TESIS – KS185411

PERAMALAN *INDONESIAN CRUDE PRICE* (ICP) DENGAN METODE *BAYESIAN MARKOV SWITCHING* GARCH

RIA RETNA HERLIYASARI
06211850010011

Dosen Pembimbing
Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
Dr.rer.pol. Heri Kuswanto, S.Si, M.Si.

Program Magister
Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember
Surabaya
2020

TESIS – KS185411

PERAMALAN *INDONESIAN CRUDE PRICE* (ICP) DENGAN METODE *BAYESIAN MARKOV SWITCHING* GARCH

RIA RETNA HERLIYASARI
06211850010011

Dosen Pembimbing
Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
Dr.rer.pol. Heri Kuswanto, S.Si, M.Si.

Program Magister
Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember
Surabaya
2020

THESIS – KS185411

FORECASTING INDONESIAN CRUDE PRICE (ICP) USING BAYESIAN MARKOV SWITCHING GARCH

RIA RETNA HERLIYASARI
06211850010011

Supervisor
Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
Dr.rer.pol. Heri Kuswanto, S.Si, M.Si.

Magister Program
Department of Statistics
Faculty of Science and Data Analytics
Institut Teknologi Sepuluh Nopember
Surabaya
2020

LEMBAR PENGESAHAN TESIS

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

Oleh:
RIA RETNA HERLIYASARI
NRP: 06211850010011

Tanggal Ujian : 29 Juli 2020
Periode Wisuda : September 2020

Disetujui oleh:
Pembimbing:

1. Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
   NIP: 19621015 198803 1 002 .....

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

Penguji :

1. Dr. Dra. Kartika Fithriasari, M.Si.
   NIP: 19691212 199303 2 002 .....

2. Santi Puteri Rahayu, S.Si, M.Si., Ph.D.
   NIP: 19750115 199903 2 003 .....

Kepala Departemen Statistika
Fakultas Sains dan Analitika Data

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

Peramalan *Indonesian Crude Price* (ICP) dengan *Bayesian Markov Switching* GARCH

Nama Mahasiswa	: Ria Retna Herliyasari
NRP	: 062118 5001 0011
Pembimbing	: Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
Co-pembimbing	: Dr.rer.pol. Heri Kuswanto, S.Si, M.Si.

ABSTRAK

Harga minyak mentah Indonesia atau *Indonesian Crude Price* (ICP) merupakan salah satu indikator makroekonomi sehingga merupakan faktor penting dalam penentuan APBN. Fluktuasi harga ICP sangat berpengaruh terhadap penerimaan dan belanja negara sehingga diperlukan prediksi ICP mendekati nilai realisasinya. ICP dapat diprediksi menggunakan analisis deret waktu karena besaran ICP merupakan observasi yang terurut. Data ICP memiliki pola yang fluktuatif sehingga memiliki variansi yang tidak homogen atau terjadi kasus heteroskedastisitas. Metode peramalan yang dapat menangkap heteroskedastisitas adalah *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). Namun, model GARCH tidak dapat mengetahui perubahan struktur dalam variansi. *Markov switching* GARCH dapat mengatasi kasus heteroskedastisitas sekaligus perubahan antar struktur variansi. Pada penelitian ini *Markov switching* GARCH diterapkan untuk memodelkan ICP dan dibandingkan dengan model GARCH tanpa *switching*. Berdasarkan nilai WAIC, MSGARCH memiliki kinerja yang lebih baik dibandingkan GARCH. Hasil peramalan menunjukkan bahwa pada Januari 2019 ICP harga ICP mengalami kenaikan menjadi US$ 67,21 dengan variansi sebesar 13,24.

Kata Kunci: *GARCH, Indonesian Crude Price, Markov switching.*

(Halaman ini sengaja dikosongkan)

Forecasting Indonesian Crude Price (ICP) Using Bayesian Markov Switching GARCH

By	: Ria Retna Herliyasari
Student Identity Number	: 062118 5001 0011
Supervisor	: Prof. Drs. Nur Iriawan, M.Ikom, Ph.D.
Co-supervisor	: Dr.rer.pol. Heri Kuswanto, S.Si, M.Si.

ABSTRACT

Indonesian Crude Price (ICP) is one of the macroeconomic indicators so it is an important factor in determining the state budget. ICP price fluctuations greatly affect state revenues and expenditures so that ICP predictions are needed close to their realization value. ICP can be predicted using time series analysis because ICP quantities are ordered observations. ICP data has a fluctuating pattern so that it has a non-homogenous variance or a case of heteroscedasticity. Forecasting methods that can capture heteroscedasticity is Generalized Autoregressive Conditional Heteroscedasticity (GARCH). However, the GARCH model cannot find out structural changes in variance. Markov switching GARCH can overcome the case of heteroscedasticity as well as changes between variance structures. In this study Markov switching GARCH is applied to model ICP and compared to the GARCH model without switching. Based on the WAIC value, MSGARCH has better performance than GARCH. Forecasting results indicate that in January 2019 ICP prices increased to US$ 67.21 with a variance of 13.24.

Keywords: GARCH, Indonesian Crude Price, Markov switching.

(Halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji Syukur atas kehadirat Allah SWT yang telah melimpahkan rahmat, taufik, hidayah, serta karunia-Nya sehingga penulis dapat menyelesaikan laporan Tesis yang berjudul "**Peramalan *Indonesian Crude Price* (ICP) dengan *Bayesian Markov Switching* GARCH**" ini tepat pada waktunya. Penulisan laporan Tesis ini tidak akan berjalan dengan lancar tanpa bantuan dan dukungan beberapa pihak. Oleh karena itu, penulis ingin mengucapkan terimakasih kepada :

1. Ibu Dr. Kartika Fithriasari, M.Si selaku Kepala Departemen Statistika ITS dan dosen penguji yang telah memberikan masukan kepada penulis untuk hasil Tesis yang lebih baik.
2. Ibu Santi Wulan Purnami, M.Si, Ph.D selaku Sekretaris Departemen I (Bidang Akademik, Kemahasiswaan, Penelitian dan Pengabdian Kepada Masyarakat) dan Ibu Dr. Vita Ratnasari, M.Si selaku Sekretaris Departemen II (Bidang Sumberdaya Keuangan, Sumber Daya Manusia, dan Sarana Prasarana yang telah memberikan banyak fasilitas sehingga membantu penyelesaian Tesis.
3. Bapak Dr.rer.pol Dedy Dwi Prastyo, M.Si selaku Kepala Program Studi Pascasarjana Statistika ITS yang telah memberikan banyak fasilitas sehingga membantu penyelesaian Tesis.
4. Prof. Drs. Nur Iriawan, M.Ikom, Ph.D dan Dr.rer.pol. Heri Kuswanto, M.Si selaku dosen pembimbing yang telah banyak meluangkan waktu untuk membimbing dan mengarahkan penulis selama menyusun Tesis.
5. Santi Puteri Rahayu, M.Si, Ph.D selaku dosen penguji yang telah memberikan masukan kepada penulis untuk hasil Tesis yang lebih baik.
6. Institut Teknologi Sepuluh Nopember karena telah memberikan beasiswa kepada penulis melalui beasiswa *fresh graduate*, sehingga penulis bisa melanjutkan pendidikan pascasarjana hingga selesai.
7. Bapak Antonius Sumitro yang telah banyak membantu kelancaran administrasi Tesis di Departemen Statistika FSAD ITS.
8. Kedua orang tua (Heriyanto dan Suherlin), Arta Gilang, Tino, Dirga, serta keluarga lainnya yang telah memberikan dukungan, do'a, dan semangat sehingga penulis dapat menyelesaikan Tesis dengan baik.

9. Bu Arifatus Sholikah yang telah banyak membantu dalam pembuatan program Stan.
10. Teman-teman seangkatan S2 Statistika ITS Angkatan 2018 Ganjil yang telah berjuang bersama untuk menyelesaikan pendidikan ini dan seluruh pihak yang belum dapat penulis sebutkan satu per satu dalam penyusunan Tesis ini.

Penulis menyadari bahwa laporan Tesis ini masih jauh dari kesempurnaan. Oleh karena itu, kritik dan saran yang bersifat membangun sangat diharapkan agar nantinya menjadi koreksi untuk penelitian yang akan datang. Penulis juga berharap agar hasil Tesis ini dapat bermanfaat bagi semua pihak. Semoga kebaikan dan bantuan yang telah diberikan kepada penulis dibalas oleh Allah SWT.

