengan  $p$  pengamatan sebelumnya atau sering disebut sebagai model regresi dengan dirinya sendiri. Bentuk umum dari model AR pada orde ke- $p$  atau  $AR(p)$  (Wei, 2006) seperti persamaan (2.1) dengan  $\phi_p$  adalah parameter model AR pada orde ke- $p$  dan  $a_t$  adalah nilai *error* pada saat ke- $t$ . Nilai  $a_t$  diasumsikan independen terhadap  $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$ . Persamaan (2.1) dapat dinotasikan dalam persamaan (2.20)

$$\phi_p(B)Z_t = a_t \quad (2.20)$$

dengan

$$\phi_p(B) = 1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p \quad (2.21)$$

disebut sebagai operator  $AR(p)$  dengan  $B$  adalah operator *backshift*.

**b. Model Moving Average**

Model *Moving Average* (MA) memiliki bentuk umum pada orde ke- $q$  atau  $MA(q)$  seperti persamaan (2.22) (Wei, 2006)

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.22)$$

atau

$$Z_t = \theta_q(B)a_t \quad (2.23)$$

dengan

$$\theta_q(B) = 1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \quad (2.24)$$

disebut operator  $MA(q)$ .  $\theta_q$  adalah parameter model MA pada orde ke- $q$  serta  $a_t$  diasumsikan independen terhadap  $a_{t-1}, a_{t-2}, \dots, a_{t-q}$  dan berdistribusi normal dengan rata-rata nol dan variansi  $\sigma_a^2$ .

**c. Model Autoregressive Moving Average**

Model gabungan dari proses  $AR(p)$  dan  $MA(q)$  adalah model  $ARMA(p, q)$ . Bentuk umum model  $ARMA(p, q)$  adalah.

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q} \quad (2.25)$$

Persamaan (2.25) dapat dinyatakan dengan menggunakan operator AR( $p$ ) dan operator MA( $q$ ) dalam persamaan (2.21) dan (2.24) menjadi persamaan (2.26).

$$\phi_p(B)Z_t = \theta_q(B)a_t \quad (2.26)$$

Syarat proses ARMA( $p, q$ ) dikatakan invertibel dan stasioner adalah akar-akar dari  $\theta_q(B) = 0$  dan  $\phi_p(B) = 0$  berada di luar lingkaran satuan. Proses ARMA( $p, q$ ) yang stasioner dan invertibel dapat ditulis dalam bentuk AR( $\infty$ ) atau MA( $\infty$ ), yaitu masing-masing (Wei, 2006) adalah

$$\pi(B)Z_t = a_t \quad (2.27)$$

atau

$$Z_t = \psi(B)a_t \quad (2.28)$$

dengan

$$\pi(B) = \phi_p(B)\theta_q^{-1}(B) = (1 - \pi_1 B - \pi_2 B^2 - \dots)$$

dan

$$\psi(B) = \theta_q(B)\phi_p^{-1}(B) = (1 + \psi_1 B + \psi_2 B^2 + \dots).$$

#### d. Model ARIMA

Model ARIMA merupakan model *time series* non-stasioner dengan orde  $p$  sebagai orde dari AR, orde  $d$  merupakan *differencing* dan orde  $q$  sebagai orde dari MA. Bentuk persamaan model ARIMA terdapat dalam persamaan (2.29).

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t \quad (2.29)$$

dengan  $\phi_p(B)$  merupakan operator AR( $p$ ) seperti dalam persamaan (2.21) dan  $\theta_q(B)$  sebagai operator MA( $q$ ) dalam persamaan (2.24) (Wei, 2006). Model ini digunakan untuk data *time series* yang telah di-*differencing* atau sudah stasioner dalam rata-rata,  $d$  adalah banyaknya proses *differencing* yang dilakukan. Misalkan orde  $q = 0$ , model ARIMA ( $p, d, q$ ) yang terbentuk disebut juga sebagai model *Autoregressive Integrated* atau ARI( $p, d$ ). Jika orde  $p = 0$ , model tersebut disebut juga model *Integrated Moving Average* atau IMA ( $d, q$ ). Jika orde  $p = q = 0$  dengan *differencing*  $d = 1$ , model ARIMA(0, 1, 0) disebut model *random walk*.

#### 2.2.1 Identifikasi Model ARIMA

Identifikasi model secara umum bertujuan untuk mendapatkan model tentatif yang akan dievaluasi pada tahapan selanjutnya, yaitu estimasi parameter,

pengujian signifikansi parameter, dan pemeriksaan diagnostik. Langkah-langkah dalam menganalisis data *time series* dengan menggunakan metode ARIMA dijelaskan oleh Wei (2006).

Langkah pertama untuk mengidentifikasi model tentatif ARIMA adalah membuat plot data *time series* dan melakukan transformasi penstabilan variansi yang sesuai dengan menggunakan transformasi Box-Cox seperti dalam Tabel 2.1. Selanjutnya, menghitung ACF dan PACF sampel dari data asli (bukan data hasil transformasi) untuk mengetahui derajat *differencing* ( $d$ ) yang dibutuhkan agar data *time series* stasioner dalam rata-rata. Jika ACF sampel turun secara lambat dan PACF sampel *cut-off* setelah lag 1, data perlu di-*differencing* satu kali. Umumnya, proses *differencing* dilakukan maksimal sebanyak dua kali ( $d = 0, 1, \text{ atau } 2$ ). Langkah ketiga, yaitu mengidentifikasi orde  $p$  dan  $q$  dengan cara menghitung ACF dan PACF sampel dari data yang telah stasioner, lalu mencocokkannya dengan pola ACF dan PACF secara teoritik dalam Tabel 2.2. Langkah keempat adalah menguji tren deterministik  $\theta_0$  ketika  $d > 0$ .

**Tabel 2.2** Karakteristik Plot ACF-PACF pada model ARIMA

Proses	ACF	PACF
AR ( $p$ )	Turun cepat secara eksponensial atau sinus teredam	<i>Cut-off</i> setelah lag $p$
MA ( $q$ )	<i>Cut-off</i> setelah lag $q$	Turun cepat secara eksponensial atau sinus teredam
ARMA ( $p, q$ )	Turun cepat secara eksponensial atau sinusoidal	Turun cepat secara eksponensial atau sinus teredam
AR ( $p$ ) atau MA( $q$ )	<i>Cut-off</i> setelah lag $p$	<i>Cut-off</i> setelah lag $q$

### 2.2.2 Estimasi Parameter

Salah satu metode estimasi yang sering digunakan dalam model *time series* adalah metode *Conditional Least Squares*, yaitu suatu metode yang digunakan dengan cara meminimumkan jumlah kuadrat *error* (selisih nilai pengamatan dan nilai ramalan). Misalkan mengestimasi parameter dari model AR(1), yaitu  $Z_t - \mu = \phi_1(Z_{t-1} - \mu) + a_t$ . Parameter yang diestimasi adalah  $\mu$  dan  $\phi_1$ . Jumlah kuadrat *error*  $a_t$  dinotasikan sebagai  $S_c(\phi_1, \mu)$  seperti dalam persamaan (2.30).

$$S_c(\phi_1, \mu) = \sum_{t=2}^n [(Z_t - \mu) - \phi_1(Z_{t-1} - \mu)]^2 \quad (2.30)$$



Persamaan (2.30) diturunkan terhadap parameter  $\mu$  dan  $\phi_1$  serta menyamakan masing-masing turunan atau derivatif dengan nol.

$$\frac{\partial S_c(\phi_1, \mu)}{\partial \mu} = \sum_{t=2}^n 2[(Z_t - \hat{\mu}) - \hat{\phi}_1(Z_{t-1} - \hat{\mu})](-1 + \hat{\phi}_1) = 0 \quad (2.31)$$

$$\frac{\partial S_c(\phi_1, \mu)}{\partial \phi_1} = \sum_{t=2}^n -2[(Z_t - \hat{\mu}) - \hat{\phi}_1(Z_{t-1} - \hat{\mu})](Z_{t-1} - \hat{\mu}) = 0 \quad (2.32)$$

Dengan menggunakan metode *Conditional Least Squares*, estimasi dari parameter  $\mu$  adalah

$$\hat{\mu} \approx \frac{\bar{Z} - \hat{\phi}_1 \bar{Z}}{1 - \hat{\phi}_1} = \bar{Z} \quad (\text{untuk } n \text{ yang besar}). \quad (2.33)$$

Estimasi dari parameter  $\phi_1$  didapatkan dengan mensubstitusikan  $\hat{\mu} = \bar{Z}$  ke dalam persamaan (2.32) sehingga diperoleh estimator  $\hat{\phi}_1$  seperti dalam persamaan (2.34).

$$\hat{\phi}_1 = \frac{\sum_{t=2}^n (Z_t - \bar{Z})(Z_{t-1} - \bar{Z})}{\sum_{t=2}^n (Z_{t-1} - \bar{Z})^2} = \hat{\rho}_1. \quad (2.34)$$

### 2.2.3 Pengujian Signifikansi Parameter

Setelah melakukan estimasi parameter, tahap berikutnya adalah menguji signifikansi parameter. Hipotesis yang digunakan untuk menguji signifikansi parameter model, misalnya model AR(1) adalah.

$$H_0 : \phi_1 = 0 \quad (\text{parameter model AR(1) tidak signifikan})$$

$$H_1 : \phi_1 \neq 0 \quad (\text{parameter model AR(1) signifikan})$$

Statistik uji yang digunakan adalah statistik uji T.

$$T = \frac{\hat{\phi}_1}{SE(\hat{\phi}_1)} \quad (2.35)$$

Keputusan menolak  $H_0$ , ketika  $|T| > t_{\frac{\alpha}{2}; n-M}$  atau *P-value*  $< \alpha$  sehingga parameter  $\phi_1$  signifikan, dengan  $n$  banyaknya pengamatan dan  $M$  banyak parameter.

### 2.2.4 Pemeriksaan Diagnostik

Setelah mengetahui parameter yang signifikan, tahap selanjutnya adalah pemeriksaan diagnostik residual, yaitu uji *white noise* dan uji distribusi normal.

### 2.2.4.1 Uji Asumsi *White Noise*

*White noise* merupakan asumsi yang harus dipenuhi dalam membuat model ARIMA, yaitu bermakna gangguan-gangguan terhadap residual telah dihilangkan. Residual dari suatu model ARIMA dikatakan *white noise*, jika antar residual saling identik dan independen. Proses ARIMA yang *whitenoise* diasumsikan memiliki rata-rata nol,  $E(a_t) = \mu_a = 0$  dan variansi yang konstan  $Var(a_t) = \sigma_a^2$  serta kovariansi  $\gamma_k = cov(a_t, a_{t-k}) = 0$  untuk  $k \neq 0$ .

Untuk menguji asumsi *white noise* dilakukan uji *Ljung-Box* dengan hipotesis sebagai berikut.

$$H_0: \rho_1 = \rho_2 = \dots = \rho_K = 0$$

$$H_1: \text{minimal ada satu } \rho_k \neq 0, k = 1, 2, \dots, K$$

Statistik uji yang digunakan adalah statistik uji Q seperti dalam persamaan (2.36)

$$Q = n(n+2) \sum_{k=1}^K \frac{\hat{\rho}_k^2}{n-k} \quad (2.36)$$

dengan  $n$  banyak pengamatan dan  $\hat{\rho}_k$  adalah autokorelasi pada lag ke- $k$ . Keputusan menolak  $H_0$ , jika  $Q > \chi_{\alpha; (K-m)}^2$  dengan  $m = p+q$  banyaknya lag atau  $P\text{-value} < \alpha$ . Jadi, dapat disimpulkan residual tidak memenuhi asumsi *whitenoise*.

### 2.2.4.2 Uji Asumsi Distribusi Normal

Uji asumsi distribusi normal dari residual model ARIMA dilakukan untuk mengetahui residual berdistribusi normal atau tidak. Pengujian dapat dilakukan dengan menggunakan uji *Kolmogorov-Smirnov* dengan hipotesis sebagai berikut.

$$H_0: F(a_t) = F_0(a_t)$$

$$H_1: F(a_t) \neq F_0(a_t)$$

Statistik uji *Kolmogorov-Smirnov* (uji *KS*) seperti dalam persamaan (2.37)

$$D = \sup_{a_t} |S(a_t) - F_0(a_t)| \quad (2.37)$$

dengan  $S(a_t)$  adalah fungsi probabilitas kumulatif data sampel,  $F_0(a_t)$  fungsi distribusi  $a_t$  yang dihipotesiskan, dan  $F(a_t)$  fungsi distribusi  $a_t$  yang belum diketahui.  $H_0$  ditolak, jika nilai  $D > K_{(1-\alpha; n)}$  dengan  $K_{(1-\alpha; n)}$  adalah nilai tabel *KS* pada kuantil  $(1-\alpha)$  dan  $n$  residual, artinya residual tidak berdistribusi normal.

### 2.2.5 Seleksi Model Terbaik

Pemilihan model terbaik digunakan untuk memilih dari dua atau lebih model yang sama baiknya, yaitu misalnya sama-sama memiliki parameter yang signifikan dan memenuhi semua asumsi residual. Oleh karena itu, diperlukan kriteria memilih model terbaik, yaitu menurut kriteria data *in sample* dan *out sample*.

#### 2.2.5.1 Berdasarkan Data *In Sample*

Kriteria yang digunakan untuk data *in sample* adalah AIC dan SBC. AIC (*Akaike's Information Criterion*) merupakan kriteria pemilihan model terbaik yang mempertimbangkan banyaknya parameter dalam model. Semakin kecil nilai AIC, semakin besar kemungkinan suatu model ARIMA tersebut akan dipilih (Wei, 2006). Nilai AIC dirumuskan seperti dalam persamaan (2.38).

$$AIC(M) = -2 \ln L(\phi, \mu, \theta, \sigma_a^2 | \mathbf{a}) + 2M \quad (2.38)$$

dengan  $M$  adalah banyaknya parameter dan  $\ln L$  adalah fungsi log-likelihood, yaitu

$$\ln L(\phi, \mu, \theta, \sigma_a^2 | \mathbf{a}) = -\frac{n}{2} \ln 2\pi\sigma_a^2 - \frac{1}{2\sigma_a^2} S(\phi, \mu, \theta)$$

dengan  $S(\phi, \mu, \theta) = \sum_{i=-M}^n [E(a_i | \phi, \mu, \theta, \mathbf{Z})]^2$  adalah fungsi *unconditional sum of squares*.

SBC (*Schwartz's Bayesian Criterion*) adalah kriteria pemilihan model terbaik yang mempertimbangkan banyaknya parameter dalam model dengan menggunakan sampel kecil. Sama halnya dengan AIC, semakin kecil nilai SBC, semakin besar kemungkinan model tersebut dipilih. Nilai SBC dirumuskan dalam persamaan (2.39)

$$SBC(M) = -2 \ln L(\phi, \mu, \theta, \sigma_a^2 | \mathbf{a}) + M \ln n \quad (2.39)$$

#### 2.2.5.2 Berdasarkan Data *Out Sample*

Untuk data *out sample*, kriteria yang digunakan adalah MSE (*Mean Square Error*) dan MAPE (*Mean Absolute Percentage Error*). MSE merupakan rata-rata kuadrat residual dari hasil peramalan dan MAPE merupakan kriteria kebaikan hasil peramalan dalam bentuk persentase, yaitu.

$$MSE = \frac{1}{l} \sum_{h=1}^l [e_n(h)]^2 \quad (2.40)$$

$$MAPE = \frac{\sum_{h=1}^l [e_n(h)/Z_{n+h}]^2}{l} 100\% \quad (2.41)$$

Nilai  $l$  adalah banyaknya residual dari hasil peramalan dan  $e_n(h) = Z_{n+h} - \hat{Z}_n(h)$  adalah residual peramalan pada  $h$  waktu ke depan dari  $n$  pengamatan aktual. Nilai  $Z_{n+h}$  adalah nilai pengamatan aktual pada waktu ke- $(n+h)$  dan  $\hat{Z}_n(h)$  adalah nilai ramalan  $h$  waktu ke depan dari  $Z_{n+h}$  dengan  $n$  pengamatan aktual. Semakin kecil nilai MSE dan MAPE, semakin besar kemungkinan model tersebut dipilih.