Surabaya, Agustus 2020

Ria Retna Herliyasari

DAFTAR ISI

HALAMAN JUDUL i
TITLE PAGE iii
LEMBAR PENGESAHAN v
ABSTRAK vii
ABSTRACT ix
KATA PENGANTAR xi
DAFTAR ISI xiii
DAFTAR TABEL xv
DAFTAR GAMBAR xvii
BAB I PENDAHULUAN 1
1.1 Latar Belakang 1
1.2 Rumusan Masalah 5
1.3 Tujuan 6
1.4 Manfaat 6
1.5 Batasan Masalah 6
BAB II TINJAUAN PUSTAKA 9
2.1 Analisis Deret Waktu 9
2.2.1 Uji Stasioneritas 9
2.2.2 Uji Heteroskedastisitas 11
2.2.3 Uji Perubahan Struktur 13
2.2 Model GARCH 14
2.3 Model MSGARCH 15
2.4 Estimasi Parameter dengan Metode *Bayesian* 17
2.4.1 Konsep *Bayesian* 18
2.4.2 Distribusi Prior 15
2.4.3 *Hamiltonian Monte Carlo* 19
2.4.4 *No-U-Turn Sampler* 22
2.5 Kriteria Konvergen 25

2.6 *Credible Interval* 26
2.7 *Watanabe-Akaike Information Criterion* 26
2.9 *Indonesian Crude Price* 28
BAB III METODOLOGI PENELITIAN..... 31
3.1 Sumber Data dan Variabel Penelitian 31
3.2 Tahapan Analisis Data 31
BAB IV ANALISIS DAN PEMBAHASAN 35
4.1 Karakteristik Data ICP 35
4.2 Pemodelan ICP 38
4.2.1 Pemodelan ICP dengan MSGARCH 38
4.2.2 Pemodelan ICP dengan GARCH 42
4.2.3 Pemilihan Model Terbaik 44
4.4 Peramalan ICP 45
BAB V KESIMPULAN DAN SARAN 49
5.1 Kesimpulan 49
5.2 Saran 49
DAFTAR PUSTAKA
LAMPIRAN
BIODATA PENULIS

DAFTAR TABEL

Tabel 2.1 Distribusi Prior 19
Tabel 3.1 Struktur Data Penelitian 31
Tabel 4.1 Statistik Deskriptif ICP 37
Tabel 4.2 *Convergence Diagnostic* Parameter MSGARCH 40
Tabel 4.3 Estimasi Parameter MSGARCH 40
Tabel 4.4 *Convergence Diagnostic* Parameter GARCH 43
Tabel 4.5 Estimasi Parameter GARCH 43
Tabel 4.6 Perbandingan Kebaikan Model MSGARH dan GARCH 44

(Halaman ini sengaja dikosongkan)

DAFTAR GAMBAR

Gambar 1.1 ICP Januari 1989 – Desember 2019.....2
Gambar 2.1 Plot Stasioneritas.....10
Gambar 2.2 Plot Heteroskedastisitas.....11
Gambar 2.3 Plot Perubahan Struktur.....13
Gambar 3.1 Diagram Alir Penelitian.....33
Gambar 4.1 *Time Series* Plot ICP yang di-*differencing* pada *lag* 1....35
Gambar 4.2 Histogram ICP yang di-*differencing* pada *lag* 1.....36
Gambar 4.3 *Structural Break Point* ICP 37
Gambar 4.4 Simulasi Historis Parameter MSGARCH 39
Gambar 4.6 Simulasi Historis Parameter GARCH.....42
Gambar 4.7 Prediksi *Mean* ICP 44
Gambar 4.8 Prediksi *Mean* dan Variansi ICP.....45

(Halaman ini sengaja dikosongkan)

BAB I
PENDAHULUAN

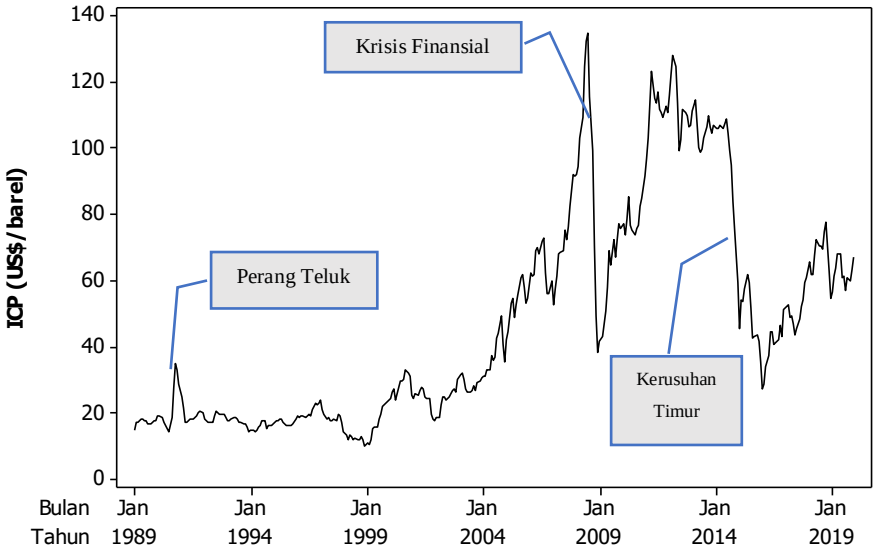
1.1 Latar Belakang

Minyak dan gas bumi merupakan sumber daya alam yang strategis tidak terbarukan yang dikuasai oleh negara dan mempunyai peranan penting dalam memberikan nilai tambah secara nyata kepada pertumbuhan ekonomi nasional. Karena peranannya yang sangat penting, maka pengelolaannya harus dilakukan secara maksimal sehingga memberikan kemakmuran dan kesejahteraan rakyat. Minyak mentah Indonesia adalah minyak bumi seperti yang telah dijelaskan pada Undang-Undang Republik Indonesia No 22 tahun 2001 terdiri atas minyak mentah utama dan minyak mentah lainnya. Berdasarkan Peraturan Menteri Energi dan Sumber daya Mineral Republik Indonesia Nomor 23 Tahun 2012, terdapat dua metode yang digunakan untuk menetapkan formula harga minyak mentah Indonesia yaitu metode *bencmarking* dan atau indeksasi. Metode *bechmarking* digunakan untuk penetapan harga minyak mentah utama Indonesia yang di perdagangkan di pasar internasional dan harganya dipublikasikan di Publikasi Internasional, sedangkan metode indeksasi digunakan untuk menentukan harga minyak mentah lainnya yang tidak diperdagangkan di pasar internasional dan atau harganya tidak dipublikasikan oleh Publikasi Internasional.

Harga minyak mentah merupakan salah satu indikator ekonomi makro, sehingga berpengaruh terhadap penyusunan Anggaran Pendapatan dan Belanja Negara (APBN). Acuan yang digunakan untuk asumsi harga minyak adalah harga minyak mentah Indonesia atau *Indonesian Crude Price* (ICP). Besaran ICP sangat berpengaruh terhadap beberapa pos penting dalam APBN seperti penerimaan migas, subsidi energi (BBM, listrik, BBN, LPG), dan Dana Bagi Hasil Migas (Handoko, 2014). Dari segi penerimaan, ICP dapat memengaruhi target penerimaan negara dari SDA migas. Sedangkan dari segi belanja negara, ICP dapat memengaruhi besaran belanja subsidi energi. Jika realisasi ICP lebih tinggi dari asumsi yang ditetapkan, maka penerimaan negara akan tinggi, tetapi beban subsidi juga tinggi. Pergerakan harga minyak seringkali tidak menentu dan fluktuatif akibat faktor fundamental seperti permintaan dan penawaran, atau karena faktor

non-fundamental seperti faktor geopolitik, spekulasi dan cuaca. Fluktuasi harga ICP sangat berpengaruh terhadap penerimaan dan belanja negara sehingga diperlukan perhitungan asumsi ICP yang tepat dimana mendekati nilai realisasinya. Fluktuasi ICP dapat diprediksi menggunakan analisis deret waktu karena besaran ICP merupakan observasi yang terurut.

Pola pergerakan harga minyak mentah Indonesia cukup berfluktuasi seperti yang ditunjukkan pada Gambar 1.1. Pada periode Januari 1989 sampai dengan Desember 1998 ICP masih cenderung stabil dengan rata-rata sekitar US$ 17,29 per barel, kecuali pada saat terjadinya Perang Teluk pada Agustus 1990 hingga Februari 1991. Perang Teluk menyebabkan penurunan produksi minyak dalam waktu singkat sehingga harga minyak mengalami kenaikan yang cukup tajam. Pada bulan September 1990 ICP naik menjadi US$ 27,63 per barel dan mencapai puncaknya yaitu sebesar US$ 34,88 per barel pada Oktober 1990 hingga kemudian perlahan turun dan stabil kembali setelah Perang Teluk dinyatakan selesai.



Gambar 1.1 ICP Januari 1989 – Desember 2019

ICP juga cenderung mengalami kenaikan pada periode-periode tertentu yakni pada periode Januari 1999 sampai dengan Juli 2008, periode Januari 2009 sampai dengan Juni 2014 dan periode Februari 2015 sampai dengan Mei 2018. Selain itu, terdapat periode-periode tertentu dimana ICP mengalami penurunan yang signifikan yaitu pada Agustus 2008 sampai dengan Desember 2008. Penurunan ini

terjadi akibat adanya krisis finansial di AS. Penurunan yang tajam juga terjadi pada periode Juli 2014 hingga Januari 2015 yang disebabkan oleh kerusuhan di Timur Tengah. Dengan demikian, dapat dilihat bahwa ICP menunjukkan ketidakpastian yang tinggi dari waktu ke waktu berdasarkan tinjauan sejarah harga minyak yang telah dijelaskan sebelumnya.

Beberapa kasus dalam *time series* dapat diselesaikan dengan asumsi bahwa struktur deret waktu dapat dideskripsikan sebagai model stasioner *Autoregressive Moving Average* (ARMA). Salah satu asumsi yang harus dipenuhi dalam pemodelan dengan ARMA adalah stasioner. Pada kasus data yang tidak stasioner dalam *mean* data dapat dimodelkan dengan *Autoregressive Integrated Moving Average* (ARIMA) (Box & Jenkins, 1976). Namun, pada beberapa kasus seperti pada data ekonomi, asumsi stasioneritas dalam variansi seringkali tidak terpenuhi karena data memiliki fluktuasi yang tidak menentu atau variansi yang tidak homogen. Apabila asumsi tersebut dilanggar, maka pemodelan menggunakan ARIMA akan menghasilkan variansi *error* yang tidak konstan atau yang disebut heteroskedastisitas (Wei, 2006).

Engle (1982) memperkenalkan model *Autoregressive Conditional Heteroscedasticity* (ARCH) untuk memodelkan variansi error. Model ARCH dalam penerapannya memiliki kelemahan yaitu ketika orde ARCH besar akan menyebabkan berkurangnya presisi estimator. Kemudian Bollerslev (1986) mengembangkan metode ARCH menjadi *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). Namun, model ARCH dan GARCH tidak dapat mendeteksi perubahan struktur variansi.