### 2.2.6 Peramalan Model ARIMA

Tujuan utama dari membuat model *time series*, khususnya model ARIMA adalah meramalkan hasil pengamatan pada waktu mendatang dengan *error* (MSE) ramalan sekecil mungkin. MSE minimum hasil ramalan ( $\hat{Z}_t(l)$ ) dari  $Z_{t+l}$  bersyarat ekspektasi kondisionalnya (Wei, 2006), yaitu

$$\hat{Z}_t(l) = E(Z_{t+l} | Z_t, Z_{t-1}, \dots, Z_1) \quad (2.42)$$

jika  $t = n$ , persamaan (2.41) dapat dinyatakan menjadi

$$\hat{Z}_n(l) = E(Z_{n+l} | Z_n, Z_{n-1}, \dots, Z_1)$$

serta peramalan  $l$  tahap ke depan yang meminimalkan MSE hasil ramalan, yaitu.

$$\hat{Z}_t(l) = \sum_{j=0}^{\infty} \psi_{l+j}^* a_{t-j} \quad (2.43)$$

### 2.3 Generalized Autoregressive Conditional Heteroscedasticity

Metode *Autoregressive Conditional Heteroscedasticity* (ARCH) diusulkan pertama kali oleh Engle (1982) untuk menangani adanya kasus heteroskedastisitas pada data *time series*, khususnya data yang memiliki volatilitas yang tinggi. Model ARCH digunakan untuk memodelkan volatilitas dengan asumsi bahwa *conditional* variansi hari ini ( $v_t^2$ ) berhubungan dengan variansi periode sebelumnya. Model ARCH( $s$ ) ditunjukkan dalam persamaan (2.44),

$$\begin{aligned} \varepsilon_t &= v_t \zeta_t \\ v_t^2 &= \omega_0 + \omega_1 \varepsilon_{t-1}^2 + \dots + \omega_s \varepsilon_{t-s}^2 \end{aligned} \quad (2.44)$$

$\varepsilon_t$  adalah residual yang tidak saling berkorelasi, tetapi memiliki variansi yang berubah seiring waktu dan  $\zeta_t$  variabel random dengan rata-rata 0 dan variansi 1.

Metode ARCH dikembangkan oleh Bollerslev (1986) menjadi *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). Dalam model GARCH, variansi residual tidak hanya berkorelasi dengan variansi residual periode lalu tetapi juga dengan residual kuadrat periode lalu, diperoleh model GARCH (r,s) seperti yang ditunjukkan dalam persamaan (2.45).

$$\begin{aligned}\varepsilon_t &= v_t \zeta_t \\ v_t^2 &= \omega_0 + \xi_1 v_{t-1}^2 + \dots + \xi_r v_{t-r}^2 + \omega_1 \varepsilon_{t-1}^2 + \dots + \omega_s \varepsilon_{t-s}^2\end{aligned}\quad (2.45)$$

Nilai  $\zeta_t$  merupakan variabel random dengan rata-rata 0 dan variansi 1. Untuk memastikan  $v_t^2 > 0$ , diasumsikan  $\omega_0 > 0$ ,  $\omega_j \geq 0$ ,  $\xi_i \geq 0$  untuk  $i, j > 0$ ,  $r < p$ ,  $s < q$ .

Efek heteroskedastisitas atau efek ARCH yang diduga terdapat dalam residual model ARIMA diuji dengan uji *Lagrange Multiplier* (LM). Hipotesis yang digunakan uji LM adalah.

$$H_0: \omega_1 = \omega_2 = \dots = \omega_s = 0 \text{ (tidak terdapat efek ARCH)}$$

$$H_1: \text{minimal ada satu } \omega_j \neq 0, j = 1, 2, \dots, s \text{ (terdapat efek ARCH)}$$

Statistik uji yang digunakan adalah (Wei, 2006).

$$LM = nR^2 \quad (2.46)$$

Nilai  $n$  adalah banyaknya pengamatan dan  $R^2$  adalah besar koefisien determinasi.

$$R^2 = \frac{\text{regression sum of squares}}{\text{total sum of squares}} = \frac{\sum_{t=1}^n (\hat{Z}_t - \bar{Z})^2}{\sum_{t=1}^n (Z_t - \bar{Z})^2} \quad (2.47)$$

Keputusan menolak  $H_0$ , jika nilai  $LM$  lebih besar dari nilai  $\chi_{(\alpha,s)}^2$ , artinya terdapat efek ARCH atau data terjadi heteroskedastisitas.

### 2.3.1 Peramalan Variansi Residual

Dalam analisis model GARCH, variansi residual yang tidak konstan atau terjadi heteroskedastisitas dimodelkan seperti persamaan (2.45). Model GARCH digunakan untuk meramalkan variansi residual yang berubah seiring waktu.

$$(1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p) e_t = \varepsilon_t \quad (2.48)$$

dengan  $e_t = \sum_{j=0}^{\infty} \psi_t B^j \varepsilon_t$

dengan  $\psi_0 = 1$  dan  $\sum_{j=0}^{\infty} \psi_t B^j = \frac{1}{1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p}$ . Dengan demikian,

ramalan variansi residual pada  $l$  langkah ke depan adalah sebagai berikut,

$$\text{Var}_t(e_{t+l}) = E_t[e_{t+l} - E_t(e_{t+l})]^2 = E_t\left[\sum_{j=0}^{l-1} \psi_j \varepsilon_{t+l-j}\right]^2 \quad (2.49)$$

dan diperoleh variansi  $e_{t+l}$  adalah

$$\text{Var}_t(e_{t+l}) = E_t\left[\sum_{j=0}^{l-1} \psi_j v_{t+l-j} e_{t+l-j}\right]^2 = \sum_{j=0}^{l-1} \psi_j^2 \varepsilon_{t+l-j}^2 \quad (2.50)$$

dengan  $v_{t+l-j}^2 = E(\varepsilon_{t+l-j}^2)$  sehingga

$$\begin{aligned} v_{t+l-j}^2 &= E_t(\varepsilon_{t+l-j}^2) \\ &= \omega_0 + \sum_{i=1}^g \alpha_i E_t(\varepsilon_{t+l-j-i}^2) + E_t(a_{t+l-j}) - \sum_{i=1}^r \xi_i E_t(a_{t+l-j-i}) \end{aligned} \quad (2.51)$$

dengan  $g$  adalah maksimum dari  $(r, s)$ .

## 2.4 Algoritma Genetika

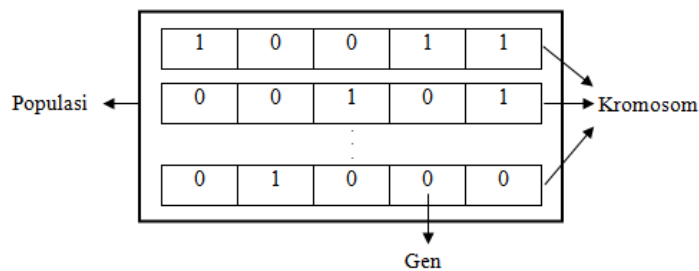
Algoritma Genetika atau *Genetic Algorithm* (GA) merupakan algoritma pembelajaran (*machine learning*) yang diperkenalkan oleh John Holland pada tahun 1975. Algoritma ini mengadopsi dari sistem genetika makhluk hidup yang terdapat operator reproduksi, *crossover*, dan mutasi dengan menggunakan prinsip seleksi alam, yaitu individu yang dapat bertahan hidup adalah individu yang kuat. Algoritma genetika diawali dengan melakukan inisialisasi himpunan solusi (populasi) dengan cara membangkitkan himpunan solusi secara acak. Dalam suatu populasi yang terdiri dari individu-individu, setiap individu merupakan representasi sebuah solusi dari permasalahan. Individu-individu tersebut dinotasikan sebagai kromosom yang terdiri dari sekumpulan gen. Menurut Engelbrecht (2002), kinerja akurasi dan waktu yang dibutuhkan untuk algoritma konvergen semakin besar, jika ukuran populasi juga semakin besar.

Untuk mendapatkan solusi terbaik, algoritma akan menyeleksi setiap individu dengan menggunakan kriteria nilai *fitness*. Setelah melakukan proses seleksi berdasarkan nilai *fitness*, individu akan melakukan proses *crossover* sehingga

akan memperoleh keturunan atau anak yang akan digunakan pada generasi berikutnya dengan menghitung nilai *fitness*-nya. Selain proses *crossover*, terdapat operator mutasi yang berfungsi untuk memperluas kemampuan algoritma dalam mencari solusi terbaik (Hung, 2011). Dengan demikian, akan terbentuk individu-individu baru dalam satu populasi pada generasi selanjutnya. Proses regenerasi tersebut berlangsung hingga didapatkan solusi yang terbaik yang ditandai dengan algoritma sudah konvergen ke arah kromosom terbaik (Gen dan Cheng, 2000). Pada setiap generasi, algoritma mengeksplorasi banyak area dari ruang parameter, kemudian memilih daerah yang memiliki probabilitas terbesar. GA mampu mengevaluasi banyak titik dalam ruang parameter secara simultan sehingga akan konvergen pada solusi yang optimal dan global (Hung, 2009).

#### 2.4.1 Kromosom

Kromosom merupakan bagian penting dalam GA karena satu kromosom atau individu mewakili satu solusi dari suatu permasalahan. Setiap kromosom disusun oleh gen-gen dengan panjang setiap kromosom adalah sama seperti yang ditunjukkan dalam Gambar 2.1. Setiap gen dalam masing-masing kromosom mewakili elemen dari solusi. Representasi kromosom dapat berupa bilangan biner, integer, ataupun bilangan riil yang dapat diperoleh melalui proses pengkodean.



**Gambar 2.1** Ilustrasi Kromosom dengan Bilangan Biner

Pada setiap generasi, kromosom dievaluasi dengan menggunakan fungsi *fitness*. Kromosom-kromosom yang diperoleh pada tahap inisialisasi akan digunakan untuk menghasilkan generasi baru dengan cara menggabungkan dua kromosom dari generasi sekarang dengan menggunakan operator *crossover* atau mengubah gen dalam sebuah kromosom dengan menggunakan operator mutasi. Generasi baru dibentuk dengan cara seleksi yang dilakukan terhadap induk (*parent*) dan keturunan (*offspring*) berdasarkan nilai *fitness* masing-masing. Proses regenerasi terus berlangsung sampai mendapatkan hasil yang konvergen ke

arah bentuk kromosom yang terbaik sehingga kromosom tersebut diharapkan menjadi solusi global dan optimal dari permasalahan.

#### 2.4.2 Fungsi *Fitness*

Fungsi *fitness* digunakan untuk mengukur tingkat kebaikan dari solusi yang sedang dicari dari suatu himpunan solusi. Sejumlah solusi yang dibangkitkan dalam populasi dievaluasi menggunakan nilai *fitness*. Kromosom yang memiliki nilai *fitness* tertinggi merupakan solusi terbaik. Pada setiap iterasi, algoritma akan mencari dan mempertahankan kromosom yang memiliki nilai *fitness* tertinggi.

Fungsi tujuan dalam estimasi parameter GARCH dengan GA adalah mendapatkan nilai *Root Mean Square Error* (RMSE) minimum, yaitu

$$RMSE = \sqrt{\frac{1}{N-n} \sum_{t=n+1}^N (\varepsilon_t^2 - \hat{v}_t^2)^2} \quad (2.52)$$

Oleh sebab itu, fungsi *fitness* yang digunakan untuk mengevaluasi setiap kromosom adalah

$$fitness = \frac{1}{RMSE} \quad (2.53)$$

dengan  $\varepsilon_t$  residual pengamatan pada waktu  $t$ , nilai ramalan *conditional variance*  $\hat{v}_t^2$ ,  $N$  merupakan semua pengamatan (*in-sample* dan *out-sample*), serta  $n$  merupakan pengamatan *in-sample* (Souza, Veiga, dan Medeiros, 2005).

#### 2.4.3 Seleksi dengan Roda *Roulette*

Metode seleksi dengan roda *Roulette* adalah salah satu cara untuk menyeleksi individu. Individu-individu dipetakan dalam segmen garis secara berurutan dan memiliki ukuran yang sama dengan nilai *fitness*.

Individu yang memiliki segmen dalam area bilangan random, individu akan terseleksi dan masuk dalam populasi. Proses diulang hingga mendapatkan sejumlah individu tertentu.

#### 2.4.4 *Crossover* atau Pindah Silang

*Crossover* adalah operator primer dalam GA yang memfasilitasi adanya eksplorasi daerah baru dalam ruang pencarian (Hung, 2011). Namun, *crossover* yang dilakukan untuk populasi yang berukuran kecil dapat berdampak tidak baik, seperti kromosom yang memiliki gen-gen yang mengarah kepada solusi akan cepat menyebar ke kromosom lainnya. Maka dari itu, proses *crossover* dibatasi



dengan ketentuan bahwa *crossover* dapat dilakukan jika suatu bilangan random  $[0,1)$  kurang dari suatu probabilitas  $P_c$  atau laju *crossover* tertentu, umumnya mendekati 1, misalnya 0,6; 0,8; dan 0,9.

1	0	0	1	1	Induk 1
1	1	1	0	1	Induk 2
1	0	1	0	1	Anak 1
1	1	0	1	1	Anak 2

**Gambar 2.2** Ilustrasi Proses *Crossover*

#### 2.4.5 Mutasi

Mutasi adalah operator sekunder dalam GA dan menghindarkan dari terjadinya berhentinya algoritma pada solusi lokal serta memperluas kemampuan algoritma untuk mencari solusi optimal (Hung, 2011). Mutasi dilakukan dengan cara membalik arah dari setiap gen dari kromosom, yaitu menukar 0 menjadi 1, atau sebaliknya.

Tingginya tingkat mutasi yang dilakukan dapat menyebabkan terjadinya degradasi algoritma menjadi pencarian secara random sehingga mutasi dilakukan seperlunya. Dengan demikian, proses mutasi dibatasi dengan ketentuan bahwa mutasi dapat dilakukan jika suatu bilangan random  $[0,1)$  kurang dari  $P_m$  atau laju mutasi tertentu, umumnya mendekati 0, misalnya 0,1; 0,01; dan 0,05.

1	0	0	1	1
0	0	0	1	1

**Gambar 2.3** Ilustrasi Proses Mutasi

#### 2.4.6 Elitisme

Suatu individu yang telah terseleksi karena memiliki nilai *fitness* yang tinggi tidak selalu akan masuk dalam populasi untuk generasi selanjutnya. Hal ini disebabkan karena nilai *fitness* tersebut memungkinkan untuk mengalami penurunan sehingga individu tersebut hilang. Oleh karena itu, dilakukan proses elitisme, yaitu mengandakan individu yang memiliki nilai *fitness* tinggi tersebut agar individu tersebut tetap bertahan/ terpilih pada generasi selanjutnya.

### 2.5 Kecepatan Angin

Angin merupakan udara yang bergerak dari tekanan tinggi ke tekanan yang lebih rendah serta dapat pula diakibatkan adanya rotasi bumi. Alat yang

digunakan untuk mengukur kecepatan angin salah satunya adalah anemometer dengan cara meletakkannya pada ketinggian seratus meter di atas permukaan air. Kecepatan angin di Indonesia rata-rata sekitar 4 sampai 5 meter per detik. Kecepatan angin penting untuk diketahui, khususnya sebagai prakiraan cuaca, mendeteksi polusi udara, penerbangan maskapai pesawat, pelayaran, dan lain-lain.

Beberapa penelitian telah dilakukan untuk meramalkan kecepatan angin, baik menggunakan model *time series*, seperti ARIMA dan GARCH maupun dengan metode *intelligent algorithm*, yaitu *Support Vector Machine*, *Support Vector Regression*, dan ANN, serta dengan metode Bayesian dan ANFIS. Penelitian tersebut dilakukan dengan tujuan untuk mendapatkan hasil peramalan yang lebih akurat.

## BAB 3

### METODOLOGI PENELITIAN

#### 3.1 Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian adalah data sekunder dari Tugas Akhir Diah Kusumawati (2011) yang berjudul Peramalan Kecepatan Angin Rata-Rata Harian di Surabaya Menggunakan Metode *Bayesian Model Averaging* dengan Pendekatan *Expectation Maximization*. Data kecepatan angin rata-rata harian (km/jam) yang digunakan merupakan data Stasiun Meteorologi Juanda Surabaya mulai dari tanggal 1 Januari 2008 sampai dengan 31 Desember 2009. Struktur data dalam penelitian ini ditunjukkan oleh Tabel 3.1.

**Tabel 3.1** Struktur Data Kecepatan Angin Rata-rata

$T$	Tanggal	Kecepatan Angin Rata-rata ( $Z_t$ )	Out sample/ in Sample
1	1 Januari 2008	$Z_1$	In sample
2	2 Januari 2008	$Z_2$	
⋮	⋮	⋮	
$n = 700$	30 November 2008	$Z_n$	
$n+1$	1 Desember 2009	$Z_{n+1}$	Out sample
$n+2$	2 Desember 2009	$Z_{n+2}$	
⋮	⋮	⋮	
$n+31$	31 Desember 2009	$Z_{n+31}$	

#### 3.2 Langkah Analisis

Langkah analisis penelitian diterapkan untuk menganalisis data kecepatan angin rata-rata dan data simulasi seperti dalam subbab 3.2.1. dan subbab 3.2.2

##### 3.2.1 Data Kecepatan Angin Rata-rata

Data kecepatan angin rata-rata diramalkan menggunakan metode ARIMA-GARCH dan ARIMA-GARCH GA dengan langkah analisis sebagai berikut.

- 1) Mengumpulkan data harian kecepatan angin rata-rata pada.
- 2) Menganalisis secara deskriptif, yaitu menghitung rata-rata, variansi, nilai minimum dan maksimum serta membuat *box-plot* dan histogram kecepatan angin rata-rata.

3) Melakukan pembagian data *in sample* untuk periode 1 Januari 2008 – 30 November 2008 dan *out sample* untuk periode 1 – 31 Desember 2009.

4) Memodelkan data dengan menggunakan ARIMA-GARCH.

Sebelum memodelkan data kecepatan angin rata-rata dengan model GARCH, dilakukan pemodelan dengan ARIMA, sebagai berikut.

- a. Membuat plot *time series* dari data *in sample* kecepatan angin rata-rata untuk mengetahui pola data secara visual, yaitu apakah terdapat pola tren, musiman, atau lainnya. Jika data berfluktuasi di sekitar nilai rata-rata dan akan kembali ke nilai rata-rata (*mean reversion*), data dikatakan sudah stasioner. Namun, jika variansi data tidak stasioner, data perlu ditransformasi dengan transformasi Box-Cox seperti dalam Tabel 2.1.
- b. Menghitung ACF dan PACF dari data asli (bukan data hasil transformasi) untuk mengetahui derajat *differencing* ( $d$ ) yang dibutuhkan agar data *time series* stasioner dalam rata-rata. Jika ACF sampel turun secara lambat dan PACF sampel *cut-off* setelah lag 1, maka dibutuhkan *differencing* satu kali. Selain menggunakan plot *time series* dan ACF, pengujian stasioneritas dalam rata-rata dilakukan dengan menggunakan uji *Augmented Dickey Fuller* untuk mengetahui apakah data stasioner dalam rata-rata atau tidak.
- c. Melakukan identifikasi orde  $p$  dan  $q$  dengan cara menghitung ACF dan PACF dari data yang telah stasioner, lalu mencocokkannya secara teoritik dalam Tabel 2.2 sehingga didapatkan beberapa model tentatif.
- d. Melakukan estimasi parameter dari model ARIMA (model tentatif) dari langkah 4c menggunakan metode estimasi *Conditional Least Square*.
- e. Melakukan pengujian signifikansi parameter model ARIMA.
- f. Melakukan pengecekan diagnostik, yaitu asumsi *white noise* dan distribusi normal menggunakan uji *Ljung-Box* dan uji *Kolmogorov-Smirnov*.
- g. Memilih model ARIMA terbaik dari beberapa model tentatif dengan menggunakan kriteria evaluasi model dan pertimbangan model *parsimony*. Setelah didapatkan model ARIMA, langkah selanjutnya adalah memodelkan data dengan GARCH, yaitu.
- h. Menguji adanya efek heteroskedastisitas dari residual kuadrat model ARIMA terbaik dari langkah 4g dengan uji *Lagrange Multiplier*.

- i. Membentuk model GARCH dengan orde  $s$  dan  $r$  tertentu.
  - j. Melakukan estimasi parameter dan pengujian signifikansi parameter model GARCH.
  - k. Melakukan evaluasi kebaikan model GARCH.
  - l. Menghitung peramalan dengan model GARCH terpilih.
- 5) Memodelkan data dengan GARCH-GA
- Model GARCH-GA dimodelkan dengan orde GARCH yang telah dibentuk pada pemodelan GARCH pada langkah 4. Langkah berikutnya adalah mengestimasi parameter model dengan GA sebagai berikut.
- a. Menentukan fungsi *fitness*,  $P_c$ ,  $P_m$ , dan banyak generasi/ iterasi.
  - b. Merepresentasikan parameter model GARCH ke dalam kromosom dan melakukan inisialisasi.
  - c. Melakukan proses seleksi kromosom dengan menggunakan metode roda *Roulette* sehingga diperoleh kromosom yang menjadi induk/ orang tua.
  - d. Melakukan *crossover*, mutasi, dan elitisme.
  - e. Melakukan pergantian populasi lama dengan populasi baru dengan memilih kromosom yang memiliki nilai *fitness* terbesar atau RMSE terkecil.
  - f. Melakukan langkah 5d dan 5e terhadap kromosom sampai mendapatkan kromosom terbaik, yaitu nilai *fitness* terbesar pada iterasi terakhir.
  - g. Meramalkan kecepatan angin untuk data *out sample* dengan model GARCH-GA selama periode ke depan dan menghitung nilai RMSE.
- 6) Melakukan perbandingan kebaikan model GARCH dan GARCH-GA berdasarkan kriteria RMSE.

### 3.2.2 Data Simulasi

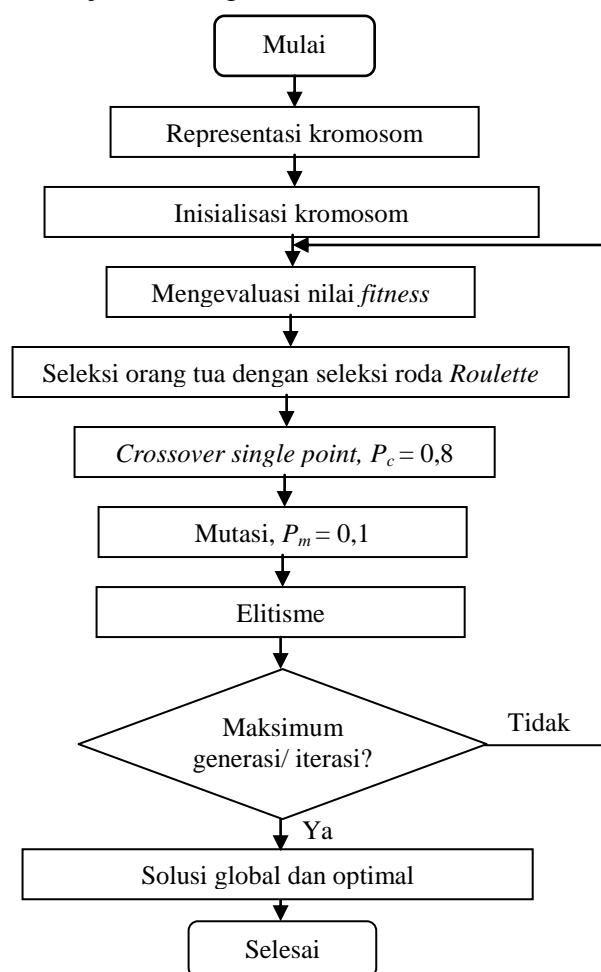
Langkah analisis penelitian yang diterapkan untuk menganalisis data yang diperoleh dari hasil simulasi.

- 1) Menentukan nilai  $P_c$ ,  $P_m$ , banyak generasi, dan fungsi objektif yang meminimumkan nilai RMSE atau mendapatkan fungsi *fitness* terbesar.
- 2) Membangkitkan data simulasi model GARCH(1,1) dengan skenario parameter yang ditentukan untuk ukuran sampel 300 dan 1000.
- 3) Mereplikasi langkah kedua sebanyak 10 kali.
- 4) Membagi data menjadi data *in sample* dan *out sample*.

- 5) Melakukan estimasi parameter model dengan GA sebagai berikut.
  - a. Menentukan fungsi *fitness*,  $P_c$ ,  $P_m$ , dan banyak generasi yang digunakan.
  - b. Merepresentasikan parameter model GARCH ke dalam kromosom dan melakukan inialisasi.
  - c. Melakukan proses seleksi kromosom.
  - d. Melakukan *crossover*, mutasi, dan elitisme.
  - e. Melakukan pergantian populasi lama dengan populasi baru dengan cara memilih kromosom yang memiliki nilai *fitness* terkecil.
  - f. Melakukan langkah 4d dan 4e terhadap kromosom sampai mendapatkan kromosom terbaik, yaitu nilai RMSE terkecil pada iterasi terakhir.
- 6) Membandingkan kebaikan hasil ramalan GARCH dan GARCH-GA.

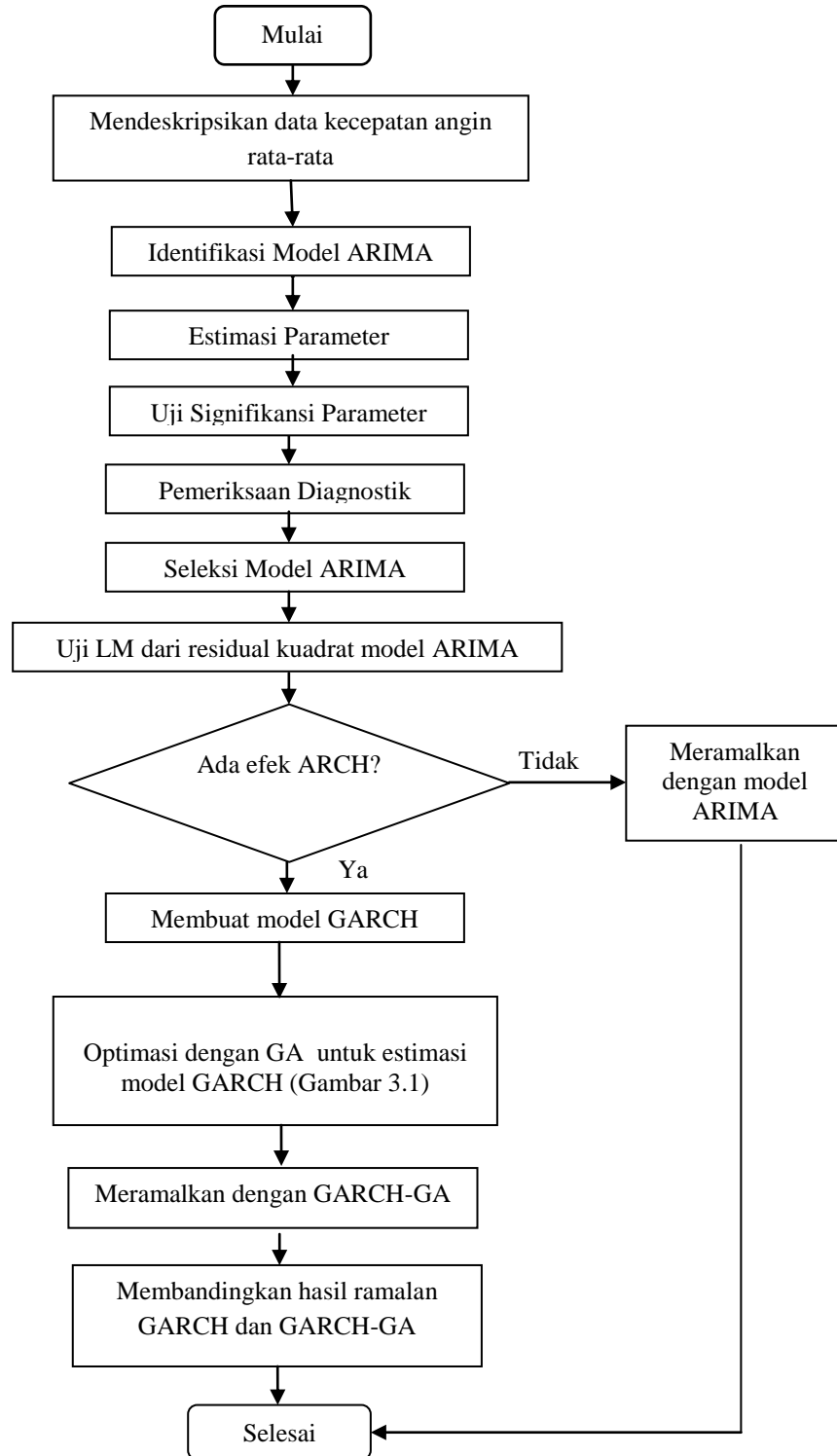
### 3.3 Diagram Alir

Gambar 3.1 menunjukkan diagram alir dari GA.



**Gambar 3.1** Diagram Alir GA

Data kecepatan angin rata-rata dianalisis sesuai dengan diagram alir dalam Gambar 3.2 hingga diperoleh model GARCH, GARCH-GA, dan perbandingan hasil ramalan dari GARCH dan GARCH-GA.



**Gambar 3.2** Diagram Alir Peramalan Data Kecepatan Angin Rata-rata dengan GARCH-GA

*(Halaman ini sengaja dikosongkan)*

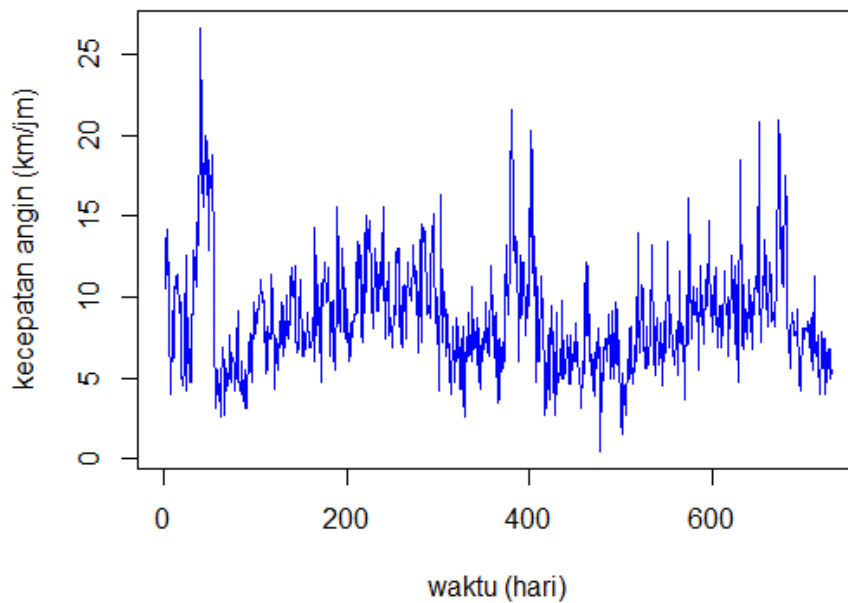


## BAB 4

### ANALISIS DAN PEMBAHASAN

#### 4.1 Karakteristik Kecepatan Angin

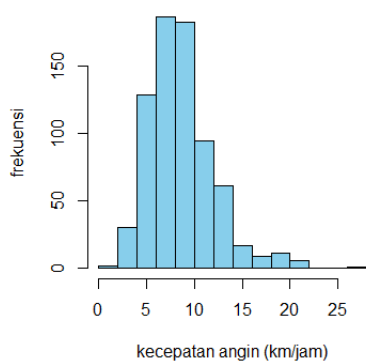
Kecepatan angin rata-rata di Juanda berfluktuasi di atas nilai nol dan tidak ada kecenderungan tren naik ataupun turun seperti yang terlihat dalam Gambar 4.1, dengan rata-rata (*mean*) kecepatan angin sebesar 8,662 km/jam. Variansi kecepatan angin rata-rata sebesar 11,482 km/jam. Kecepatan angin terbesar terjadi pada 9 Februari 2008 sebesar 26,6 km/jam dan terkecil pada 22 April 2009 sebesar 0,5 km/jam. Kecepatan angin yang sangat kencang dan tidak stabil dapat berdampak negatif terhadap lingkungan, misalnya memengaruhi penerbangan pesawat di sekitar stasiun BMKG Juanda.



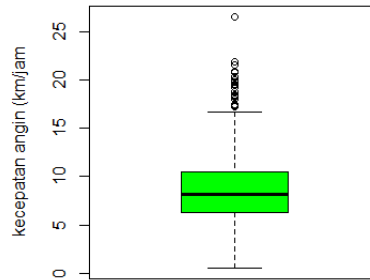
**Gambar 4.1** Kecepatan Angin Rata-rata Juanda Surabaya

Sebaran data kecepatan angin rata-rata miring ke kanan (*right skewed*) sebesar 1,16. Kemencengan data kecepatan angin dapat ditunjukkan oleh adanya ekor (*tail*) pada sisi kanan nilai rata-rata pada histogram dalam Gambar 4.2. Hal ini menunjukkan bahwa terdapat kecepatan angin rata-rata yang cukup besar, nilai tersebut diidentifikasi sebagai *outlier* yang ditunjukkan oleh lingkaran-lingkaran di atas garis *whisker* dalam Gambar 4.3. Bentuk kotak pada *boxplot* relatif kecil dan sama antara nilai di atas rata-rata dan di bawah rata-rata sehingga sebaran data cukup simetris tanpa memperhatikan amatan *outlier*. Data *outlier* tersebut yang

menjadikan sebaran data menjadi menceng ke kanan. Namun, data *outlier* tersebut tetap digunakan dalam analisis lebih lanjut.