Sebagai solusi pemodelan data runtun waktu yang mengalami perubahan struktur (*structural break*), Hamilton (1989) memperkenalkan model *Markov switching*. Pada model *Markov switching*, perubahan struktur yang terjadi pada data runtun waktu tidak dianggap sebagai kejadian yang deterministik, tetapi dianggap sebagai suatu kejadian random yang tidak teramati yang diwujudkan sebagai variabel *state* atau *regime*. Hamilton (1989) menggabungkan model *Markov switching* dengan model *autoregressive* sehingga menghasilkan model yang dapat menangkap pergeseran *mean*. Namun model tersebut tidak dapat mendeteksi pergeseran variansi. Selanjutnya, Hamilton dan Susmel (1994)

mengembangkan model *Markov switching* ARCH yang mampu menjelaskan adanya persgeseran variansi. Kemudian Gray (1996) memperkenalkan model *Markov switching* GARCH (MSGARCH) yang memiliki karakteristik hampir sama dengan *Markov switching* ARCH tetapi parameter yang dilibatkan lebih sederhana.

Penelitian menggunakan model MSGARCH telah banyak diterapkan. Marcucci (2005) dan Klassen (2002) menerapkan MSGARCH untuk memodelkan perubahan volatilitas pada *stock market*. Hadiyat, Iriawan, dan Mulyono (2007) mengidentifikasi dan menerapkan *Markov switching* GARCH untuk menangkap keberadaan *regime* volatilitas yang berbeda yaitu *regime* volatilitas tinggi dan *regime* volatilitas rendah dalam *return* indeks Dowjones. Hasil matriks transisi menunjukkan bahwa *regime* volatilitas rendah lebih sering terjadi daripada model *regime* dengan volatilitas tinggi. Gunay (2015)menggunakan model MSGARCH untuk memodelkan harga minyak mentah jenis *Brent* dan menyimpulkan bahwa model *Markov switching* GARCH memiliki performa yang lebih baik dibandingkan model GARCH tanpa pergantiang *regime*. Runfang, Jiangze, & Xiaotao (2017) meramalkan *Oil Market Volatility* dengan MSGARCH.

Beberapa peneliti menerapkan pendekatan *Bayesian* dalam estimasi parameter MSGARCH. Hal ini dikarenakan ketika melakukan estimasi parameter pada model *Markov Swithcing* GARCH menggunakan *Maximum Likelihood Estimation* menimbulkan kendala dalam proses implementasinya karena struktur dari GARCH yaitu pada model variansi bergantung pada semua informasi variabel *state* pada waktu sebelumnya. Jika dimisalkan ukuran *state* adalah J dan ukuran observasi adalah T, maka harus mempertimbangkan sebanyak JT kombinasi untuk mendapatkan fungsi *likelihood* sehingga hal ini mustahil untuk dilakukan. Kemudian, Das (2004) menunjukkan bahwa permasalahan dalam estimasi parameter dapat diatasi dengan mudah menggunakan *Bayesian*. Pada *Bayesian inference*, variabel *state* dianggap sebagai variabel random dan fungsi *likelihood* dibangun dengan asumsi bahwa variabel *state* telah diketahui sebelumnya. Dengan demikian pembentukan fungsi *likelihood* menjadi mudah yaitu dengan mendefinisikan fungsi kontribusi *likelihood* pada setiap *state*, kemudian

menetapkan fungsi *likelihood* dengan menambahkan bobot probabilitas transisi dan probabilitas berada pada *state* tertentu.

Dalam penelitian ini, parameter MSGARCH diestimasi menggunakan pendekatan *Bayesian* menggunakan bahasa pemrograman Stan sehingga algoritma yang digunakan dalam proses integrasi adalah *No-U-Turn Sampler* (NUTS). NUTS merupakan pengembangan dari *Hamiltonian Monte Carlo* (HMC). Sedangkan HMC sendiri adalah salah satu metode MCMC yang menggunakan konsep dinamika Hamiltonian. Metode HMC dapat menghindari perilaku *random walk* dalam pengambilan sampel yang merupakan kelemahan dari metode MCMC. Namun kinerja HMC sangat sensitif terhadap dua hiperparameter, sehingga NUTS diperkenalkan untuk mengotomatiskan pengaturan hiperparameter tersebut. Menurut Nishio dan Arakawa (2019) metode NUTS memiliki kemampuan yang lebih baik daripada MCMC maupun HMC dalam hal kualitas sampel yang dihasilkan yaitu ukuran sampel efektif yang besar, autokorelasi rendah dan kemiringan (*skewness*) rendah dari distribusi posterior marginal.

Dengan demikian model MSGARCH dengan pendekatan *Bayesian* digunakan untuk memodelkan *Indonesian Crude Price* (ICP) yang memiliki pola fluktuatif. Metode ini diharapkan mampu meramalkan ICP yang memiliki pergantian pola yaitu pola fluktuasi rendah pada periode tertentu dan mengalami fluktuasi yang tinggi pada periode yang lain dengan akurat.

1.2 Rumusan Masalah

Proses penyusunan APBN memerlukan beberapa asumsi ekonomi makro, salah satunya adalah asumsi harga minyak mentah. Acuan yang digunakan untuk asumsi harga minyak mentah adalah *Indonesian Crude Price* (ICP). Penentuan asumsi ICP dapat dilakukan dengan cara meramalkan nilai ICP yang akan terjadi pada periode ke depan melalui informasi nilai ICP pada periode sebelumnya. Oleh karena itu, dibutuhkan metode peramalan ICP yang akurat mendekati nilai yang sebenarnya. Jika hasil ramalan ICP tidak akurat, misalnya nilai ICP yang sebenarnya lebih tinggi dari asumsi yang ditetapkan, maka akan mengakibatkan penerimaan negara akan tinggi, tetapi beban subsidi juga tinggi. ICP memiliki pola fluktuasi yang tidak menentu yaitu terjadi fluktuasi rendah pada periode tertentu, dan fluktuasi yang tinggi pada saat yang lain. Hal ini menyebabkan resiko variansi

tidak homogen sehingga memungkinkan terjadinya kasus heteroskedastisitas. Adanya kasus heteroskedastisitas dapat menyebabkan prediksi ICP tidak akurat. Agar dapat menangkap kasus heteroskedastisitas, dapat digunakan model *Generalized Autoregressive Conditional Heteroscedasticity* (GARCH). Selain itu, ICP merupakan data ekonomi yang diduga memiliki perubahan struktur variansi. Pemodelan yang dapat menangkap perubahan struktur variansi salah satunya adalah *Markov switching* GARCH. Dengan demikian, rumusan masalah yang pertama pada penelitian ini adalah bagaimana memodelkan *Indonesian Crude Price* (ICP) dengan *Markov switching* GARCH dengan pendekatan *Bayesian* yaitu dengan menggunakan integrasi numerik *No-U-Turn Sampler* (NUTS) dan membandingkan kinerjanya dengan model GARCH. Setelah didapatkan model terbaik, kemudian meramalkan *Indonesian Crude Price* (ICP) menggunakan model terbaik.

1.3 Tujuan

Berdasarkan rumusan masalah pada penelitian ini, didapatkan tujuan penelitian ini adalah:

1. Mendapatkan karakteristik data ICP.
2. Mendapatkan model *Indonesian Crude Price* (ICP) melalui metode *Bayesian Markov switching* GARCH.
3. Mendapatkan hasil ramalan satu langkah ke depan *Indonesian Crude Price* (ICP) model terbaik.

1.4 Manfaat

Manfaat yang diharapkan dari penelitian ini adalah sebagai berikut.

1. Memberikan masukan kepada pemerintah mengenai metode peramalan yang baik untuk meramalkan besaran ICP.
2. Hasil penelitian ini dapat digunakan sebagai bahan referensi bagi penelitian-penelitian selanjutnya.

1.5 Batasan Masalah

Berikut ini adalah batasan-batasan masalah dalam penelitian ini.

1. Data yang digunakan dalam penelitian ini adalah data rata-rata harga minyak mentah Indonesia atau *Indonesian Crude Price* (ICP) bulanan dari bulan Januari 1989 sampai bulan Desember 2019.

2. *State* pada pemodelan *Markov switching* yang digunakan dalam penelitian ini adalah dua *state*.

(Halaman ini sengaja dikosongkan)

BAB II
TINJAUAN PUSTAKA

Bab II ini berisi tentang teori dan penjelasan yang mendasari analisis dalam penelitian ini. Pada subbab pertama membahas mengenai analisis deret waktu yang terdiri dari uji stasioneritas, uji heteroskedastisitas dan uji perubahan struktur. Subbab kedua dan subbab ketiga masing-masing membahas tentang model GARCH dan model MSGARCH. Subbab keempat estimasi parameter dengan metode *Bayesian* yang terdiri dari konsep *Bayesian*, distribusi prior, *Hamiltonian Monte Carlo*, dan *No-U Turn Sampler*. Subbab selanjutnya yaitu subbab kelima sampai dengan subbab kedelapan masing-masing membahas mengenai kriteria konvergen, *credible interval*, *Watanabe Akaike Information Criterion*, dan *Indonesian Crude Price*.