**Gambar 4.2** Histogram Kecepatan Angin Rata-rata



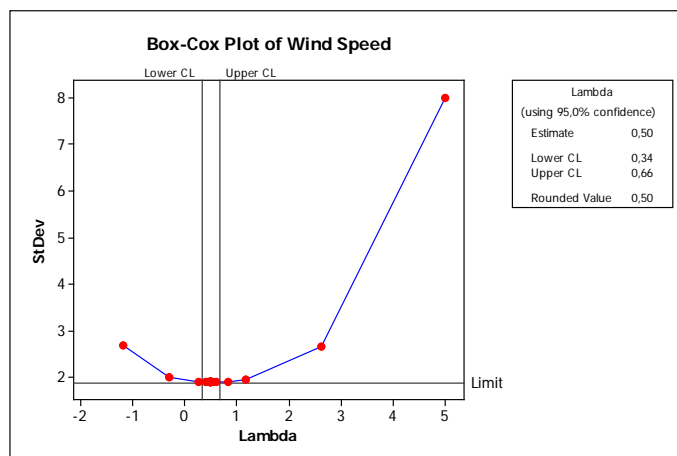
**Gambar 4.3** Boxplot Kecepatan Angin Rata-rata

## 4.2 Model ARIMA

Kecepatan angin rata-rata dimodelkan menggunakan ARIMA, yaitu kecepatan angin hari ini dimodelkan dari kecepatan angin periode sebelumnya dan/atau dengan residual periode sebelumnya. Tahapan dijelaskan dalam Subbab 2.2.

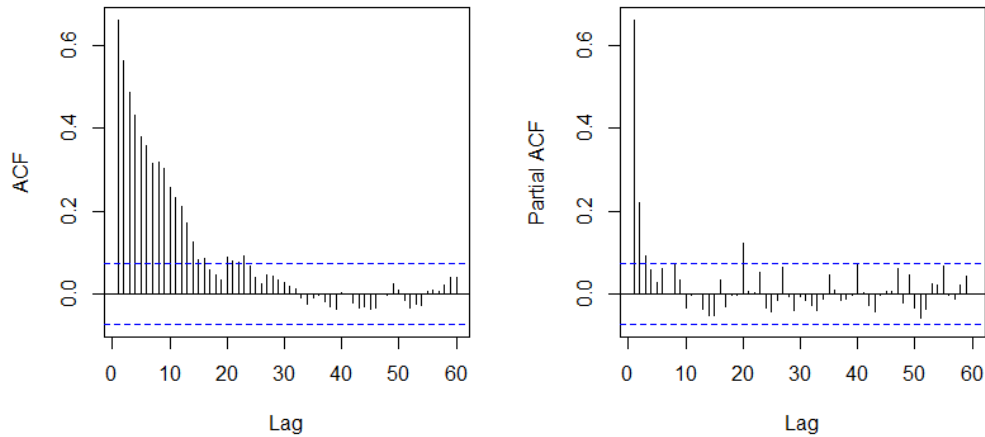
### 4.2.1 Identifikasi Model ARIMA

Identifikasi model ARIMA dilakukan untuk mengidentifikasi proses yang sesuai dengan data kecepatan angin rata-rata. Sebelum menentukan orde model ARIMA yang sesuai, proses diuji terlebih dahulu apakah sudah stasioner dalam *mean* dan varians atau belum. Plot kecepatan angin rata-rata (Gambar 4.1) menunjukkan adanya proses yang tidak stasioner. Oleh karena itu, dilakukan transformasi Box-Cox untuk mengetahui apakah proses sudah stasioner dalam varians.



**Gambar 4.4** Transformasi Box-Cox Kecepatan Angin Rata-rata

Hasil transformasi Box-Cox dalam Gambar 4.4 menunjukkan bahwa nilai lambda dan *rounded value* sebesar 0,50 serta batas bawah dan batas atas masing-masing sebesar 0,34 dan 0,66 dengan tingkat kepercayaan 95%. Berdasarkan Tabel 2.1, data kecepatan angin rata-rata ditransformasi dengan transformasi akar pangkat 2 karena memiliki nilai lambda sebesar 0,5.



**Gambar 4.5** ACF dan PACF Kecepatan Angin Rata-rata

Stasioneritas dalam *mean* dapat diidentifikasi dengan menggunakan plot ACF dan PACF yang ditampilkan dalam Gambar 4.5. Plot ACF turun secara eksponensial dan plot PACF *cut-off* pada lag-3 dan 20. Secara visual, belum dapat dipastikan proses sudah stasioner dalam *mean* atau belum. Dengan demikian, dilakukan pengujian ADF dengan hipotesis sebagai berikut.

$H_0 : \delta = 0$  (data tidak stasioner atau non-stasioner)

$H_1 : \delta < 0$  (data stasioner)

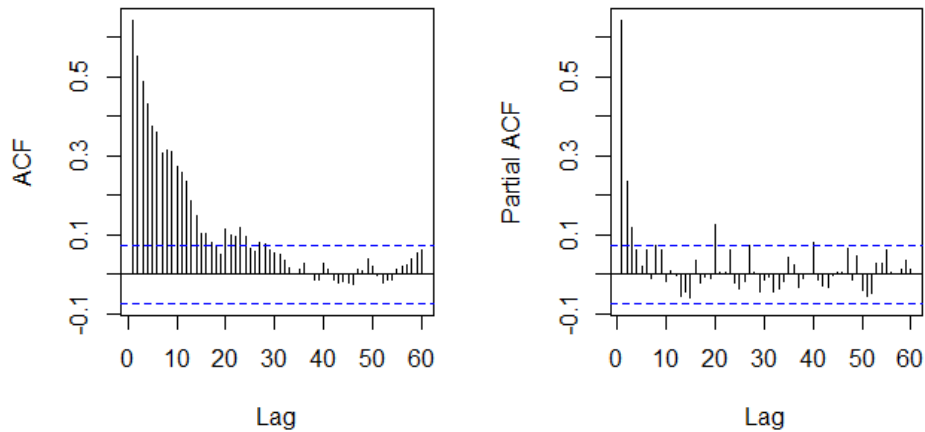
Uji ADF Tipe 2 yang digunakan dalam uji stasioneritas karena kecepatan angin rata-rata memiliki *drift* (amatan kecepatan angin rata-rata berada di atas nol) serta tidak memiliki tren. Nilai statistik uji ADF ( $\tau$ ) dihitung dengan menggunakan persamaan (2.6) dan didapatkan hasil dalam Tabel 4.1.

**Tabel 4.1** Uji ADF

Lag	$\tau$	P-value
1	-11,91	0,01
2	-8,68	0,01
3	-7,54	0,01
4	-6,79	0,01
5	-6,39	0,01
6	-5,78	0,01
7	-5,62	0,01

Nilai  $\tau$  lebih besar dari  $t$  dalam tabel Dickey Fuller serta  $P$ -value untuk setiap lag mulai dari lag ke-1 sampai dengan lag ke-7 sebesar 0,01 kurang dari 0,05 sehingga  $H_0$  ditolak. Dengan demikian, data kecepatan angin rata-rata stasioner dalam *mean*. Kecepatan angin berfluktuasi di sekitar nilai rata-rata sebesar 8,662 km/jam, tetapi variasi fluktuasinya tidak stasioner sehingga dilakukan transformasi agar data stasioner dalam varians.

Setelah menguji stasioneritas data dan melakukan transformasi, model ARIMA diidentifikasi dengan menggunakan plot ACF dan PACF dari data hasil transformasi. Plot ACF dan PACF data hasil transformasi dalam Gambar 4.6 hampir sama dengan plot dalam Gambar 4.5, yaitu plot ACF eksponensial turun dan PACF *cut-off*.



**Gambar 4.6** ACF dan PACF Data Transformasi

Identifikasi model ARIMA dilakukan dengan memperhatikan pola ACF dan PACF seperti dalam Tabel 2.2. Plot ACF eksponensial turun serta PACF data transformasi *cut-off* pada lag ke-1 sampai lag ke-3, lag ke-20, dan lag ke-40, artinya kecepatan angin rata-rata hari ini dipengaruhi kecepatan angin 3 hari sebelumnya serta hari ke-20 dan ke-40 sebelum pengamatan.

#### 4.2.2 Estimasi Parameter dan Uji Signifikansi

Berdasarkan plot ACF dan PACF dalam Gambar 4.6, dibentuk beberapa model dugaan/ tentatif ARIMA. Oleh karena pola PACF *cut-off* pada lag ke-3, 20, dan 40, tetapi lag ke-40 tidak dimasukkan dalam model karena ordenya sangat panjang sehingga model dugaan adalah AR(3) dan AR([3,20]). Selain itu, PACF lag orde ke-8 dan 9 meningkat setelah lag ke-7 sehingga dimodelkan dengan model AR([1,2,3,9]). Model dugaan lain adalah ARMA([3],3), ARMA([3,20],3),

dan ARMA([1,2,3,9],[19]). Parameter dari model tentatif diestimasi dengan menggunakan metode *Conditional Least Square* dan didapatkan hasil seperti dalam Tabel 4.2.

**Tabel 4.2** Estimasi Parameter Model ARIMA

Model	Parameter	Estimasi	SE	T	P
AR([1,2,3,9])	$\mu$	2,91616	0,07833	37,23	<0,0001
	$\phi_1$	0,45265	0,03771	12	<0,0001
	$\phi_2$	0,17227	0,04086	4,22	<0,0001
	$\phi_3$	0,09688	0,03821	2,54	0,0115
	$\phi_9$	0,08355	0,03043	2,75	0,0062
ARMA([1,2,3,9], [19])	$\mu$	2,91543	0,07477	38,99	<0,0001
	$\phi_1$	0,4551	0,03779	12,04	<0,0001
	$\phi_2$	0,17095	0,04092	4,18	<0,0001
	$\phi_3$	0,09794	0,03818	2,56	0,0105
	$\phi_9$	0,09693	0,03058	3,17	0,0016
	$\theta_{19}$	0,11981	0,03854	3,11	0,002
ARMA([3,20], 3)	$\mu$	2,91914	0,08039	36,31	<0,0001
	$\phi_3$	0,64366	0,06139	10,49	<0,0001
	$\phi_{20}$	0,06624	0,02961	2,24	0,0256
	$\theta_1$	-0,45074	0,03703	-12,17	<0,0001
	$\theta_2$	-0,37404	0,03767	-9,93	<0,0001
	$\theta_3$	0,28853	0,06417	4,5	<0,0001

Ketiga model ARIMA dalam Tabel 4.2 merupakan model dari proses yang stasioner dalam *mean* sehingga tidak dilakukan *di-differencing*. Hasil estimasi parameter model AR([1,2,3,9]), ARMA([1,2,3,9], [19]), dan ARMA([3,20], 3) diuji signifikansi parameter dengan menggunakan uji T dalam persamaan (2.35). Dengan tingkat signifikansi sebesar 0,05, semua parameter model memiliki *P-value* kurang dari 0,05 sehingga  $H_0$  ditolak. Jadi, disimpulkan bahwa semua parameter untuk setiap model ARIMA signifikan.

#### 4.2.3 Cek Diagnostik

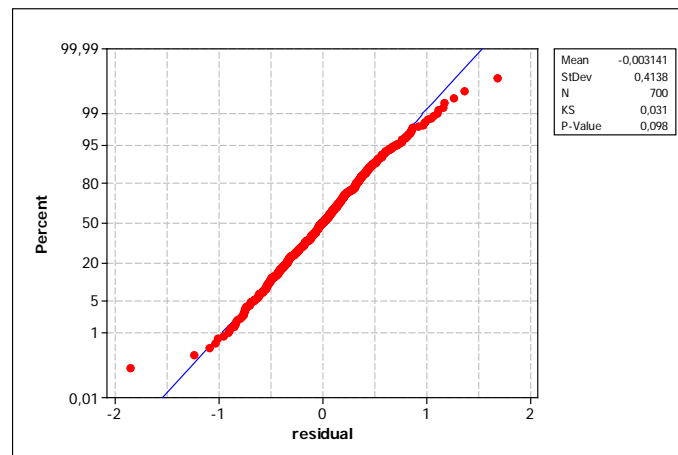
Model ARIMA yang memiliki parameter signifikan dilakukan tahap analisis ARIMA selanjutnya, yaitu melakukan cek diagnostik residual model ARIMA.

Model ARIMA merupakan model yang terikat oleh asumsi, yaitu asumsi residual *whitenoise* dan berdistribusi normal. Uji Ljung-Box digunakan untuk menguji residual *whitenoise* dengan persamaan (2.36). Uji normalitas residual model ARIMA menggunakan uji *Kolmogorov-Smirnov* (KS) dalam persamaan (2.37).

**Tabel 4.3** Uji Ljung-Box Residual Model AR([1,2,3,9])

Lag	Q	Df	P
6	3,75	2	0,1536
12	7,42	8	0,4919
18	13,32	14	0,5016
24	27,98	20	0,1099
30	32,24	26	0,1852
36	35,79	32	0,2951
42	44,15	38	0,2277
48	48,48	44	0,2971

Tabel 4.3 menunjukkan hasil uji Ljung-Box untuk residual model AR(3[9]), yaitu kumulatif lag ke-6 hingga lag ke-48 semua nilai P lebih dari 0,05 (tingkat signifikansi) sehingga  $H_0$  ditolak. Jadi, residual model AR([1,2,3,9]) sudah memenuhi asumsi *whitenoise*, yaitu antar residual identik dan independen dengan diasumsikan rata-rata nol dan variansi residual konstan.



**Gambar 4.7** Uji KS Residual Model AR([1,2,3,9])

Uji KS residual model AR([1,2,3,9]) menghasilkan nilai statistik uji KS sebesar 0,031 dengan *P-value* sebesar 0,098. Dengan tingkat signifikansi 0,05,  $H_0$  gagal ditolak karena *P-value* lebih dari 0,05 sehingga residual model AR([1,2,3,9]) memenuhi asumsi residual berdistribusi normal.

Residual model ARMA([1,2,3,9], [19]), dan ARMA([3,20], 3) diuji *white noise* dan hasilnya ditunjukkan dalam Tabel 4.4. *P-value* untuk kedua model

kurang dari 0,05 pada lag ke-6 hingga lag ke-48 sehingga residual model ARMA([1,2,3,9], [19]), dan ARMA([3,20], 3) telah memenuhi asumsi *whitenoise*.

**Tabel 4.4** Uji Ljung-Box Residual Model ARMA([1,2,3,9], [19]), dan ARMA([3,20], 3)

Model	Lag	Q	df	P
ARMA([1,2,3,9], [19])	6	3,51	1	0,061
	12	7,32	7	0,3963
	18	14,32	13	0,3518
	24	20,89	19	0,3428
	30	26,06	25	0,4046
	36	30,82	31	0,4751
	42	39,18	37	0,3724
	48	43,58	43	0,4465
ARMA([3,20], 3)	6	1,46	1	0,2276
	12	10,82	7	0,1466
	18	15,58	13	0,2726
	24	22,18	19	0,2754
	30	28,16	25	0,3004
	36	31,86	31	0,4236
	42	39,41	37	0,3624
	48	43,8	43	0,4374

Selanjutnya, residual model diuji apakah berdistribusi normal atau tidak dengan uji KS dan hasil pengujian KS disajikan dalam Tabel 4.5. Kedua model menghasilkan *P-value* sebesar  $>0,1500$  lebih besar daripada 0,05 sehingga  $H_0$  gagal ditolak. Jadi, residual model ARMA([1,2,3,9],[19]), dan ARMA([3,20],3) berdistribusi normal.

**Tabel 4.5** Uji KS Residual Model ARMA([1,2,3,9],[19]), dan ARMA([3,20], 3)

Model	D	P
ARMA([1,2,3,9], [19])	0,028907	$>0,1500$
ARMA([3,20], 3)	0,026269	$>0,1500$

#### 4.2.4 Evaluasi Model

Model dugaan yang telah dibentuk pada Subbab 4.2.2 telah memenuhi asumsi *whitenoise* dan berdistribusi normal sehingga model tersebut dapat digunakan untuk meramalkan kecepatan angin rata-rata pada periode berikutnya. Untuk mengevaluasi kebaikan ketiga model, digunakan kriteria evaluasi AIC dan BIC untuk data *in sample* serta kriteria MSE dan MAPE untuk data *out sample*.

Model AR([1,2,3,9]) memiliki nilai AIC terkecil dibandingkan dua model lain-nya seperti dalam Tabel 4.6, yaitu 760,2223. Kriteria SBC terkecil terdapat

pada model ARMA([1,2,3,9], [19]). Hasil evaluasi kebaikan model terbaik untuk data *out sample* adalah model ARMA([3,20], 3) karena nilai MSE dan MAPE terkecil dibandingkan model lainnya. Jadi, model ARMA([3,20], 3) dapat menghasilkan hasil ramalan yang lebih baik daripada dua model lainnya berdasarkan nilai MSE dan MAPE yang terkecil.

**Tabel 4.6** Evaluasi Model ARIMA

Model	In Sample		Out Sample	
	AIC	SBC	MSE	MAPE
AR([1,2,3,9])	760,2223	782,9777	0,167429	13,64813
ARMA([1,2,3,9], [19])	753,7115	781,018	0,182856	14,61766
ARMA([3,20], 3)	760,5595	787,866	0,166939	13,62945

Selain memperhatikan aspek kebaikan model, pemilihan model juga memperhatikan aspek kesederhanaan model (*parsimony*). Model AR([1,2,3,9]) lebih sederhana dan mudah diinterpretasikan, yaitu kecepatan angin saat ini dipengaruhi kecepatan angin 3 hari sebelumnya dan hari ke-9 sebelum hari ini. Selain itu, hasil evaluasi model untuk data *out sample* antara AR([1,2,3,9]) dan ARMA([3,20], 3) hampir sama. Jadi, model yang digunakan untuk meramalkan kecepatan angin rata-rata periode berikutnya adalah model AR([1,2,3,9]). Secara matematis, model AR([1,2,3,9]) dinotasikan dengan

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \phi_3(Z_{t-3} - \mu) + \phi_9(Z_{t-9} - \mu) + a_t$$

$$Z_t = \mu + \phi_1(Z_{t-1} - \mu) + \phi_2(Z_{t-2} - \mu) + \phi_3(Z_{t-3} - \mu) + \phi_9(Z_{t-9} - \mu) + a_t$$

sehingga kecepatan angin rata-rata di Juanda Surabaya dimodelkan dengan menggunakan metode ARIMA sebagai berikut,

$$Z_t = 2,91616 + 0,45265(Z_{t-1} - 2,91616) + 0,17227(Z_{t-2} - 2,91616) + 0,09688(Z_{t-3} - 2,91616) + 0,08355(Z_{t-9} - 2,91616) + a_t$$

atau disederhanakan menjadi

$$Z_t = 0,567634 + 0,45265Z_{t-1} + 0,17227Z_{t-2} + 0,09688Z_{t-3} + 0,08355Z_{t-9} + a_t.$$

Konstanta model AR([1,2,3,9]) sebesar 0,567634. Nilai koefisien lag  $Z_t$  semakin menurun seiring bertambah besarnya lag  $Z_t$  terhadap pengamatan sehingga korelasi atau pengaruh kecepatan angin rata-rata sehari sebelumnya terhadap pengamatan kecepatan angin rata-rata hari ini lebih besar daripada kecepatan angin 2 hari sebelumnya serta 3 dan 9 hari sebelumnya, yaitu masing-masing 0,45265; 0,17227; 0,09688; dan 0,08355.



#### 4.2.5 Peramalan Kecepatan Angin dengan Model ARIMA

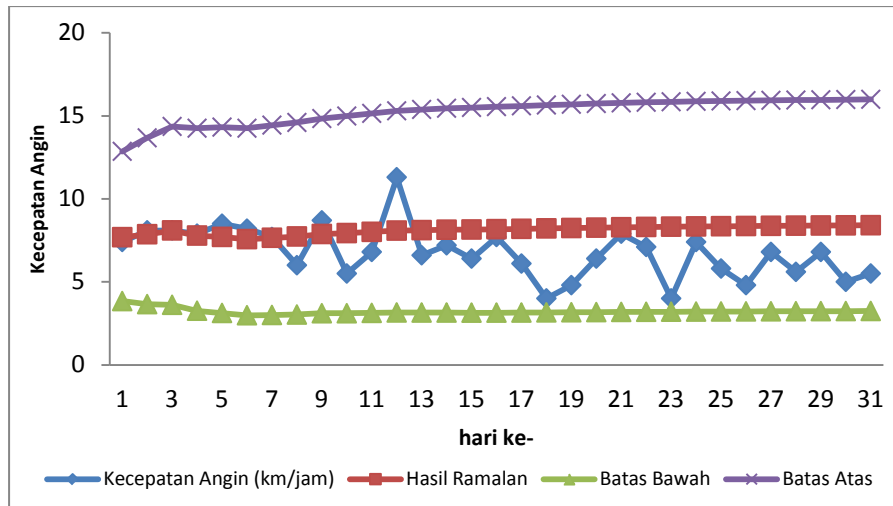
Kecepatan angin rata-rata di Juanda Surabaya selama satu bulan ke depan diramalkan dengan model AR([1,2,3,9]). Hasil ramalan model AR([1,2,3,9]) dan interval kepercayaan 95% untuk ramalan ditunjukkan dalam Tabel 4.7.

**Tabel 4.7** Hasil Ramalan Model AR([1,2,3,9])

<i>Outsample</i>	<b>Ramalan</b>	<b>SE</b>	<b>Batas Bawah</b>	<b>Batas Atas</b>
2,72029	2,77276	0,41500	1,95938	3,58614
2,84605	2,80489	0,45553	1,91206	3,69772
2,84605	2,84544	0,48167	1,90138	3,78951
2,81069	2,79049	0,50257	1,80547	3,77552
2,91548	2,77572	0,51449	1,76735	3,78409
2,86356	2,75167	0,52225	1,72809	3,77526
2,77489	2,76635	0,52739	1,73269	3,80001
2,44949	2,78063	0,53072	1,74044	3,82082
2,94958	2,80721	0,53290	1,76274	3,85169
2,34521	2,81700	0,53800	1,76255	3,87146
2,60768	2,83008	0,54169	1,76838	3,89178
3,36155	2,84365	0,54490	1,77566	3,91164
2,56905	2,84840	0,54785	1,77463	3,92217
2,68328	2,85292	0,55024	1,77448	3,93137
2,52982	2,85510	0,55217	1,77286	3,93733
2,77489	2,85854	0,55372	1,77328	3,94381
2,46982	2,86211	0,55492	1,77448	3,94974
2,00000	2,86675	0,55586	1,77729	3,95621
2,19089	2,87062	0,55672	1,77946	3,96177
2,52982	2,87460	0,55745	1,78203	3,96718
2,81069	2,87866	0,55807	1,78487	3,97245
2,66458	2,88195	0,55861	1,78711	3,97680
2,00000	2,88491	0,55906	1,78916	3,98065
2,72029	2,88738	0,55945	1,79087	3,98389
2,40832	2,88962	0,55978	1,79247	3,98678
2,19089	2,89165	0,56006	1,79395	3,98934
2,60768	2,89357	0,56029	1,79544	3,99171
2,36643	2,89534	0,56048	1,79682	3,99385
2,60768	2,89700	0,56064	1,79816	3,99583
2,23607	2,89858	0,56078	1,79947	3,99768
2,34521	2,90002	0,56089	1,80069	3,99935

Hasil ramalan kecepatan angin rata-rata dalam Tabel 4.7 cenderung semakin naik mulai dari periode ke-10 sehingga metode ARIMA baik digunakan untuk jangka pendek (*short term*). Ramalan dari model ARIMA dalam waktu jangka panjang

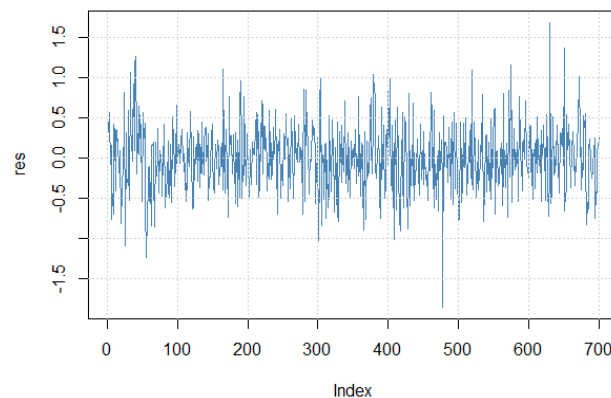
cenderung hampir sama dengan ramalan periode hari sebelumnya sehingga cenderung kurang representatif. Pada tahap identifikasi, data mengalami transformasi akar pangkat dua sehingga hasil ramalan sebaiknya dikembalikan dalam bentuk data awal dengan menghitung kuadrat data ramalan dalam Tabel 4.7 sehingga didapatkan grafik data *out sample* kecepatan angin rata-rata dengan ramalannya dalam Gambar 4.8.



**Gambar 4.8** Grafik Ramalan Model ARIMA dan Data *Outsample*

### 4.3 Model GARCH

Residual model kecepatan angin rata-rata yang ditunjukkan Gambar 4.9 terlihat memiliki volatilitas yang cukup besar sehingga diduga terjadi kasus heteroskedastisitas. Dengan demikian, kecepatan angin rata-rata dimodelkan dengan model ARIMA-GARCH. Model ARIMA digunakan untuk memodelkan *conditional mean* dan GARCH untuk memodelkan *conditional variance* sehingga dapat merepresentasikan fenomena sebenarnya dan meningkatkan kebaikan ramalan.



**Gambar 4.9** Plot Residual Model AR([1,2,3,9])

### 4.3.1 Uji Lagrange Multiplier

Uji *Lagrange Multiplier* (LM) digunakan untuk menguji adanya efek ARCH atau heteroskedastisitas. Residual dari model AR([1,2,3,9]) diuji dengan uji LM yang dihitung menggunakan persamaan (2.45) dan hasil uji LM disajikan dalam Tabel 4.8.

**Tabel 4.8** Hasil Uji LM

s	LM	$\chi^2_{0,05; s}$	P
1	5,2008	3,84	0,0226
2	6,1049	5,99	0,0472
3	6,7525	7,81	0,0802
4	7,8807	9,49	0,096
5	7,9496	11,07	0,159
6	8,075	12,59	0,2327
7	8,2413	14,07	0,3118
8	8,2418	15,51	0,4102
9	8,6164	16,92	0,4734
10	8,6292	18,31	0,5676
11	8,6487	19,68	0,6543
12	8,6887	21,03	0,7293

Nilai LM lebih besar dari  $\chi^2_{0,05; s}$  sampai lag kedua sehingga residual model AR([1,2,3,9]) memiliki efek ARCH atau variansi kondisional dari residual terjadi heteroskedastisitas (*conditional heteroscedasticity*). Dengan demikian, variansi residual dimodelkan dengan menggunakan model GARCH. Model GARCH yang digunakan dalam penelitian ini adalah model GARCH (1,1) dan GARCH(0,1) atau ARCH(1) karena orde LM yang signifikan hanya pada lag yang kecil. Selain itu, model GARCH memenuhi prinsip *parsimony*.

### 4.3.2 Estimasi Parameter Model GARCH

Variansi residual model AR([1,2,3,9]) dimodelkan dengan menggunakan GARCH(1,1) sehingga parameter yang diestimasi adalah parameter  $\omega_0$ ,  $\omega_1$ , dan  $\xi_1$ . Hasil estimasi parameter model AR([1,2,3,9])-GARCH(1,1) dalam Tabel 4.9 menunjukkan bahwa parameter  $\omega_0$  dan  $\omega_1$  signifikan pada tingkat signifikansi 5%, sedangkan parameter  $\xi_1$  tidak signifikan.

**Tabel 4.9** Estimasi Model GARCH(1,1)

Model	Parameter	Estimasi	SE	T	P
AR([1,2,3,9])	$\mu$	2,91616	0,07833	37,23	<0,0001
	$\phi_1$	0,45265	0,03771	12	<0,0001
	$\phi_2$	0,17227	0,04086	4,22	<0,0001
	$\phi_3$	0,09688	0,03821	2,54	0,0115
	$\phi_9$	0,08355	0,03043	2,75	0,0062
GARCH(1,1)	$\omega_0$	0,115	0,0474	2,43	0,0152
	$\omega_1$	0,1035	0,0506	2,05	0,0408
	$\xi_1$	0,225	0,28	0,79	0,4306

Selanjutnya, variansi residual dimodelkan tanpa menggunakan parameter  $\xi_1$  atau dengan kata lain model yang dibentuk adalah model ARCH(1). Estimasi model AR([1,2,3,9])-ARCH(1) ditampilkan dalam Tabel 4.10. Parameter model ARCH(1) signifikan pada tingkat signifikansi 5% sehingga model yang digunakan untuk meramalkan kecepatan angin rata-rata di Juanda Surabaya adalah model AR([1,2,3,9])-ARCH(1).

**Tabel 4.10** Estimasi Model ARCH(1)

Model	Parameter	Estimasi	SE	T	P
AR([1,2,3,9])	$\mu$	2,91616	0,07833	37,23	<0,0001
	$\phi_1$	0,45265	0,03771	12	<0,0001
	$\phi_2$	0,17227	0,04086	4,22	<0,0001
	$\phi_3$	0,09688	0,03821	2,54	0,0115
	$\phi_9$	0,08355	0,03043	2,75	0,0062
ARCH(1)	$\omega_0$	0,1545	0,008801	17,56	<0,0001
	$\omega_1$	0,097	0,0481	2,02	0,0438

Secara matematis, model AR([1,2,3,9])-ARCH(1) yang digunakan untuk meramalkan kecepatan angin rata-rata di Juanda Surabaya adalah

$$Z_t = 2,91616 + 0,45265(Z_{t-1} - 2,91616) + 0,17227(Z_{t-2} - 2,91616) + 0,09688(Z_{t-3} - 2,91616) + 0,08355(Z_{t-9} - 2,91616) + a_t$$

atau

$$Z_t = 0,567634 + 0,45265Z_{t-1} + 0,17227Z_{t-2} + 0,09688Z_{t-3} + 0,08355Z_{t-9} + a_t,$$

*conditional variance* dari residual  $a_t$  dimodelkan dengan

$$v_t^2 = 0,1545 + 0,097\varepsilon_{t-1}^2.$$

### 4.3.3 Evaluasi Keباikan Model GARCH

Model AR([1,2,3,9])-ARCH(1) dievaluasi kinerjanya dengan menggunakan kriteria AIC, SBC, dan MSE seperti yang disajikan dalam Tabel 4.11. Hasil evaluasi kebaikan model AR([1,2,3,9])-ARCH(1) lebih baik daripada nilai kriteria kebaikan model yang didapatkan dari model AR([1,2,3,9]) karena nilai AIC, SBC, dan MSE lebih kecil. Dengan demikian, pemodelan variansi residual AR(3[9]) yang mengalami heteroskedastis dengan ARCH(1) dapat meningkatkan kebaikan model dalam meramalkan kecepatan angin rata-rata di Juanda Surabaya.

**Tabel 4.11** Evaluasi Model AR([1,2,3,9])-ARCH(1)

Kriteria	Nilai	Kriteria	Nilai
AIC	747,905446	MSE	0,04479
SBC	757,007607	RMSE	0,2116364

### 4.3.4 Peramalan Variansi Residual

Variansi residual AR([1,2,3,9]) dimodelkan dengan ARCH(1) dan didapatkan model  $\hat{v}_t^2 = 0,1545 + 0,097\varepsilon_{t-1}^2$ . Variansi residual pada  $t = 1, 2, \dots, n$  dihitung menggunakan persamaan  $\hat{v}_t^2 = 0,1545 + 0,097\varepsilon_{t-1}^2$  dengan  $\hat{v}_t^2$  estimasi *conditional variance* residual dan  $\varepsilon_{t-1}^2$  residual AR([1,2,3,9]) yang *white noise*.