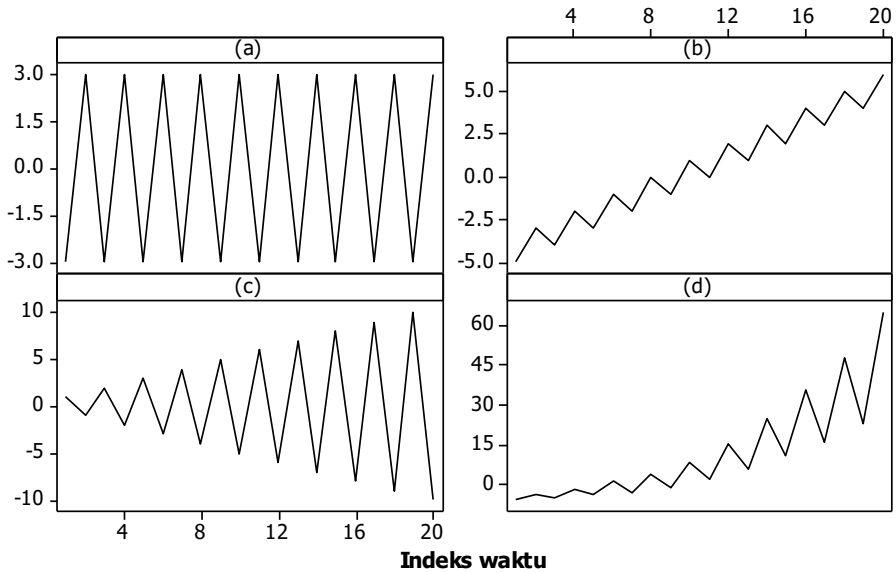
2.1 Analisis Deret Waktu

Data deret waktu didefinisikan sebagai himpunan observasi yang terurut terhadap dimensi waktu. Secara umum terdapat beberapa macam pola data dalam analisis deret waktu yaitu tren, siklis, musiman, dan *irregular* (Hanke & Wichern, 2014). Pola tren didefinisikan sebagai kecenderungan arah data dalam jangka panjang yang dapat berupa kenaikan atau penurunan. Pola siklis terjadi ketika data dipengaruhi oleh fluktuasi jangka panjang. Pola musiman terjadi ketika data mempunyai pola berulang dari periode satu ke periode berikutnya, misalnya pola harian, mingguan, bulanan, atau tahunan. Sedangkan pola *irregular* adalah pola data yang tidak terprediksi atau memiliki fluktuasi yang random.

Pemodelan data deret waktu dapat digunakan untuk meramalkan data periode waktu ke depan. Peramalan data kuantitatif dapat diterapkan jika memenuhi tiga kondisi yaitu tersedianya informasi masa lalu, informasi tersebut dapat dibentuk menjadi data numerik, dan dapat diasumsikan bahwa aspek pola data yang terjadi di masa lalu akan terus berlanjut di masa mendatang. Asumsi yang diperlukan untuk menentukan model adalah data dalam keadaan stasioner.

Stasioneritas adalah tidak adanya kenaikan atau penurunan dalam data. Fluktuasi data berada di sekitar nilai *mean* yang konstan, tidak bergantung waktu

dan tidak menunjukkan penambahan variansi yang signifikan dari waktu ke-waktu. Secara visual, kondisi stasioneritas dapat dilihat pada Gambar 2.1.



Gambar 2.1 Plot Stasioneritas: (a) Stasioner dalam mean dan variansi (b) Stasioner dalam variansi namun tidak stasioner dalam mean (c) Stasioner dalam mean namun tidak stasioner dalam variansi (d) Tidak stasioner dalam mean dan variansi

2.1.1 Uji Stasioneritas

Stasioneritas dalam *mean* dapat diuji menggunakan uji *Augmented Dickey Fuller* (ADF) dengan persamaan (2.1) .

$$\Delta y_t = \gamma + \varphi y_{t-1} + \sum_{i=1}^{l} \psi_i \Delta y_{t-1} + e_t \qquad (2.1)$$

dimana:

Δy_t = $y_t - y_{t-1}$

γ = intersep

φ = koefisien dari y_{t-1}

ψ_i = koefisien dari Δy_{t-1}

e_t = *error* yang *white noise*

l = banyak *lag*.

Hipotesis yang digunakan dalam pengujian *Augmented Dickey-Fuller* adalah sebagai berikut.

$H_0 : \varphi = 0$ (data tidak stasioner dalam *mean*)

$H_1 : \varphi < 0$ (data stasioner dalam *mean*)

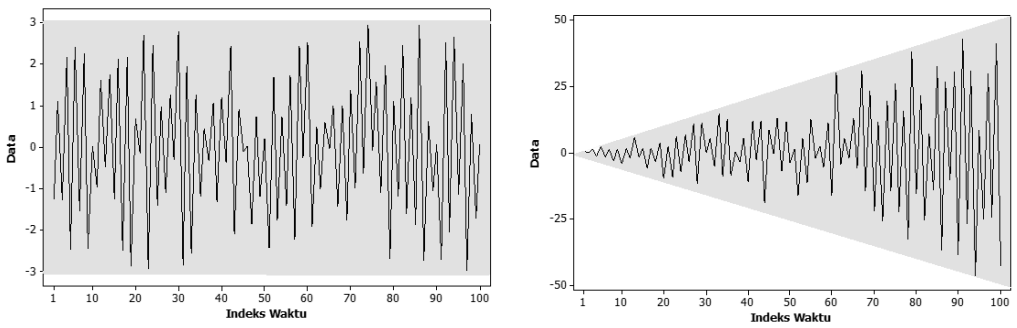
Penolakan hipotesis akan dilakukan dengan menggunakan statistik uji dalam persamaan (2.2)

$$\delta = \frac{\hat{\varphi}}{SE\left(\hat{\varphi}\right)}. \tag{2.2}$$

Keputusan menolak H_0 diperoleh ketika nilai $|\delta|$ lebih besar daripada nilai δ dengan derajat bebas T (banyak pengamatan) dan taraf signifikasnsi α dalam Tabel *Dickey Fuller*. Jika hal ini terpenuhi, maka berarti data memiliki pola stasioner dalam *mean* (Wei, 2006).

2.1.2 Uji Heteroskedastisitas

Suatu data dikatakan heteroskedastisitas apabila suatu data memiliki variansi residual yang tidak konstan atau tidak homogen untuk setiap observasi. Gambaran suatu data yang memiliki kasus heteroskedastisitas dapat dilihat pada Gambar 2.2.



Gambar 2.2 Plot Heteroskedastisitas (a) Tidak memiliki kasus heteroskedastisitas (b) Memiliki kasus heteroskedastisitas

Efek heteroskedastisitas atau yang biasa disebut efek *autoregressive conditional heteroscedasticity* (ARCH) dapat diperiksa melalui uji *Lagrange Multiplier* (LM). Prinsip dalam uji LM adalah dengan meregresikan kuadrat

residual dengan *lag*-nya sendiri. Misalkan diberikan persamaan model *mean* pada persamaan (2.3):

$$y_t = \mu + \varepsilon_t \tag{2.3}$$

dimana ε_t merupakan residual dari model *mean* dan tidak berkorelasi namun memiliki variansi yang berubah sepanjang waktu. Menurut Engle (1982), bentuk residual dapat dimodelkan sebagai $\varepsilon_t = \sigma_t u_t$, diman σ_t adalah variansi pada waktu ke t dan u_t adalah IIDN(0,1). Untuk menguji efek ARCH, makadiberikan persamaan regresi kuadrat residual pada persamaan (2.4)

$$\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \ldots + \alpha_l \varepsilon_{t-l}^2 \tag{2.4}$$

Hipotesis yang digunakan dalam pengujian LM adalah sebagai berikut.

H_0 : $\alpha_1 = \ldots = \alpha_l = 0$ (tidak terdapat efek ARCH sampai *lag* ke-*l*).

H_1 : minimal ada satu $\alpha_i \neq 0$; $i = 1, \ldots, l$ (terdapat efek ARCH sampai *lag* ke-*l*).

Penolakan hipotesis akan dilakukan dengan menggunakan statistik uji dalam persamaan (2.5)

$$\text{LM} = TR^2 \tag{2.5}$$

dimana:

T = ukuran sampel

R^2 = koefisien determinasi.

$$R^2 = 1 - \frac{SSE}{SST} \tag{2.6}$$

dimana:

SSE = *sum square error*

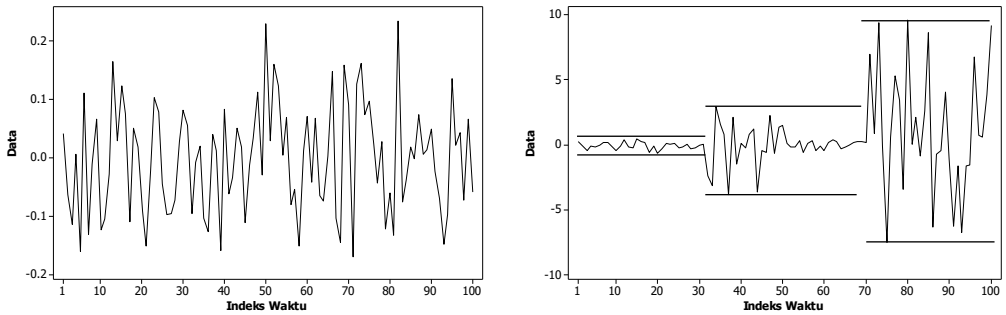
SST = *sum square total*

Statistik uji LM kemudian dibandingkan dengan nilai tabel χ_l^2 untuk mengambil keputusan penolakan hipotesis nul, H_0. Jika estimasi nilai α_i mendekati nol sehingga nilai koefisien determinasi kecil, maka H_0 cenderung gagal ditolak. Jika estimasi nilai α_i tidak mendekati nol sehingga nilai koefisien determinasi besar, maka H_0 cenderung ditolak. Nilai koefisien determinasi yang besar akan mengakibatkan nilai statistik uji LM juga besar. Dengan demikian, jika

nilai statistik uji LM lebih besar daripada nilai χ_l^2 sehingga didapatkan kesimpulan bahwa terdapat efek ARCH atau heteroskedastisitas sampai *lag-l* pada data runtun waktu.