Ramalan variansi residual selama sebulan ke depan disajikan dalam Tabel 4.12. Variansi residual ramalan dihitung dengan menggunakan persamaan yang bersifat rekursif. Variansi residual pada  $h$  periode selanjutnya dihitung dengan persamaan  $Var_{t+h}(\varepsilon_t) = E_{t+h}(\varepsilon_t^2) = \hat{v}_{t+h}^2$  sehingga didapatkan *conditional variance* ramalan pada  $t=701, 702, \dots, 731$ , sebagai berikut.

$$\begin{aligned}\hat{v}_{701}^2 &= Var_{700}(\varepsilon_{701}) = E_{700}(\varepsilon_{701}^2) = E_{700}(0,1545 + 0,097\varepsilon_{700}^2) \\ &= 0,1545 + E_{700}(0,097\varepsilon_{700}^2) = 0,1545 + 0,097\varepsilon_{700}^2\end{aligned}$$

$$\begin{aligned}\hat{v}_{702}^2 &= Var_{700}(\varepsilon_{702}) = E_{700}(\varepsilon_{702}^2) = E_{700}(0,1545 + 0,097\varepsilon_{701}^2) \\ &= 0,1545 + E_{700}(0,097\varepsilon_{701}^2) = 0,1545 + 0,097E_{700}(\varepsilon_{701}^2) \\ &= 0,1545 + 0,097(0,1545 + 0,097\varepsilon_{700}^2) = 0,1545 + 0,097(0,1545) + 0,097^2\varepsilon_{700}^2\end{aligned}$$

$$\begin{aligned}\hat{v}_{703}^2 &= Var_{700}(\varepsilon_{703}) = E_{700}(\varepsilon_{703}^2) = E_{700}(0,1545 + 0,097\varepsilon_{702}^2) \\ &= 0,1545 + E_{700}(0,097\varepsilon_{702}^2) = 0,1545 + 0,097(0,1545 + 0,097(0,1545) + 0,097^2\varepsilon_{700}^2) \\ &= 0,1545 + 0,097(0,1545) + 0,097^2(0,1545) + 0,097^3\varepsilon_{700}^2\end{aligned}$$

dan seterusnya sampai  $\hat{v}_{731}^2$ .

Secara umum, ramalan *conditional variance* dihitung dengan persamaan berikut

$$\hat{v}_{t+h}^2 = \left( \sum_{h=1}^t \omega_1^{h-1} \right) \omega_0 + \omega_1^h \varepsilon_t^2$$

dengan  $t = 1, 2, \dots, n$  pengamatan *in-sample* dan  $h$  periode ke depan. Tabel 4.12 merupakan hasil ramalan variansi residual untuk 31 hari ke depan. Pada periode ramalan hari ke-9 sampai dengan ke-31, nilai ramalan variansi residual sama.

**Tabel 4.12** Ramalan variansi residual

$h$	$\hat{v}_h^2$	$h$	$\hat{v}_h^2$
1	0,160368155	17	0,171111017
2	0,170068902	18	0,171111017
3	0,171009927	19	0,171111017
4	0,171101211	20	0,171111017
5	0,171110066	21	0,171111017
6	0,171110925	22	0,171111017
7	0,171111008	23	0,171111017
8	0,171111016	24	0,171111017
9	0,171111017	25	0,171111017
10	0,171111017	26	0,171111017
11	0,171111017	27	0,171111017
12	0,171111017	28	0,171111017
13	0,171111017	29	0,171111017
14	0,171111017	30	0,171111017
15	0,171111017	31	0,171111017
16	0,171111017		

#### 4.4 Model GARCH-GA

Orde GARCH yang digunakan dalam meramalkan *conditional variance* sudah didapatkan dan dijelaskan dalam Subbab 4.3, yaitu  $r = 0$  dan  $s = 1$  atau terbentuk model ARCH(1). Parameter model ARCH(1), yaitu  $\omega_0$  dan  $\omega_1$  dioptimasi menggunakan *genetic algorithm* (GA) agar didapatkan estimasi yang optimum global. Tahapan pertama dalam estimasi dengan GA adalah membentuk kromosom awal yang terdiri dari 2 gen, yaitu  $\omega_0$  dan  $\omega_1$  sebagai solusi dari algoritma. Kromosom yang berisi hasil estimasi model ARCH(1) digunakan sebagai inisialisasi salah satu kromosom pada generasi pertama yang digunakan untuk megestimasi parameter ARCH(1). Kromosom akan mengalami tahap seleksi, *crossover* atau pindah silang, mutasi, dan elitisme dengan kromosom-

kromosom lainnya. Kromosom yang dibangkitkan pada setiap generasi sebanyak 100 kromosom seperti dalam Gambar 4.11.

	$\omega_0$	$\omega_1$
Kromosom 1	0,1545	0,097
Kromosom 2	0,8081	0,7716
	⋮	
Kromosom 100	1,1995	0,8955

**Gambar 4.10** Ilustrasi Solusi Metode GA

Fungsi *fitness* yang digunakan dalam mengevaluasi kromosom pada setiap populasi adalah  $\frac{1}{RMSE}$  seperti dalam persamaan (2.53). Nilai *fitness* dihitung

untuk setiap kromosom pada setiap iterasi (generasi) sebagai kriteria dalam melakukan seleksi, pindah silang, mutasi, dan elitisme. Dalam penelitian ini, fungsi tujuan dalam algoritma untuk mengestimasi parameter ARCH(1) adalah mendapatkan nilai RMSE *out-sample* minimum sehingga kromosom yang lebih baik adalah kromosom yang memiliki nilai *fitness* lebih besar.

Setelah mendapatkan nilai *fitness* untuk semua kromosom, kromosom diseleksi dengan menggunakan seleksi *Rhoulette Whell* untuk menentukan orang tua/ induk pada generasi selanjutnya. Dalam metode seleksi *Rhoulette Whell*, setiap kromosom dihitung proporsi dari nilai *fitness* dan nilai *fitness* kumulatif serta dibangkitkan bilangan random berdistribusi Uniform (0,1). Nilai *fitness* kumulatif dibandingkan dengan bilangan random uniform. Jika nilai *fitness* kumulatif pada kromosom ke-*i* lebih besar daripada nilai bilangan random ke-*i*, kromosom ke-*i* terpilih sebagai orang tua pada generasi selanjutnya.

**Tabel 4.13** Ilustrasi Seleksi *Rhoulette Whell*

Kromosom	Nilai <i>Fitness</i>	Proporsi Nilai <i>Fitness</i>	Nilai <i>Fitness</i> Kumulatif	Bilangan Random U(0,1)
1	0,2116	0,2116	0,2116	0,5229
2	3,0524	0,1564	0,368	0,3476
...	...	...	...	...
100	8,8683	0,4955	1	0,6364

Nilai bilangan random kromosom pertama dalam Tabel 4.13 lebih besar daripada nilai *fitness* kumulatif pertama sehingga kromosom pertama berubah menjadi indi-

vidu baru. Proses seleksi terus berlangsung hingga semua kromosom terevaluasi sehingga didapatkan individu baru pada generasi berikutnya.

Tahapan berikutnya adalah melakukan pindah silang untuk menghasilkan kromosom baru dari sepasang kromosom orang tua/ induk. Probabilitas pindah silang yang digunakan dalam penelitian ini adalah 0,8. Oleh karena kromosom terdiri dari 2 gen, pindah silang dilakukan dengan pindah silang satu titik (*single point*). Selanjutnya, membangkitkan bilangan random Uniform(0,1) sebanyak ukuran populasi. Bilangan random dibandingkan dengan probabilitas *crossover* atau pindah silang (0,8). Jika bilangan random lebih kecil daripada probabilitas pindah silang, kromosom induk/ orang tua tersebut disilangkan (dikawinkan) sehingga didapatkan ketu-runan atau individu baru dari perkawinan tersebut.

0,1545	0,097	Induk 1
0,8081	0,7716	Induk 2
0,1545	0,7716	Anak 1
0,8081	0,097	Anak 2

**Gambar 4.11** Ilustrasi Kromosom ARCH(1) yang Melakukan Pindah Silang

Induk 1 dan induk 2 dalam Gambar 4.11 melakukan pindah silang dengan memindahkan nilai estimasi gen kedua induk 1 dengan gen kedua induk 2 sehingga dihasilkan 2 individu baru.

Proses selanjutnya adalah proses mutasi dengan cara mengganti nilai dalam gen pada sebuah kromosom dengan nilai random lain sehingga didapatkan keberagaman gen untuk menghindari terjadinya berhentinya algoritma pada solusi lokal. Sama seperti proses pindah silang, bilangan random dibangkitkan sebanyak ukuran sampel, tetapi setiap gen memiliki bilangan random. Nilai bilangan random dalam gen dibandingkan dengan probabilitas mutasi (0,1). Jika nilai bilangan random kurang dari probabilitas mutasi, nilai estimasi dalam gen tersebut diganti dengan nilai random lainnya seperti dalam Gambar 4.12.

Sebelum mutasi	Kromosom	0,8081	0,7716
	Bilangan Random	0,0473	0,8455
Setelah mutasi	Kromosom	0,5851	0,7716
	Bilangan Random	0,8438	0,8455

**Gambar 4.12** Ilustrasi Proses Mutasi

Proses elitisme merupakan proses terakhir pada satu generasi (iterasi) untuk mengestimasi parameter. Elitisme bertujuan untuk mempertahankan kromosom



yang memiliki nilai *fitness* terbaik agar kromosom tetap bertahan/ berada pada generasi berikutnya, yaitu sebanyak 5 kromosom. Jadi, setiap generasi mempertahankan 5 kromosom dengan nilai *fitness* terbaik untuk bertahan/ berada pada generasi selanjutnya. Iterasi dilakukan sebanyak 1000 kali hingga didapatkan kromosom terbaik pada generasi terakhir untuk estimasi parameter ARCH(1) ditunjukkan dalam Tabel 4.14.

**Tabel 4.14** Solusi Estimasi Parameter ARCH(1) dengan GA

Parameter ARCH(1)	Estimasi	1/RMSE	RMSE
$\omega_0$	0,028777	5,046601	0,198153
$\omega_1$	0,879281		

Nilai RMSE model ARCH-GA lebih kecil daripada nilai RMSE model ARCH dalam Tabel 4.11 sehingga estimasi dengan GA menghasilkan ramalan lebih baik daripada tanpa menggunakan GA. Secara matematis, model AR([1,2,3,9]) ARCH(1)-GA adalah.

$$Z_t = 0,567634 + 0,45265Z_{t-1} + 0,17227Z_{t-2} + 0,09688Z_{t-3} + 0,08355Z_{t-9} + a_t$$

dengan  $v_t^2 = 0,028777 + 0,879281\varepsilon_{t-1}^2$ .

#### 4.5 Studi Simulasi

Studi simulasi dilakukan untuk mengetahui kemampuan model GARCH dengan optimasi GA dalam mengakomodasi kasus heteroskedastisitas serta optimasi parameter GARCH sehingga didapatkan hasil ramalan yang baik dan akurat. Dalam penelitian ini, data dibangkitkan mengikuti proses GARCH(1,1) untuk skenario 1 dan 2 dengan parameter model yang ditentukan seperti dalam Tabel 4.15 dan direplikasi sebanyak 10 kali. Model GARCH orde 1 umum dan seringkali digunakan dalam peramalan karena modelnya yang sederhana. Data dibangkitkan dengan distribusi normal. Karakteristik yang dapat menggambarkan sebaran data adalah kurtosis dan *skewness*.

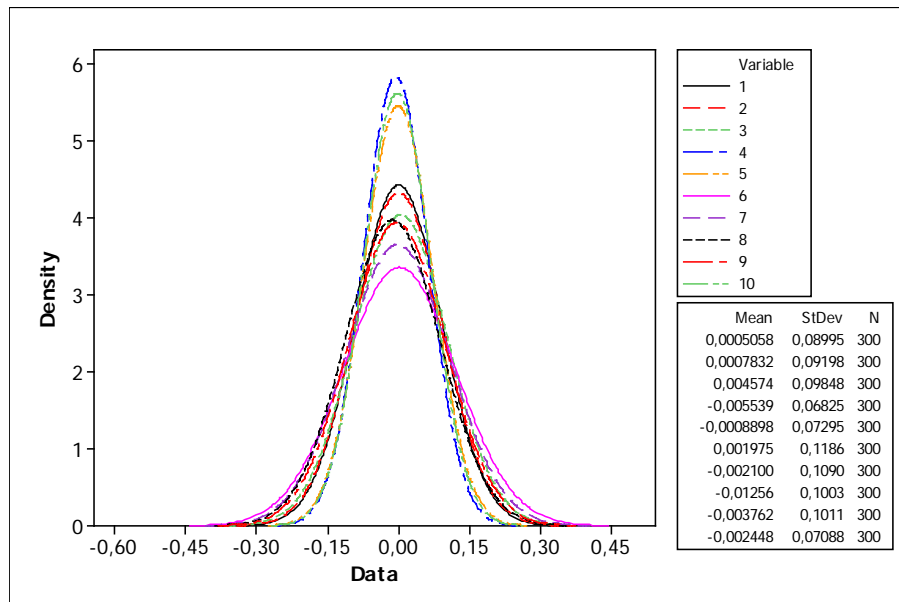
**Tabel 4.15** Skenario Simulasi

Parameter	Skenario 1	Skenario 2
$\omega_0$	0,00054	0,000011
$\omega_1$	0,25	0,09
$\xi_1$	0,69	0,9

Dalam melakukan simulasi, data dibangkitkan sesuai parameter yang ditentukan. Data yang digunakan untuk analisis studi simulasi adalah data yang memiliki nilai estimasi yang mendekati dengan parameter yang ditentukan atau berada dalam batas interval kepercayaan 95%. Selain itu, nilai estimasi parameter yang mendekati parameter bangkitan tersebut juga signifikan pada tingkat signifikansi 5%. Proses membangkitkan data tersebut diulang hingga mendapatkan 10 data bangkitan.

#### 4.5.1 Data Simulasi Skenario 1

Data yang mengikuti model GARCH(1,1) dengan parameter skenario 1 dibangkitkan sebanyak 300 dan 1000 observasi. Sebaran data simulasi model GARCH(1,1) skenario 1 dengan ukuran sampel 300 ditampilkan dalam bentuk histogram seperti dalam Gambar 4.13.



Gambar 4.13 Skenario 1 untuk  $n = 300$

Data simulasi skenario 1 untuk ukuran sampel 300 observasi memiliki kurtosis yang positif sehingga mengindikasikan sebaran data *fat-tailed*. Kurtosis terbesar terdapat di data replikasi ke-6 (5,95) dengan *skewness* (kemiringan) sebesar -1,02 yang ditunjukkan dalam Tabel 4.16 sehingga sebarannya miring ke kiri (*left-skewed*). Data replikasi ke-4 memiliki kurtosis terkecil sebesar 0,24 sehingga bagian ekor atau *tailed* dari distribusi datanya lebih mendekati distribusi normal, sedangkan data replikasi ke-9 memiliki *skewness* paling mendekati nol sebesar 0,05 sehingga data replikasi ke-9 lebih simetris.

**Tabel 4.16** *Skewness* dan *Kurtosis* Data Simulasi Skenario 1 untuk  $n = 300$ 

Replikasi	<i>Skewness</i>	<i>Kurtosis</i>
1	0,06	2,60
2	-0,14	1,32
3	0,36	2,64
4	0,27	0,24
5	-0,21	0,41
6	-1,02	5,95
7	0,28	2,22
8	0,10	2,47
9	0,05	1,98
10	-0,15	1,11

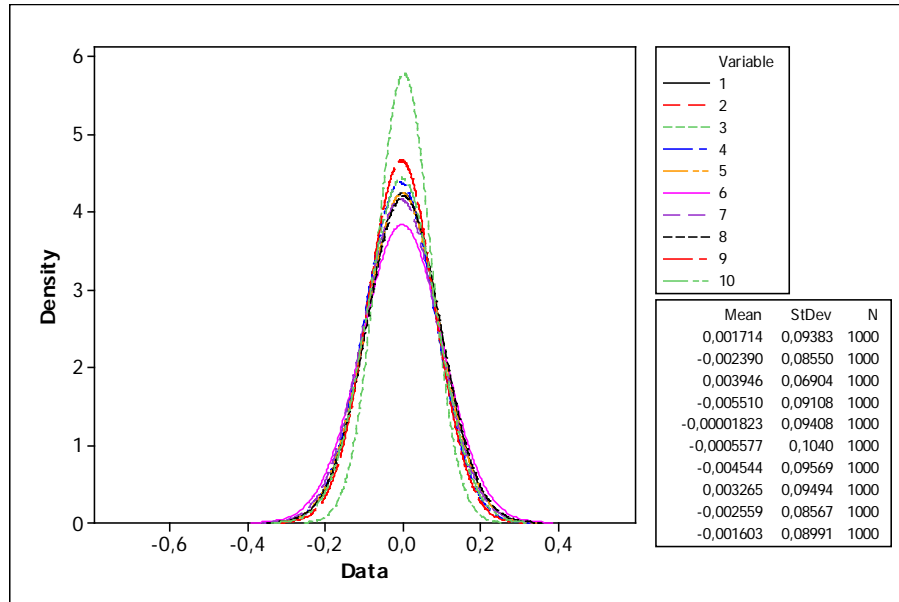
Sepuluh data replikasi tersebut selanjutnya dianalisis dengan menggunakan GARCH-GA. Data dibagi menjadi data *in sample* (270) dan *out sample* (30). Data *in sample* digunakan untuk membentuk model dan didapatkan estimasi parameternya dalam Tabel 4.17, sedangkan data *out sample* digunakan untuk meramalkan periode ke depan dan mengevaluasi kebaikan model.