2.1.3 Uji Perubahan Struktur

Algoritma *iterated cumulative sum of squares* (ICSS) adalah uji yang digunakan untuk mengetahui adanya perubahan struktur (*structural breaks*) pada volatilitas energi. Beberapa peneliti telah menerapkan uji ICSS ini untuk mendeteksi perubahan mendadak pada data return energi (Jiang, et al., 2018). Perubahan ini yang akan digunakan sebagai dasar adanya perubahan *state* pada data runtun waktu. Gambaran adanya perubahan struktur (*structural breaks*) dalam variansi dapat dilihat melalui gambar 2.3 sebagai berikut.



Gambar 2.3 Plot Perubahan Struktur (a) Tidak ada perubahan struktur (b) Ada perubahan struktur

Diasumsikan model deret waktu:

$$y_t = \mu + \varepsilon_t \tag{2.7}$$

dimana:

μ = *mean*

ε_t = residual IIDN $\left(0, \sigma^2\right)$

Dicurigai di setiap *t* akan mempunyai variansi yang berbeda, σ_t^2. Oleh sebab itu perlu dilakukan pengujian dengan menggunakan hipotesis sebagai berikut.

H_0 : tidak ada perubahan struktur variansi

H_1 : ada perubahan struktur variansi

Statistik uji yang digunakan dalam pengujian ini adalah statistik Inclan Tiao (IT) sebagai berikut (Inclan & Tiao, 1994).

$$\text{IT} = \sup_k \left| \sqrt{\frac{T}{2} D_k} \right| \tag{2.8}$$

dimana:

$$D_k = \frac{d_k}{d_T} - \frac{k}{T} \; ; k = 1, 2, ..., T$$

$$d_k = \sum_{t=1}^{k} \varepsilon_t^2$$

$$d_T = \sum_{t=1}^{T} \varepsilon_t^2$$

Statistik uji IT dibandingkan dengan nilai pada tabel Inclan & Tiao dengan taraf signifikansi 0,05 dan derajat bebas *T*. Jika nilai IT lebih besar daripada nilai tabel, maka terjadi perubahan struktur pada variansi. Jika *break point* pada variansi tidak terjadi dalam proses iterasi, maka nilai statistik D_k akan berada di sekitar nol. Sebaliknya, jika ada satu atau lebih *break points* maka nilai D_k akan menyimpang secara signifikan di atas atau di bawah nol.

2.2 Model GARCH

Model GARCH pertama kali diperkenalkan oleh Bollerslev (1986) yang merupakan pengembangan dari model ARCH yang diusulkan oleh Engle (1982). Model GARCH digunakan untuk memodelkan deret waktu yang memiliki kasus heteroskedastisitas atau variansi yang tidak konstan. Berikut adalah model GARCH.

$$\begin{aligned} y_t &= \mu + \varepsilon_t \\ \varepsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \end{aligned} \tag{2.9}$$

dimana:

y_t = data runtun waktu

μ = *mean* model

ε_t = residual pada waktu *t*

σ_t = volatilitas pada waktu *t*

u_t = residual berdistribusi $N(0,1)$

$\alpha_0, \alpha_1, \beta_1$ = parameter model GARCH

Untuk menjamin kepositifan, maka terdapat beberapa fungsi batasan yaitu $\alpha_0 > 0,\ \alpha_1 \geq 0,$ dan $\beta_1 \geq 0$. Stasioneritas dalam variansi didapatkan dengan syarat $\alpha_1 + \beta_1 < 1$ (Ardia, et al., 2018).

2.3 Model MSGARCH

Parameter model GARCH umumnya dianggap konstan, tetapi pergerakan return finansial antara resesi dan ekspansi yang berbeda mengakibatkan volatilitas yang berbeda pula. Gray (1996) dalam Sopipan, Sattayaham & Premanode (2012) memperluas model GARCH ke model *Markov switching* GARCH (MSGARCH) untuk menangkap perubahan *regime* dalam volatilitas dengan variabel *state* yang tidak dapat diobservasi secara langsung. Variabel *state* yang tidak dapat diobservasi secara langsung diasumsikan memenuhi proses *first order Markov Chain*.

Model *Markov Switching* GARCH dengan dua *state* dapat disajikan pada persamaan berikut Bauwens, Preminger, & Rombouts (2010).

$$\begin{aligned} y_t &= \mu_{s_t} + \varepsilon_t \\ \varepsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= \alpha_{0,s_t} + \alpha_{1,s_t,} \varepsilon_{t-1}^2 + \beta_{1,s_t} \sigma_{t-1}^2 \end{aligned} \tag{2.10}$$

dimana:

y_t = data runtun waktu

s_t = variabel *state*.

μ_{S_t} = *mean* model pada *state* s_t

ε_t = residual pada waktu *t*

σ_t = volatilitas pada waktu *t*

u_t = residual berdistribusi $N(0,1)$

$\alpha_{0,s_t}, \alpha_{1,s_t}, \beta_{1,s_t}$ = parameter model GARCH pada *state* s_t

Untuk memastikan variansi positif maka terdapat beberapa fungsi batasan yaitu $\alpha_{0,s_t} > 0$, $\alpha_{1,s_t} \geq 0$, dan $\beta_{1,s_t} \geq 0$. Stasioneritas variansi pada setiap *state* didapatkan dengan syarat $\alpha_{1,s_t} + \beta_{1,s_t} < 1$ (Ardia, et al., 2018). Pada model *Markov switching* dengan dua *state*, nilai $s_t = 1$ atau $s_t = 2$. Variabel *state* yang tak teramati (s_t) merupakan realisasi dari *first order Markov chain* dengan probabilitas transisi konstan pada persamaan (2.11)

$$P(s_t = j \mid s_{t-1} = i) = p_{ij} \tag{2.11}$$

dengan $j = 1, 2$ dan $i = 1, 2$. Persamaan (2.11) dapat dinyatakan dalam matriks probabilitas transisi pada persamaan (2.12).

$$\begin{aligned} \mathbf{P} &= \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix} \\ &= \begin{pmatrix} p_{11} & 1 - p_{11} \\ 1 - p_{22} & p_{22} \end{pmatrix} \end{aligned} \tag{2.12}$$

dimana:

p_{11} = probabilitas proses tetap bertahan di *state* 1

p_{12} = probabilitas proses akan bertransisi dari *state* 1 ke *state* 2

p_{21} = probabilitas proses akan bertransisi dari *state* 2 ke *state* 1

p_{22} = probabilitas proses tetap bertahan di *state* 2

Menurut Hamilton (2005) s_t tidak bisa diamati secara langsung, tetapi inferensia mengenai s_t dapat dibuat melalui perilaku y_t yang teramati. Permasalahanya adalah *state* pada periode sebelumnya tidak diketahui sehingga diharuskan menggunakan algoritma iteratif dimana algoritma yang digunakan dalam penelitian ini adalah algoritma *Hamiltonian Monte Carlo* (HMC) dan *No-U Turn Sampler* (NUTS) yang akan dijelaskan pada subbab 2.4.3 dan 2.3.4.

Langkah awal yang diperlukan untuk melakukan iterasi ini adalah mendefinisikan kontribusi likelihood dari setiap observasi y_t pada masing masing *state* melalui persamaan (2.13).

$$\eta_{jt} = f\left(y_t \mid s_t = j;\boldsymbol{\theta}\right) = \frac{1}{\sqrt{2\pi\left(\alpha_{0j} + \alpha_{1j,}\varepsilon_{t-1}^2 + \beta_{1j}\sigma_{t-1}^2\right)}} e^{-\frac{\left(y_t-\mu_j\right)^2}{2\pi\left(\alpha_{0j}+\alpha_{1j,}\varepsilon_{t-1}^2+\beta_{1j}\sigma_{t-1}^2\right)}} \quad (2.13)$$

dengan $j = 1, 2$. dimana $\boldsymbol{\theta}$ adalah vektor parameter populasi, yaitu $\boldsymbol{\theta} = \left(p_{11}, p_{22}, \mu_1, \mu_2, \alpha_{01}, \alpha_{02}, \alpha_{11}, \alpha_{12}, \beta_{11}, \beta_{12}\right)'$. Dalam hal ini, probabilitas berada pada setiap *state* tidak diketahui, sehingga dimisalkan $\xi_{i,t-1}$ = probabilitas berada di *state i* pada periode *t*. Dengan demikian kontribusi likelihood dari observasi ke-*t* melalui persamaan (2.14)

$$\begin{aligned} f\left(y_t \mid \boldsymbol{\theta}\right) &= \sum_{i=1}^{2}\sum_{j=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt} \\ &= p_{11}\xi_{1,t-1}\eta_{1t} + p_{12}\xi_{1,t-1}\eta_{2t} + p_{21}\xi_{2,t-1}\eta_{1t} + p_{22}\xi_{2,t-1}\eta_{2t} \end{aligned} \quad (2.14)$$

sehingga probabilitas berada di *state* 1 dan 2 dapat dihitung melalui persamaan (2.15)

$$\xi_{jt} = \frac{\sum_{i=1}^{2} p_{ij}\xi_{i,t-1}\eta_{jt}}{f\left(y_t \mid \boldsymbol{\theta}\right)}. \quad (2.15)$$

Dengan demikian fungsi *likelihood* dan log-*likelihood* dapat didefinisikan melalui persamaan (2.16) dan persamaan (2.17).

$$\mathcal{L}\left(y \mid \boldsymbol{\theta}\right) = \prod_{t=1}^{T} f\left(y_t \mid \boldsymbol{\theta}\right) \quad (2.16)$$

$$\ln \mathcal{L}\left(y \mid \boldsymbol{\theta}\right) = \ln\left(\prod_{t=1}^{T} f\left(y_t \mid \boldsymbol{\theta}\right)\right)$$

$$\ln \mathcal{L}\left(y \mid \boldsymbol{\theta}\right) = \sum_{t=1}^{T} \ln f\left(y_t \mid \boldsymbol{\theta}\right) \quad (2.17)$$

2.4 Estimasi Parameter dengan Metode *Bayesian*

Dalam penelitian ini parameter model MSGARCH diestimasi menggunakan metode *Bayesian*. Berikut ini dijelaskan mengenai konsep *Bayesian*, distribusi prior dan estimasi parameter menggunakan metode *Hamiltonian Monte Carlo* dan *No-U-Turn Sampler* (NUTS).

2.4.1 Konsep *Bayesian*

Bayesian merupakan salah konsep dalam statistika yang memiliki pandangan berbeda dengan konsep statistika frekuentis. Pada statistika frekuentis, parameter dianggap sesuatu yang tetap (*fixed*) dan tidak diketahui, tetapi dalam *Bayesian*, parameter dianggap sebagai variabel random yang mempunyai distribusi yang biasa disebut distribusi prior.

Dalam hal inferensia, statistika klasik hanya melakukan inferensia berdasarkan sampel data yang diperoleh atau menggunakan fungsi *likelihood*. Sedangkan pada konsep *Bayesian*, selain menggunakan fungsi *likelihood*, juga mempertimbangkan informasi masa lalu tentang parameter distribusi prior, sehingga diperoleh distribusi posterior. Secara proporsional, persamaan distribusi posterior dapat dituliskan seperti persamaan (2.18)

$$f\left(\boldsymbol{\theta} \mid y\right) \propto f\left(y \mid \boldsymbol{\theta}\right) f\left(\boldsymbol{\theta}\right) \tag{2.18}$$

dimana $f\left(\boldsymbol{\theta}\right)$ adalah fungsi densitas prior dan $f\left(y \mid \boldsymbol{\theta}\right)$ adalah fungsi *likelihood*. Fungsi *likelihood* untuk pengamatan dari 1 sampai T dapat didefinisikan dalam persamaan (2.13).

Estimasi parameter dilakukan dengan menggunakan metode numerik. Metode yang paling populer dalam *Bayesian* adalah *Markov Chain Monte Carlo* (MCMC). MCMC adalah sebuah metode yang digunakan untuk membangkitkan variabel random yang didasarkan pada rantai *Markov*. Namun, penelitian ini menggunakan bahasa pemrograman Stan sehingga metode estimasi parameter yang digunakan adalah *Hamiltonian Monte Carlo* (HMC) dengan *No-U-Turn Sampler* (NUTS) yang akan dijelaskan pada subbab 2.4.3 dan 2.4.4.

2.4.2 Distribusi Prior

Pada pendekatan *Bayesian*, selain mempertimbangkan fungsi *likelihood*, penentuan distribusi prior merupakan tahapan yang sangat penting sehingga diperoleh distribusi posterior yang *proper*. Berikut adalah beberapa jenis prior yang digunakan dalam pendekatan *Bayesian*.

1. *Conjugate prior* atau *non-conjugate prior* (Gelman, et al., 2014), yaitu prior yang berkaitan atau bersesuaian dengan fungsi *likelihood*.

2. *Informative prior* atau *non-informative prior*, yaitu prior yang terkait dengan diketahui atau belum diketahuinya pola distribusi dari data.
3. *Proper prior* atau *improper prior*, yaitu prior yang berkaitan dengan pemberian bobot/densitas pada setiap titik apakah berdistribusi uniform atau tidak.
4. *Pseudo prior* adalah prior yang berkaitan dengan hasil elaborasi dengan pendekatan statistika frekuentis.

Pada penelitian sebelumnya, Bauwens, Preminger dan Rombouts (2010) menggunakan prior uniform dalam mengestimasi setiap parameter *Markov switching* GARCH. Ardia, et al (2018) menggunakan parameter *truncated normal* dalam mengestimasi parameter GARCH disesuaikan dengan fungsi batasan pada masing-masing parameter. Pada penelitian ini distribusi prior yang digunakan untuk mengestimasi parameter MSGARCH adalah seperti disajikan dalam Tabel 2.1.

Tabel 2.1 Distribusi Prior

Parameter	Distribusi Prior
$p_{11},\ p_{22}$	Uniform $(0,1)$
$\alpha_{01},\ \alpha_{02}$	Normal $\left(\mu_{\alpha_0}, \sigma_{\alpha_0}\right)$
$\alpha_{11},\ \alpha_{12}$	Uniform $\left(a_{\alpha_1}, b_{\alpha_1}\right)$
$\beta_{11},\ \beta_{12}$	Normal $\left(\mu_{\beta_1}, \sigma_{\beta_1}\right)$

2.4.3 *Hamiltonian Monte Carlo*

Hamiltonian Monte Carlo (HMC) adalah variasi dari algoritma *Metropolis Hasting* yang menggunakan dinamika *Hamiltonian* untuk membuat distribusi proposal (Betancourt, 2017). Dalam fisika, *Hamiltonian* (*H*) didefinisikan pada persamaan (2.19).

$$H\left(\boldsymbol{\theta}, \mathbf{r}\right) = U\left(\boldsymbol{\theta}\right) + K\left(\mathbf{r}\right) \tag{2.19}$$

dimana:

$U\left(\boldsymbol{\theta}\right)$ = energi potensial

$K\left(\mathbf{r}\right)$ = energi kinetik.

Dalam metode HMC, jika mengestimasi variabel random $\boldsymbol{\theta}$ dengan fungsi densitas $f(\boldsymbol{\theta} \mid y)$, maka didefinisikan sebuah variabel momentum $\mathbf{r}$ yang mengikuti distribusi normal:

$$f(\mathbf{r}) \sim N_d(0, \mathbf{I}) \tag{2.20}$$

dimana $\mathbf{I}$ adalah matriks identitas dan $N_d(0, \mathbf{I})$ adalah distribusi normal multivariat dengan *mean* 0 dan matriks kovariansi $\mathbf{I}$ dan q adalah ukuran vektor $\boldsymbol{\theta}$. *Joint density function* dari $f(\boldsymbol{\theta}, \mathbf{r})$ mengikuti persamaan (2.21).

$$\begin{aligned} f(\boldsymbol{\theta}, \mathbf{r}) &= \exp\{\log f(\boldsymbol{\theta} \mid y) + \log f(\mathbf{r})\} \\ &\propto \exp\left(\log f(\boldsymbol{\theta} \mid y) - \frac{1}{2}\mathbf{r}'\mathbf{r}\right) \end{aligned} \tag{2.21}$$

Dalam HMC, $U(\boldsymbol{\theta})$ dan $K(\mathbf{r})$ didefinisikan sebagai:

$$U(\boldsymbol{\theta}) = -\log f(\boldsymbol{\theta} \mid y) \tag{2.22}$$

$$K(\mathbf{r}) = \frac{1}{2}\mathbf{r}'\mathbf{r} \tag{2.23}$$

Jadi *joint density function* dari $f(\boldsymbol{\theta}, \mathbf{r})$ dapat ditulis seperti pada persamaan (2.24).

$$\begin{aligned} f(\boldsymbol{\theta}, \mathbf{r}) &= \exp\{-H(\boldsymbol{\theta}, \mathbf{r})\} \\ &= \exp\{-U(\boldsymbol{\theta}) - K(\mathbf{r})\} \\ &= \exp\left\{-\left(-\log f(\boldsymbol{\theta} \mid y)\right) - \frac{1}{2}\mathbf{r}'\mathbf{r}\right\} \\ &= f(\boldsymbol{\theta} \mid y)\exp\left\{-\frac{1}{2}\mathbf{r}'\mathbf{r}\right\} \end{aligned} \tag{2.24}$$

Tahapan dalam iterasi algoritma *Hamiltonian Monte Carlo* adalah sebagai berikut:

1. Menetapkan nilai inisial dari $\boldsymbol{\theta}^0, \mathbf{r}^0$ dan menentukan nilai π, L, dan M, serta menentukan fungsi log posterior $\log f(\boldsymbol{\theta} \mid y)$.
2. Inisial momentum $\mathbf{r}^0$ didapatkan dengan cara membangkitkan variabel random yang mengikuti distribusi normal standar yaitu $\mathbf{r}^0 \sim N(0, I_q)$ sebagai nilai inisial. Pembangkitan variabel random berdistribusi normal multivariat standard dengan variat sebanyak q sama dengan membangkitkan variabel

random berdistribusi normal univariat standard $r^0 \sim N(0,1)$ sebanyak q kali. Algoritma yang digunakan dalam membangkitkan variabel random $r^0 \sim N(0,1)$ adalah *slice sampling* dengan tahapan seperti pada Algoritma 2.1.

<u>Algoritma 2.1.</u> <u>Membangkitkan Variabel Random Berdistribusi Normal Standar dengan *Slice Sampling*</u>

1. Mendefinisikan fungsi kepadatan peluang dari distribusi normal standar $f(r) = \frac{1}{\sqrt{2(3,14)}} \mathrm{e}^{-\frac{1}{2}r^2}$.
2. Menentukan nilai awal r_0, dimana r_0 memenuhi $f(r_0) > 0$.
3. Melakukan pengambilan sampel nilai h yang berada pada interval $0 < h < f(r_0)$.
4. Menggambar garis horizontal yang memotong kurva $f(r)$ dan melewati titik h dengan batas kiri garis adalah $a^* = -\sqrt{-2\ln\left(h\sqrt{2(3,14)}\right)}$ dan batas kanan garis adalah $b^* = \sqrt{-2\ln\left(h\sqrt{2(3,14)}\right)}$.
5. Melakukan pengambilan sampel nilai (r,h) dari garis yang telah digambar pada tahap 5.
6. Menetapkan r sebagai nilai sampel yang baru.
7. Ulangi langkah 3 sampai langkah 5 menggunakan nilai r yang baru hingga mendapatkan sebanyak *M* sampel.