**Tabel 4.17** Hasil Estimasi dan RMSE GARCH(1,1)-GA Skenario 1 untuk  $n = 300$ 

Replikasi	$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\xi}_1$	RMSE GARCH-GA
1	0,002703	0,137307	0,090852	0,003718
2	0,004844	0,186965	0,097197	0,007284
3	0,003278	0,046269	0,071336	0,003646
4	0,001544	0,109294	0,468806	0,004578
5	0,004594	0,023129	0,047193	0,004980
6	0,005297	0,003750	0,010372	0,009010
7	0,001644	0,116281	0,515635	0,005364
8	0,003381	0,020132	0,006683	0,003594
9	0,007276	0,351897	0,376573	0,018230
10	0,003263	0,146711	0,231266	0,005797
Rata-rata	0,003782	0,114173	0,191591	0,006620
Deviasi Standar	0,001755	0,104014	0,194092	0,004428

Data *in sample* setiap replikasi dimodelkan dengan GARCH(1,1) dan didapatkan estimasi parameter dengan optimasi GA seperti dalam Tabel 4.17 dengan rata-rata untuk estimasi parameter  $\omega_0$ ,  $\omega_1$ , dan  $\xi_1$  masing-masing adalah 0,003782; 0,114173; dan 0,191591. Deviasi standar dari estimasi parameter hampir sama dengan nilai rata-rata untuk semua replikasi. Hal ini disebabkan karena hasil estimasi yang bervariasi untuk setiap replikasi. Meskipun demikian, hasil estimasi tersebut merupakan solusi yang diperoleh dari algoritma genetika, yaitu

solusi global optimum. Rata-rata nilai RMSE yang didapatkan adalah 0,006620 dengan RMSE terkecil terdapat dalam replikasi 8, sedangkan RMSE terbesar ada di replikasi ke-9.



**Gambar 4.14** Skenario 1 untuk  $n = 1000$

Data simulasi skenario 1 dibangkitkan sebanyak 1000 observasi dengan sebaran data ditampilkan dalam Gambar 4.14. Data 10 replikasi cenderung memiliki sebaran yang hampir sama dengan deviasi standar data sekitar 0,09 dengan *mean* mendekati nol. Menurut Gambar 4.14, replikasi ke-3 memiliki puncak yang paling tinggi dan kurtosis terbesar (7,32) serta deviasi standar terkecil, yaitu 0,06904 sehingga data replikasi ke-3 paling homogen dibandingkan yang lainnya. Replikasi ke-3 memiliki sebaran data yang paling simetris karena nilai *skewness* sebesar 0,00 yang ditunjukkan Tabel 4.18.

**Tabel 4.18** *Skewness* dan Kurtosis Data Simulasi Skenario 1 untuk  $n = 1000$

Replikasi	<i>Skewness</i>	Kurtosis
1	0,26	2,02
2	-0,00	1,01
3	0,00	3,79
4	-0,66	7,32
5	0,32	3,34
6	-0,04	4,77
7	0,07	1,50
8	-0,06	2,46
9	0,03	1,93
10	-0,45	6,94

Sepuluh data replikasi tersebut selanjutnya dianalisis dengan menggunakan GARCH-GA. Data dibagi menjadi data *in sample* (970) dan *out sample* (30). Data *in sample* digunakan untuk membentuk model, sedangkan data *out sample* digunakan untuk meramalkan dan mengevaluasi kebaikan model dalam meramalkan pada periode ke depan.

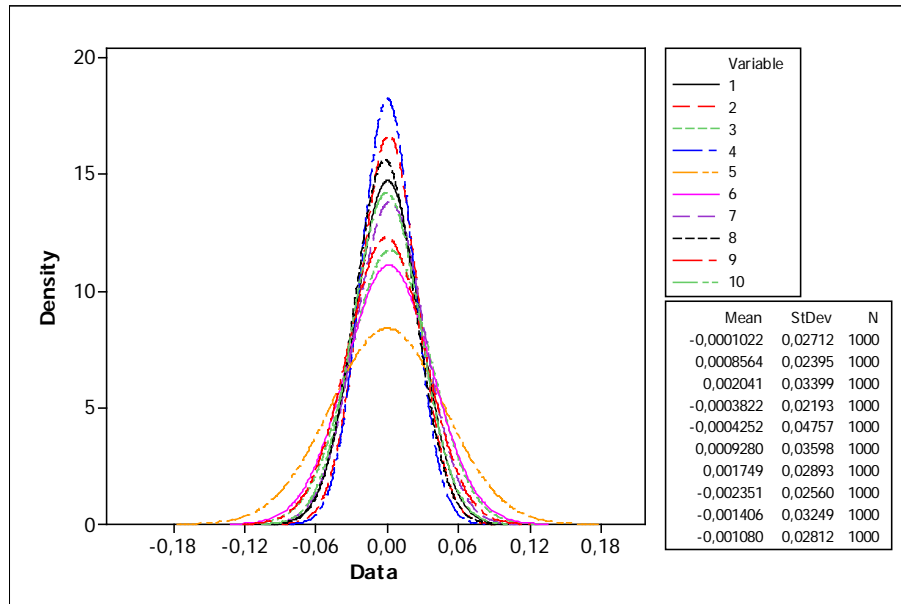
**Tabel 4.19** Hasil Estimasi dan RMSE GARCH(1,1)-GA Skenario 1 untuk  $n = 1000$

Replikasi	$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\xi}_1$	RMSE GARCH-GA
1	0,004105	0,022538	0,017227	0,006345
2	0,009978	0,040480	0,085433	0,019483
3	0,002794	0,003286	0,005965	0,003361
4	0,004696	0,283912	0,212983	0,011578
5	0,002205	0,007872	0,003669	0,002986
6	0,004960	0,097707	0,091090	0,008362
7	0,005178	0,017260	0,023851	0,005937
8	0,003617	0,019083	0,009312	0,003907
9	0,005652	0,224346	0,144837	0,008670
10	0,004580	0,135994	0,113066	0,006268
Rata-rata	0,004777	0,085248	0,070743	0,007690
Deviasi Standar	0,002119	0,099548	0,071239	0,004918

Berdasarkan Tabel 4.19, rata-rata untuk estimasi parameter  $\omega_0$ ,  $\omega_1$ , dan  $\xi_1$  masing-masing adalah 0,004777; 0,085248; dan 0,070743. Deviasi standar dari estimasi parameter juga hampir sama dengan nilai rata-rata untuk semua replikasi seperti dalam Tabel 4.17. Meskipun demikian, hasil estimasi tersebut merupakan solusi yang diperoleh dari algoritma genetika yang merupakan solusi global optimum. Rata-rata nilai RMSE model GARCH(1,1)-GA skenario 1 untuk ukuran sampel 1000 observasi adalah 0,007690 dengan RMSE terkecil terdapat dalam replikasi ke-5, sedangkan RMSE terbesar ada di replikasi ke-2, yaitu masing-masing sebesar 0,002986 dan 0,019483 yang terdapat dalam Tabel 4.19.

#### 4.5.2 Data Simulasi Skenario 2

Model GARCH(1,1) dengan parameter seperti dalam skenario 2 dibangkitkan sebanyak 10 replikasi dengan ukuran sampel 1000 observasi memiliki sebaran data seperti dalam Gambar 4.15. Data yang digunakan untuk memodelkan sebanyak 970 observasi dan sisanya digunakan sebagai data *out sample*. Sebaran data hasil simulasi lebih bervariasi dibandingkan hasil simulasi pada Subbab 4.5.1 terlihat dari bentuk puncak, kemiringan, serta bagian ekor (*tailed*) yang beragam.



**Gambar 4.15** Skenario 2 untuk  $n = 1000$

Data 10 replikasi memiliki kurtosis positif dengan deviasi standar antara 0,02 sampai 0,04 dan *mean* mendekati nol. Menurut Gambar 4.15, replikasi ke-4 memiliki puncak tertinggi serta deviasi standar terkecil, yaitu 0,02193 seperti dalam Tabel 4.20 sehingga data replikasi ke-4 paling homogen dibandingkan lainnya. Replikasi ke-5 memiliki deviasi standar terbesar dan bentuk kurvaanya lebih lebar sehingga data replikasi ke-5 paling bervariasi.

**Tabel 4.20** *Skewness* dan *Kurtosis* Data Simulasi Skenario 2 untuk  $n = 1000$

Replikasi	<i>Skewness</i>	<i>Kurtosis</i>
1	-0,12	0,88
2	-0,06	0,71
3	0,17	1,85
4	0,02	0,29
5	-0,07	2,18
6	0,15	0,95
7	0,24	0,61
8	0,10	0,78
9	-0,31	2,76
10	0,03	0,87

Berdasarkan Tabel 4.21, rata-rata dan deviasi standar untuk estimasi parameter  $\omega_0$ ,  $\omega_1$ , dan  $\xi_1$  lebih kecil dibandingkan hasil dalam Tabel 4.17 dan 4.19. Deviasi standar dari estimasi parameter hampir sama dengan nilai rata-rata untuk semua replikasi karena hasil estimasi bervariasi. Meskipun demikian, hasil estimasi tersebut merupakan solusi global optimum dari algoritma genetika.

**Tabel 4.21** Hasil Estimasi dan RMSE GARCH(1,1)-GA Skenario 2 untuk  $n = 1000$ 

Replikasi	$\hat{\omega}_0$	$\hat{\omega}_1$	$\hat{\xi}_1$	RMSE GARCH-GA
1	0,001119	0,032837	0,006387	0,002060
2	0,000346	0,002849	0,011536	0,001565
3	0,000519	0,092363	0,051909	0,001689
4	0,000342	0,008113	0,004792	0,001610
5	0,001164	0,005143	0,002775	0,001726
6	0,000359	0,013537	0,003281	0,001536
7	0,000338	0,009320	0,005410	0,001221
8	0,001067	0,006951	0,004550	0,001256
9	0,000329	0,051750	0,031401	0,001521
10	0,000788	0,007065	0,005342	0,001664
Rata-rata	0,000637	0,022993	0,012738	0,001585
Deviasi Standar	0,000360	0,028808	0,016165	0,000238

Rata-rata nilai RMSE model GARCH(1,1)-GA skenario 2 untuk ukuran sampel 1000 observasi adalah 0,001585 dengan RMSE terkecil terdapat dalam replikasi ke-7, sedangkan RMSE terbesar ada di replikasi ke-1 seperti yang ditunjukkan dalam Tabel 4.21. Hasil RMSE untuk data dengan nilai kurtosis yang lebih rendah cenderung lebih rendah daripada untuk data dengan kurtosis yang lebih besar.

*(Halaman ini sengaja dikosongkan)*



## **BAB 5**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Kesimpulan yang didapatkan dari penelitian adalah

1. Kecepatan angin rata-rata di Juanda Surabaya memiliki rata-rata sebesar 8,662 km/jam dengan variansi sebesar 11,482 km/jam serta memuat adanya data *outlier*.
2. Model ARIMA-GARCH yang terpilih untuk meramalkan kecepatan angin rata-rata di Juanda Surabaya adalah AR([1,2,3,9]) ARCH(1) dengan RMSE sebesar 0,2116. Parameter ARCH(1) diestimasi dengan GA dan didapatkan nilai RMSE yang lebih kecil, yaitu sebesar 0,198153 sehingga GA mampu meningkatkan kebaikan hasil ramalan.
3. Simulasi dari skenario yang dibuat menunjukkan bahwa GARCH-GA dapat menghasilkan RMSE yang lebih kecil.

#### **5.2 Saran**

Saran yang diberikan untuk penelitian selanjutnya adalah melakukan analisis data *outlier* sebelum memodelkan kecepatan angin rata-rata dengan model GARCH. Simulasi model dapat dibangun dari model ARMA-GARCH yang lainnya dengan skenario lainnya. Selain itu, studi simulasi dapat lebih komprehensif membahas distribusi asal data dibangkitkan dan kesesuaiannya dengan hasil simulasi yang didapatkan.

*(Halaman ini sengaja dikosongkan)*

## DAFTAR PUSTAKA

- Anityaloka, R.N. (2012). *Peramalan Kecepatan Angin Rata-Rata di Bandara Juanda Surabaya Menggunakan Metode ANFIS*, Tesis. Institut Teknologi Sepuluh Nopember, Surabaya.
- Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity, *Econometrics*, 31, 307-327.
- Chen, K. dan Yu, J. (2014). Short-term Wind Speed Prediction Using an Unscented Kalman Filter Based State-Space Support Vector Regression Approach, *Applied Energy*, 113, 690-705.
- Engelbrecht, A.P. (2002), *Computational Intelligence : An Introduction*, John Wiley and Sons.
- Engle, R.F. (1982). Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation, *Econometrica*, 55, 987-1007.
- Gao, S., He, Y., dan Chen, H. (2009). Wind Speed Forecast for Wind Farms Based on ARMA-ARCH Model, 1-4.
- Gen, M. dan Cheng, R. (2000), *Genetic Algorithm and Optimization Engineering*, John Wiley and Sons.
- Gujarati, D.N. (2004). *Basic Econometrics (4th ed.)*. The McGraw-Hill: New York.
- Holland, J.H. (1975). *Adaptive in Natural and Artificial Systems*, Univ. Mich. Press, Ann Arbor, MI.
- Hung, J.C. (2009). A Fuzzy GARCH Model Applied to Stock Market Scenario Using a Genetic Algorithm. *Expert System with Applications*, 36, 11710-11717.
- Hung, J.C. (2011). Applying a Combined Fuzzy Systems and GARCH Model to Adaptively Forecast Stock Market Volatility, *Applied Soft Computing*, 11, 3938-3945.
- Kavasseri, R.G. dan Seetharaman, K. (2009). Day-ahead Wind Speed Forecasting Using f-ARIMA Models, *Renewable Energy*, 34, 1388-1393.
- Kusumawati, D. (2011). *Peramalan Kecepatan Angin Rata-Rata Harian di Surabaya Menggunakan Metode Bayesian Model Averaging dengan*

*Pendekatan Expectation Maximization*, Skripsi. Institut Teknologi Sepuluh Nopember, Surabaya.

- Liu, D., Niu, D., Wang, H., dan Fan, L. (2014). Short-Term Wind Speed Forecasting Using Wavelet Transform and Support Vector Machines Optimized by Genetic Algorithm, *Renewable Energy*, 62, 592-597.
- Liu, H., Erdem, E., dan Shi, J. (2011). Comprehensive Evaluation of ARMA-GARCH(-M) Approaches for Modeling the Mean and Volatility of Wind Speed, *Applied Energy*, 88, 724-732.
- Permai, S.D., Iriawan, N., dan Irhamah. (2013). Pemodelan Kecepatan Angin Rata-Rata Di Sumenep Menggunakan Mixture of Anfis, *Jurnal Statistika*, 1(2).
- Souza, L., Veiga, A., dan Medeiros, M.C. (2005). Evaluating the Forecasting Performance of GARCH Models Using White's Reality Check, *Brazilian Review of Econometrics*, 25(1), 43-66.
- Torres, J.L., Garcia, A., Blas, M.D., dan Francisco, A.D. (2005). Forecast of Hourly Average Wind Speed with ARMA Models in Navarre (Spain), *Solar Energy*, 79, 65-77.
- Wahyuningtyas, R.B. (2015). *Prediksi Kecepatan Angin dan Ketinggian Gelombang pada Cuaca Maritim Menggunakan Metode ARIMA Box Jenkins-Filter Kalman*, Skripsi. Institut Teknologi Sepuluh Nopember, Surabaya.
- Wang, F., Liu, D.Y., Zeng, L.H., Chen, S.L., dan Zhou, L. (2009). Study on the Wind Speed Frequency Distribution with AR-GARCH Model, 1-5.
- Wei, W. W. (2006). *Time Series Analysis, Univariate and Multivariate Method*. New York: Greg Tobin.