3. Memperbarui secara simultan parameter $\boldsymbol{\theta}$ dan momentum $\mathbf{r}$. Dalam proses ini terdapat algoritma *leapfrog* yang masing-masing melibatkan faktor π. Pada tahap *leapfrog*, nilai parameter $\boldsymbol{\theta}$ dan momentum $\mathbf{r}$ berubah. Tahapan *leapfrog* dilakukan dengan cara mengulangi langkah a. sampai c. berikut sebanyak *L* kali, dimana $\tau = 1, 2, \ldots, L$.

   a. Gunakan gradien dari log $f(\boldsymbol{\theta} \mid y)$ untuk membuat setengah tahap dari $\mathbf{r}$

$$\mathbf{r}^{\tau+\frac{\pi}{2}} \leftarrow \mathbf{r}^{\tau} + \frac{\pi}{2}\frac{d\log f\left(\boldsymbol{\theta}^{\tau} \mid y\right)}{d\boldsymbol{\theta}^{\tau}}.$$

b. Gunakan momentum $\mathbf{r}$ untuk memperbarui posisi dari vektor parameter $\boldsymbol{\theta}$

$$\boldsymbol{\theta}^{\tau+\pi} \leftarrow \boldsymbol{\theta}^{\tau} + \pi\mathbf{r}^{\tau+\frac{\pi}{2}}.$$

c. Sekali lagi gunakan gradien untuk memperbarui setengah bagian dari $\mathbf{r}$

$$\mathbf{r}^{\tau+\pi} \leftarrow \mathbf{r}^{\tau+\frac{\pi}{2}} + \frac{\pi}{2}\frac{d\log f\left(\boldsymbol{\theta}^{\tau+\pi} \mid y\right)}{d\boldsymbol{\theta}^{\tau+\pi}}.$$

d. Diperoleh nilai parameter $\boldsymbol{\theta}$ dan momentum $\mathbf{r}$ yang baru.

4. Untuk $m = 1, 2, ..., M$, beri label parameter $\boldsymbol{\theta}$ dan momentum $\mathbf{r}$ yang baru dengan $\boldsymbol{\theta}^{m}$ dan $\mathbf{r}^{m}$. Kriteria *accept* atau *reject* dari proposal didasarkan pada algoritma *Metropolis Hasting* dengan probabilitas penerimaan sebagai berikut.

$$\lambda = \min\left\{1, \frac{\exp\left(\log f\left(\boldsymbol{\theta}^{m} \mid y\right) - \frac{1}{2}\mathbf{r}^{m}\cdot\mathbf{r}^{m}\right)}{\exp\left(\log f\left(\boldsymbol{\theta}^{m-1} \mid y\right) - \frac{1}{2}\mathbf{r}^{0}\cdot\mathbf{r}^{0}\right)}\right\} \tag{2.25}$$

Penerimaan proposal dilakukan dengan cara membangkitkan variabel random $u_1 \sim uniform(0,1)$, kemudian jika $u_1 \le \lambda$, maka $\left(\boldsymbol{\theta}^{m}, \mathbf{r}^{m}\right)$ sebagai nilai baru parameter $\boldsymbol{\theta}$ dan momentum $\mathbf{r}$. Jika $u_1 > \lambda$, maka parameter $\boldsymbol{\theta}$ dan momentum $\mathbf{r}$ tidak mengalami pembaharuan yaitu tetap $\left(\boldsymbol{\theta}^{m-1}, \mathbf{r}^{0}\right)$.

5. Ulangi langkah 3 dan langkah 4 sampai *M* iterasi. Kriteria konvergen dari *M* sampel dinilai melalui nilai $\hat{R}$ yang mendekati 1 seperti yang akan dijelaskan pada subab 2.5.

Kelemahan dari metode HMC adalah efisiensi sampling sangat bergantung pada pengaturan dua parameter π dan *L* (Hoffman & Gelman, 2014). Di sisi lain, nilai π yang besar mengarah ke tingkat penerimaan yang rendah karena peningkatan kesalahan integrasi oleh integrasi *leapfrog*. Sebaliknya, jika π terlalu kecil, diperlukan waktu komputasi yang lama untuk mendapatkan panjang lintasan yang memadai. Disisi lain, banyaknya *L* mempengaruhi efisiensi pengambilan sampel, jika *L* terlalu kecil, sampel yang dihasilkan oleh HMC menunjukkan

autokorelasi yang cukup tinggi antara iterasi yang berurutan. Sebaliknya, jika L besar menyebabkan panjang lintasan yang besar yang dapat memindahkan parameter kembali ke keadaan semula.

2.4.4 *No-U-Turn Sampler*

Kelemahan dari HMC kemudian diperbaiki melalui *No-U-Turn Sampler* (NUTS). NUTS secara otomatis memilih nilai L yang sesuai pada setiap iterasi untuk memaksimalkan jarak di setiap *step leapfrog* dan menghindari perilaku *randomwalk*. Prosedur pengambilan sampel dengan NUTS adalah sebagai berikut (Hoffman & Gelman, 2014) (Nishio & Arakawa, 2019).

1. Menentukan nilai inisial untuk $\boldsymbol{\theta}^0, \pi^0$ dan nilai $\varsigma, \phi, \nu, w, \kappa$, M.
2. Membangkitkan variabel momentum $\mathbf{r}^0$ dari distribusi normal standar $\mathbf{r}^0 \sim N\left(0, \mathbf{I}_q\right)$, dimana q adalah ukuran vektor $\boldsymbol{\theta}$ atau banyaknya parameter yang diestimasi. Pembangkitan variabel random berdistribusi normal multivariat standar dengan variat sebanyak q sama dengan membangkitkan variabel random berdistribusi normal univariat standar $r^0 \sim N(0,1)$ sebanyak q kali. Algoritma yang digunakan dalam membangkitkan variabel random $r^0 \sim N(0,1)$ adalah *slice sampling* seperti pada Algoritma 2.1.
3. Membangkitkan *auxiliary variable* ι dari distribusi uniform. $\iota \sim \text{uniform}\left(0, \exp\left(\log f\left(\boldsymbol{\theta} \mid y\right) - \frac{1}{2}\mathbf{r}'\mathbf{r}\right)\right)$.
4. Membangkitkan C yaitu *subset* dari proposal $(\boldsymbol{\theta}, \mathbf{r})$. Di dalam NUTS, C dipilih melalui $(\boldsymbol{\theta}, \mathbf{r})$ menggunakan metode *doubling* dengan kernel transisi Ψ. Selanjutnya, dilakukan pengambilan sampel $\left(\boldsymbol{\theta}^m, \mathbf{r}^m\right)$ dari C. Penggabungan $\left(\boldsymbol{\theta}^m, \mathbf{r}^m\right)$ dengan C dinamakan C^{new}. Berikut adalah fungsi kernel transisi Ψ.

$$\Psi\left(\boldsymbol{\theta}^m, \mathbf{r}^m \mid \boldsymbol{\theta}, \mathbf{r}, C\right) = \begin{cases} \dfrac{I\left[\boldsymbol{\theta}^m, \mathbf{r}^m \in C^{new}\right]}{\left|C^{new}\right|} & ;\text{(a)} \\ \dfrac{\left|C^{new}\right|}{\left|C^{old}\right|} \dfrac{I\left[\boldsymbol{\theta}^m, \mathbf{r}^m \in C^{new}\right]}{\left|C^{new}\right|} + \left(1 - \dfrac{\left|C^{new}\right|}{\left|C^{old}\right|}\right) I\left[\left(\boldsymbol{\theta}^m, \mathbf{r}^m\right) = \left(\boldsymbol{\theta}, \mathbf{r}\right)\right] & ;(b) \end{cases} \tag{2.26}$$

(a) Ketika $\left|C^{new}\right| > \left|C^{old}\right|$

(b) Ketika $\left|C^{new}\right| \leq \left|C^{old}\right|$

Keterangan:

$I[\cdot]$ = bernilai 1 jika pernyataan yang ada di dalam kurung benar, dan bernilai 0 jika salah.

C = *subset* dari proposal $(\boldsymbol{\theta}, \mathbf{r})$.

C^{new} = *subset* dari proposal $(\boldsymbol{\theta}, \mathbf{r})$ yang didapatkan dengan menambahkan $(\boldsymbol{\theta}^*, \mathbf{r}^*)$ pada proses *doubling* terakhir.

C^{old} = *disjoint subset* dari C dimana $C = C^{new} \cup C^{old}$.

Ψ = kernel transisi dari C^{old} ke C^{new}.

5. Untuk $m = 1, 2, \ldots, M$, dilakukan proses penerimaan proposal $(\boldsymbol{\theta}^m, \mathbf{r}^m)$ dengan probabilitas λ^m. Penerimaan proposal dilakukan dengan cara membangkitkan variabel random $u_2 \sim uniform(0,1)$. Jika $u_2 \leq \lambda^m$, maka $(\boldsymbol{\theta}^m, \mathbf{r}^m)$ sebagai parameter baru. Jika $u_2 > \lambda^m$, maka parameter $(\boldsymbol{\theta}^{m-1}, \mathbf{r}^0)$ dipertahankan, dimana:

$$\lambda^m = \frac{1}{\left|F^m\right|} \sum_{\boldsymbol{\theta}, \mathbf{r} \in F^m} \min\left(1, \frac{P\left(\boldsymbol{\theta}^m, \mathbf{r}^m\right)}{P\left(\boldsymbol{\theta}^{m-1}, \mathbf{r}^0\right)}\right). \tag{2.27}$$

Keterangan:

$\boldsymbol{\theta}^m, \mathbf{r}^m$ = nilai proposal

$\boldsymbol{\theta}^{m-1}, \mathbf{r}^0$ = nilai inisial

F^m = himpunan semua *state* yang dieksplorasi selama *final doubling* untuk iterasi ke-*m* dari *Markov chain*.