## LAMPIRAN

**Lampiran 1.** Data Kecepatan Angin Rata-rata

Hari ke-	Kecepatan Angin Rata-rata	Transformasi Kecepatan Angin
1	10.5	3.240370349
2	13.2	3.633180425
3	14.2	3.768288736
4	11.4	3.376388603
5	12.2	3.492849839
6	6.6	2.569046516
7	6.0	2.449489743
8	4.0	2
9	7.9	2.810693865
10	9.2	3.033150178
11	6.3	2.50998008
12	10.0	3.16227766
13	11.3	3.361547263
14	10.5	3.240370349
15	11.4	3.376388603
⋮	⋮	⋮
721	7.1	2.664582519
723	4.0	2
724	7.4	2.720294102
725	5.8	2.408318916
726	4.8	2.19089023
727	6.8	2.607680962
728	5.6	2.366431913
729	6.8	2.607680962
730	5.0	2.236067977
731	5.5	2.34520788

## Lampiran 2. Uji ADF Dengan *Software R*

```
> adf.test(data1)
Augmented Dickey-Fuller Test
alternative: stationary

Type 1: no drift no trend
  lag   ADF p.value
[1,]  0 -4.04 0.0100
[2,]  1 -2.88 0.0100
[3,]  2 -2.47 0.0146
[4,]  3 -2.18 0.0297
[5,]  4 -2.02 0.0434
[6,]  5 -1.75 0.0803
[7,]  6 -1.63 0.0979
Type 2: with drift no trend
  lag   ADF p.value
[1,]  0 -11.91  0.01
[2,]  1  -8.68  0.01
[3,]  2  -7.54  0.01
[4,]  3  -6.79  0.01
[5,]  4  -6.39  0.01
[6,]  5  -5.78  0.01
[7,]  6  -5.62  0.01
Type 3: with drift and trend
  lag   ADF p.value
[1,]  0 -11.91  0.01
[2,]  1  -8.68  0.01
[3,]  2  -7.53  0.01
[4,]  3  -6.79  0.01
[5,]  4  -6.39  0.01
[6,]  5  -5.78  0.01
[7,]  6  -5.62  0.01
-----
Note: in fact, p.value = 0.01 means p.value <= 0.01
```

### Lampiran 3. *Syntax* Model AR([1,2,3,9]) dengan SAS

```
data angin;
input y;
datalines;
3.240370349
3.633180425
3.768288736
.
.
.
2.449489743
2.607680962
2.846049894
;
proc arima data=angin;
identify var=y nlag=48;
run;
estimate p=(1,2,3,9) method=cls;
run;

forecast out=frfst lead=31;
run;
proc univariate data=frfst normal;
var residual;
run;
proc export data=work.frfst
outfile = 'D:\outp1.xls'
dbms = excel
replace;
run;
```

#### Lampiran 4. *Syntax* Model AR([1,2,3,9])-GARCH(1,1) dengan SAS

```
data angin;
input y;
datalines;
3.240370349
3.633180425
3.768288736
.
.
.
2.449489743
2.607680962
2.846049894
;
data angin;
set angin;
run;
proc arima data= angin;
identify var=y;
estimate p=(1,2,3,9) method=cls;
forecast out=ramalan lead=31;
run;
proc print data=ramalan;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc autoreg data=ramalan;
model residual=/archtest noint;
model residual=/noint garch=(p=1,q=1);
output out=ramgarch cev=vhat;
run;
proc print data=ramgarch;
run;
```



**Lampiran 5.** *Syntax* Model AR([1,2,3,9])-ARCH(1,1) dengan SAS

```
data angin;
input y;
datalines;
3.240370349
3.633180425
3.768288736
.
.
.
2.449489743
2.607680962
2.846049894
;
data angin;
set angin;
run;
proc arima data= angin;
identify var=y;
estimate p=(1,2,3,9) method=cls;
forecast out=ramalan lead=31;
run;
proc print data=ramalan;
run;
proc univariate data=ramalan normal;
var residual;
run;
proc autoreg data=ramalan;
model residual=/archtest noint;
model residual=/noint garch=(q=1);
output out=ramgarch cev=vhat;
run;
proc print data=ramgarch;
run;
proc export data=work.ramgarch
outfile = 'D:\outgarch.xls'
dbms = excel
replace;
run;
```

## Lampiran 6. Syntax Estimasi ARCH(1)-GA Data Kecepatan Angin Rata-rata

```
rm(list=ls())
library(iterators)
library(foreach)
library(GA)
alldata=read.table("D://datatesis.csv",sep=";",header=TRUE)
mydata=as.matrix(alldata)
training=mydata[1:700,] #observasi Zt dan residual nt
testing=mydata[-c(1:700),]
nt=mydata[,2]

#Fungsi Fitness
#Parameter yang diestimasi
my_fitness=function(x){
  omega=x[1]
  alpha=x[2]
  #Model GARCH(1,0) atau ARCH(1) data training #notasi R GARCH(s,r)
  #conditional variance out sample/ forecast
  for (i in 701:731){
    k=length(testing[,2])
    koefal = function(alpha, k) {
      al = 0
      for(j in 1:k) al = al + alpha^(j-1)
      return(al)
    }
    cv=0
    sigma2=0
    for (j in 1:k){
      sigma2[700+j] = cv+(koefal(alpha, j))*omega + (alpha^j) * (nt[700]^2)
    }
  }

  #fungsi fitness untuk optimasi GA / fungsi tujuan
  E=(nt[701:731])^2-(sigma2[701:731])
  for (i in 701:731){
    RMSE=(mean(E^2))^(1/2)
    invRMSE=1/RMSE
  }
  return(invRMSE)
}

#Range nilai parameter
par.min=c(omega=1e-3, alpha=1e-3)
par.max=c(omega=2, alpha=2)

#Run genetic algorithm
set.seed(1234)
optim=ga(type="real-valued", fitness=my_fitness, min=par.min, names=names(par.min),
         max=par.max, popSize=100, maxiter=1000, pcrossover=0.8, pmutation=0.1)
my_result=summary(optim)
my_result
```

### Lampiran 7. Output ARCH(1)-GA Data Kecepatan Angin Rata-Rata

```
+-----+
|           Genetic Algorithm           |
+-----+

GA settings:
Type           = real-valued
Population size = 100
Number of generations = 1000
Elitism        = 5
Crossover probability = 0.8
Mutation probability = 0.1
Search domain =
  omega alpha
Min 0.001 0.001
Max 2.000 2.000

GA results:
Iterations           = 1000
Fitness function value = 5.046601
Solution =
  omega      alpha
[1,] 0.02877711 0.8792811
```

### Lampiran 8. Simulasi GARCH(1,1) Skenario 1 untuk $n = 300$

```
library(Rcpp)
library(RcppArmadillo)
library(stats)
library(rugarch)
library(MASS)
library(nnet)
library(car)
library(FinTS)
library(TSA)
garch11.1= ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm", fixed.pars =
list(omega=0.00054, alpha1=0.25, beta1=0.69))
set.seed(123)
garch11.sim1 = ugarchpath(garch11.1, n.sim = 300)
garch.spec11.1 = ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm")
modelfit11.1=ugarchfit(spec=garch.spec11.1,data=garch11.sim1 @path$seriesSim)
show(modelfit11.1)

#Membagi data in dan out sample dan mengestimasi parameter data in sample
data11.1=garch11.sim1 @path$seriesSim[1:270,]
spec11.1 = ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm")
fit11.1=ugarchfit(spec=spec11.1,data=data11.1)
show(fit11.1)
write.csv(garch11.sim1 @path, "D://GARCH11.1.csv")
```

### Lampiran 9. Simulasi GARCH(1,1) Skenario 1 untuk $n = 1000$

```
garch12.1= ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm", fixed.pars =
list(omega=0.00054, alpha1=0.25, beta1=0.69))
set.seed(123)
garch12.sim1 = ugarchpath(garch12.1, n.sim = 1000)
garch.spec12.1 = ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm")
modelfit12.1=ugarchfit(spec=garch.spec12.1,data=garch12.sim1 @path$seriesSim)
show(modelfit12.1)

#Membagi data in dan out sample dan mengestimasi parameter data in sample
data12.1=garch12.sim1 @path$seriesSim[1:970,]
spec12.1 = ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm")
fit12.1=ugarchfit(spec=spec12.1,data=data12.1)
show(fit12.1)
write.csv(garch12.sim1 @path, "D://GARCH12.1.csv")
```

### Lampiran 10. Simulasi GARCH(1,1) Skenario 2 untuk $n = 1000$

```
garch22.1= ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm", fixed.pars =
list(omega=0.000011, alpha1=0.09, beta1=0.9))
set.seed(123)
garch22.sim1 = ugarchpath(garch22.1, n.sim = 1000)
garch.spec22.1 = ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm")
modelfit22.1=ugarchfit(spec=garch.spec22.1,data=garch22.sim1 @path$seriesSim)
show(modelfit22.1)

#Membagi data in dan out sample dan mengestimasi parameter data in sample
data22.1=garch22.sim1 @path$seriesSim[1:970,]
spec22.1 = ugarchspec (variance.model = list(garchOrder = c(1,1)), mean.model =
list(armaOrder = c(0,0), include.mean=FALSE), distribution.model="norm")
fit22.1=ugarchfit(spec=spec22.1,data=data22.1)
show(fit22.1)
write.csv(garch22.sim1 @path, "D://GARCH22.1.csv")
```

### Lampiran 11. Syntax GARCH(1,1)-GA

```
rm(list=ls())
library(iterators)
library(foreach)
library(GA)
alldata=read.table("D://GARCH11.1.csv",sep=";",header=TRUE)
mydata=as.matrix(alldata)
training=mydata[1:270,]
testing=mydata[-c(1:270),]
nt=mydata[,4]

#Fungsi Fitness
my_fitness=function(x){
  omega=x[1]
  alpha=x[2]
  beta=x[3]
  #Model GARCH(1,1) data training
  for (i in 1:270){
    sigma2[1] = omega/(1-alpha-beta) #unconditional variance
    {
      sigma2[i+1] = omega + (alpha * nt[i]^2) + (beta * sigma2[i])
      #conditional variance out sample/ forecast
      for(j in 1:29){
        sigma2[271+j] = omega + (alpha * sigma2[270+j]) + (beta * sigma2[270+j])
      }
    }
    #fungsi fitness untuk optimasi GA / fungsi tujuan
    E=(nt[271:300])^2-(sigma2[271:300])
    for (i in 271:300){
      RMSE=(mean(E^2))^(1/2)
      invRMSE=1/RMSE
    }
    return(invRMSE)
  }
}

#Range nilai parameter
par.min=c(omega0=1e-10, alpha1=1e-5, beta1=1e-5)
par.max=c(omega0=2, alpha1=1, beta1=1)
#Run genetic algorithm
set.seed(1234)
optim=ga(type="real-valued", fitness=my_fitness, min=par.min,
names=names(par.min), max=par.max, popSize=100, maxiter=1000,
pcrossover=0.8, pmutation=0.1)
my_result=summary(optim)
my_result
#RMSE out sample
RMSE=1/my_result$fitness
RMSE
```

**Lampiran 12. Output Estimasi GARCH(1,1) Replikasi ke-1 Skenario 1 ( $n = 300$ )**

```

*-----*
*           GARCH Model Fit           *
*-----*

Conditional Variance Dynamics
-----
GARCH Model      : sGARCH(1,1)
Mean Model       : ARFIMA(0,0,0)
Distribution      : norm

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
omega   0.000535   0.000233   2.3006  0.021413
alpha1  0.274427   0.075663   3.6270  0.000287
beta1   0.659740   0.081079   8.1370  0.000000

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
omega   0.000535   0.000204   2.6241  0.008689
alpha1  0.274427   0.062855   4.3660  0.000013
beta1   0.659740   0.065146  10.1270  0.000000

LogLikelihood : 350.114

Information Criteria
-----
Akaike          -2.3141
Bayes           -2.2771
Shibata         -2.3143
Hannan-Quinn   -2.2993

Weighted Ljung-Box Test on Standardized Residuals
-----
                        statistic p-value
Lag[1]                  0.282  0.5954
Lag[2*(p+q)+(p+q)-1][2]  1.968  0.2682
Lag[4*(p+q)+(p+q)-1][5]  5.230  0.1357
d.o.f=0
H0 : No serial correlation

Weighted Ljung-Box Test on Standardized Squared Residuals
-----
                        statistic p-value
Lag[1]                  0.258  0.6115
Lag[2*(p+q)+(p+q)-1][5]  1.939  0.6328
Lag[4*(p+q)+(p+q)-1][9]  4.199  0.5554
d.o.f=2

Weighted ARCH LM Tests
-----
      Statistic Shape Scale P-Value
ARCH Lag[3]    0.5537 0.500 2.000 0.4568
ARCH Lag[5]    3.8322 1.440 1.667 0.1897
ARCH Lag[7]    4.4183 2.315 1.543 0.2902

Nyblom stability test
-----
Joint Statistic: 0.2452
Individual Statistics:
omega 0.1185
alpha1 0.1677
beta1 0.1554

```

Asymptotic Critical values (10% 5% 1%)  
 Joint Statistic: 0.846 1.01 1.35  
 Individual Statistic: 0.35 0.47 0.75

Sign Bias Test

	t-value	prob	sig
Sign Bias	0.9258	0.3553	
Negative Sign Bias	0.3677	0.7134	
Positive Sign Bias	0.3558	0.7223	
Joint Effect	0.9070	0.8237	

Adjusted Pearson Goodness-of-Fit Test:

group	statistic	p-value	(g-1)
1	20	13.20	0.8282
2	30	27.40	0.5501
3	40	39.73	0.4372
4	50	49.33	0.4598

Elapsed time : 0.1248

**Lampiran 13.** *Output* Estimasi GARCH(1,1) Data *In Sample* Replikasi ke-1 Skenario 1 ( $n = 300$ )

```

*-----*
*                GARCH Model Fit                *
*-----*

Conditional Variance Dynamics
-----
GARCH Model      : sGARCH(1,1)
Mean Model       : ARFIMA(0,0,0)
Distribution      : norm

Optimal Parameters
-----
      Estimate  Std. Error  t value  Pr(>|t|)
omega    0.000501   0.000242   2.0696  0.038494
alpha1   0.287891   0.080972   3.5554  0.000377
beta1    0.658320   0.083104   7.9216  0.000000

Robust Standard Errors:
      Estimate  Std. Error  t value  Pr(>|t|)
omega    0.000501   0.000231   2.1648  0.030399
alpha1   0.287891   0.069210   4.1597  0.000032
beta1    0.658320   0.070229   9.3739  0.000000

LogLikelihood : 308.2306

Information Criteria
-----
Akaike      -2.2610
Bayes       -2.2210
Shibata     -2.2612
Hannan-Quinn -2.2449

```





## SURAT PERNYATAAN

Saya yang bertanda tangan di bawah ini, mahasiswa Departemen Statistika FMKSD ITS.

Nama : Indah Fahmiyah  
NRP : 06211650010032  
Prodi : Magister Statistika  
Fakultas : Matematika, Komputasi dan Sains Data

menyatakan bahwa data yang digunakan dalam Tesis ini merupakan data sekunder yang diambil dari Tugas Akhir dari Diah Kusumawati (2011) yang berjudul Peramalan Kecepatan Angin Rata-Rata Harian di Surabaya Menggunakan Metode *Bayesian Model Averaging* dengan Pendekatan *Expectation Maximization*.

Surat pernyataan ini dibuat dengan sebenarnya. Apabila terdapat pemalsuan data maka saya siap menerima sanksi sesuai aturan yang berlaku.

Surabaya, 27 Juli 2018

Mengetahui,

Pembimbing Tesis

Mahasiswa



Irhamah, M.Si., Ph.D.  
NIP. 19780406 200112 2 002



Indah Fahmiyah  
NRP. 06211650010032

## BIODATA PENULIS



Nama lengkap penulis adalah Indah Fahmiyah. Penulis merupakan anak bungsu dari pasangan suami-istri, yaitu Ishaq dan Nafsiyah. Penulis berasal dari Sidoarjo dan berdomisili di Desa Suko, Sukodono, Sidoarjo. Pendidikan dasar penulis ditempuh di SDN Suko 1 (2000-2006). Kemudian, penulis menempuh pendidikan menengah di SMPN 1 Taman (2006-2009) dan SMAN 1 Sidoarjo (2009-2012). Pada tahun 2012, penulis diterima sebagai mahasiswa Program Sarjana Statistika ITS dan dinyatakan lulus pada 2016. Selanjutnya, penulis mendaftar beasiswa *fresh graduate* semester gasal Program Magister Statistika ITS dan resmi diterima pada tahun 2016 dengan NRP 1316201032 atau 06211650010032. Selama menjadi mahasiswa S1 Statistika ITS, penulis aktif mengikuti beberapa kepanitiaan dan lomba. Organisasi yang pernah diikuti selama S1 antara lain: HIMASTA-ITS, FORSIS ITS, JMMI ITS, dan Kopma dr. Angka ITS. Peneliti tetap mengikuti organisasi selama menempuh perkuliahan di S2 Statistika ITS, yaitu HMP-ITS (Himpunan Mahasiswa Pascasarjana ITS) dan FORSIMMPAS-ITS (Forum Silaturahmi Mahasiswa Muslim Pascasarjana ITS). Kritik dan saran terkait penelitian dalam Tesis dapat disampaikan kepada penulis. Penulis dapat dihubungi melalui email: [indahfahmiyah@gmail.com](mailto:indahfahmiyah@gmail.com).

*(Halaman ini sengaja dikosongkan)*