6. Memperbarui parameter *tuning* π^m melalui *dual averaging* pada persamaan (2.28) sampai persamaan (2.30).

$$\log\left(\pi^{m+1}\right) \leftarrow \phi - \frac{\sqrt{m}}{\nu} \frac{1}{m+w} \sum_{m=1}^{M} \left(\varsigma - \lambda^m\right) \tag{2.28}$$

$$\log\left(\bar{\pi}^{m+1}\right) \leftarrow \omega^m \log\left(\pi^{m+1}\right) + \left(1 - \omega^m\right) \log\left(\bar{\pi}^m\right) \tag{2.29}$$

$$\pi^{m+1} \leftarrow \bar{\pi}^{m+1}. \tag{2.30}$$

Keterangan:

λ^m = probabilitas penerimaan aktual untuk iterasi ke-*m*

ς = probabilitas penerimaan yang diinginkan $(\varsigma \approx 0,6)$

ϕ = titik yang dipilih secara bebas yang menunjukkan arah penyusutan iterasi ke-*m* $\left(\phi = \log\left(10\pi_1\right)\right)$

ν = parameter bebas yang mengontrol jumlah penyusutan menuju ϕ

w = parameter bebas yang menghambat eksplorasi awal

$\omega^m = m^{-\kappa}$; $\kappa < 1$.

7. Ulangi langkah 2 sampai langkah 6 hingga *M* iterasi. Kriteria konvergen dari *M* sampel dinilai melalui nilai $\hat{R}$ yang mendekati 1 seperti yang akan dijelaskan pada subab 2.5.

2.5 Kriteria Konvergen

Berdasarkan Gelman (2014), konvergensi dapat didiagnosa dengan memeriksa *mixing* dan stasioneritas. Ada berbagai cara untuk melakukan *convergence diagnostic* diantaranya dengan menerapkan pendekatan yang cukup sederhana yaitu dengan membagi masing-masing *chain* menjadi dua dan memeriksa bahwa semua hasil berikutnya telah bercampur (*mixing*). Hal ini secara bersamaan menguji *mixing* (jika semua *chain* telah tercampur dengan baik, bagian-bagian dari *chain* yang berbeda juga harus bercampur) dan stasioneritas (bagian pertama dan kedua dari setiap urutan harus melintasi distribusi yang sama). Untuk setiap skalar pada vektor parameter $\boldsymbol{\theta}$ misalkan θ, simulasi dilabeli dengan θ_{na} dimana $n = 1, 2, \ldots, N;\ a = 1, 2, \ldots, A$. Kemudian dihitung :

$$B = \frac{N}{A-1}\sum_{a=1}^{A}\left(\bar{\theta}_{.a} - \bar{\theta}_{..}\right)^2 \tag{2.31}$$

dimana $\bar{\theta}_{.a} = \frac{1}{N}\sum_{n=1}^{N}\theta_{na}$, $\bar{\theta}_{..} = \frac{1}{A}\sum_{a=1}^{A}\bar{\theta}_{.a}$

$$W = \frac{1}{A}\sum_{a=1}^{A}\upsilon_a^2 \tag{2.32}$$

dimana $\upsilon_a^2 = \frac{1}{N-1}\sum_{n=1}^{N}\left(\theta_{na} - \bar{\theta}_{.a}\right)^2$

keterangan:

B = *between-sequence variance*

W = *within-sequence variance*

A = banyaknya *chain* (setelah dipisah akibat *warm-up*)

N = panjang setiap *chain*

Between-sequence variance (*B*) mengandung faktor N karena didasarkan pada variansi dari *within-sequence mean* $\left(\bar{\theta}_{.a}\right)$ yang masing-masing merupakan rata-rata dari θ_{na} sebanyak *N*.

$$\text{var}^{+}\left(\theta \mid y\right) = \frac{N-1}{N}W + \frac{1}{N}B \tag{2.33}$$

$$\hat{R} = \sqrt{\frac{\text{var}^{+}\left(\theta \mid y\right)}{W}}\,. \tag{2.34}$$

Berdasarkan Carpenter, et al (2017) estimasi parameter dikatakan telah mencapai konvergen ketika nilai $\hat{R}$ mendekati satu.

2.6 *Credible Interval*

Interval *Bayesian* dikenal sebagai *credible interval*. Diberikan data, maka *credible interval* $\left[b_b, b_a\right]$ $100\left(1-\alpha\right)\%$ untuk $\boldsymbol{\theta}$ memenuhi persamaan (2.35).

$$P\left(\boldsymbol{\theta} \in \left[b_b, b_a\right] \mid y\right) = \int_{b_b}^{b_a} f\left(\boldsymbol{\theta} \mid y\right) = 1-\alpha \tag{2.35}$$

Pada konsep *credible interval*, $\boldsymbol{\theta}$ merupakan variabel random dan interval adalah variabel *fixed*. *Credible interval* dapat digunakan sebagai acuan penentuan signifikansi parameter. Parameter $\boldsymbol{\theta}$ dikatakan signifikan jika berada pada batas interval $\left[b_b, b_a\right]100\left(1-\alpha\right)\%$ (King, et al., 2010).

2.7 *Watanabe Akaike Information Criterion*

Setelah didapatkan model *Bayesian*, seringkali ingin diukur akurasi prediktifnya untuk tujuam perbandingan atau memilihan model terbaik. Akurasi prediktif dapat diukur melalui validasi silang (*cross validation*) dan kriteria

informasi (*information criterion*). *Watanabe-Akaike Information Criterion* atau *Widely Applicable Information Criterion* (WAIC) merupakan perbaikan dari *Deviance Information Criterion* (DIC) untuk model *Bayesian*. DIC sangat populer dalam mengukur kebaikan model *Bayesian* yaitu pada penerapannya di *package* BUGS, namun DIC memiliki beberapa permasalahan yaitu tidak semua *fully Bayesian* berdasarkan estimasi titik. Sebagai contoh, DIC dapat menghasilkan estimasi negatif dari banyaknya parameter yang efektif di dalam sebuah model dan tidak didefinisikan sebagai model singular. WAIC adalah *fully Bayesian* dan mendekati pendekatan *Bayesian cross-validation* (Vehtari, et al., 2017).

Misalkan data $y_1, y_2, \ldots, y_T$ dimodelkan dengan melibatkan parameter $\boldsymbol{\theta}$, sehingga:

$$\begin{aligned}\text{Fungsi } \textit{likelihood} &= f\left(y \mid \boldsymbol{\theta}\right) \\ &= \prod_{t=1}^{T} f\left(y_t \mid \boldsymbol{\theta}\right)\end{aligned}$$

$$\text{Distribusi prior} = f\left(\boldsymbol{\theta}\right)$$

$$\begin{aligned}\text{Distribusi posterior} &= f_{post}\left(\boldsymbol{\theta}\right) \\ &= f\left(\boldsymbol{\theta} \mid y\right)\end{aligned}$$

$$\begin{aligned}\text{Distribusi posterior prediktif} &= f_{post}\left(\tilde{y}\right) \\ &= \int f\left(\tilde{y}_i \mid \boldsymbol{\theta}\right) f_{post}\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta}\end{aligned}.$$

Ukuran akurasi prediktif untuk T titik data diberikan oleh *expected log pointwise predictive density* (elpd) pada persamaan (2.36).

$$\text{elpd} = \sum_{t=1}^{T} E_{f_t}\left(\log f_{post}\left(\tilde{y}_t\right)\right) \tag{2.36}$$

dimana $f_t\left(y\right)$ adalah distribusi yang merepresentasikan proses pembangkitan data asli daripada y_t. $f_t\left(y\right)$ bergantung pada $\boldsymbol{\theta}$, sehingga tidak diketahui dan kemudian digunakan WAIC untuk mendekatinya. Sebuah persamaan yang sangat membantu dalam analisis ini adalah *log pointwise predictive density* (lpd) yang disajikan pada persamaan (2.37).

$$\begin{aligned}\text{lpd} &= \sum_{t=1}^{T} \log f_{post}\left(y_t\right) \\ &= \sum_{t=1}^{T} \log \int f\left(y_t \mid \boldsymbol{\theta}\right) f_{post}\left(\boldsymbol{\theta}\right) d\boldsymbol{\theta}\end{aligned} \tag{2.37}$$

Untuk menghitung lpd, dapat dilakukan dengan mengevalusi ekspektasi menggunakan *draw* dari $f_{post}\left(\boldsymbol{\theta}\right)$, yaitu simulasi posterior pada umumnya yang diberi label $\boldsymbol{\theta}^m$ dimana $m = 1, 2, \ldots, M$, sehingga didapatkan *computed log pointwise predictive density* ($\widehat{\text{lpd}}$) pada persamaan (2.38).

$$\widehat{\text{lpd}} = \sum_{t=1}^{T} \log\left(\frac{1}{M}\sum_{m=1}^{M} f\left(y_t \mid \boldsymbol{\theta}^m\right)\right) \tag{2.38}$$

WAIC dapat diinterpretasikan sebagai pendekatan komputasi yang mudah untuk akurasi prediktif dan didefinisikan berdasarkan persamaan (2.39).

$$\widehat{\text{elpd}}_{\text{waic}} = \widehat{\text{lpd}} - \hat{p}_{waic} \tag{2.39}$$

dimana $\hat{p}_{waic}$ adalah estimasi banyaknya parameter efektif yang diperoleh melalui persamaan (2.40) berikut.

$$\hat{p}_{waic} = \sum_{t=1}^{T} \text{var}\left(\log f\left(y_t \mid \boldsymbol{\theta}^m\right)\right) \tag{2.40}$$

Dengan demikian WAIC diperoleh melalui persamaan (2.41) berikut.

$$\text{WAIC} = -2\,\widehat{\text{elpd}}_{\text{waic}} \tag{2.41}$$

2.8 *Indonesian Crude Price*

Harga minyak mentah Indonesia atau *Indonesian Crude Price* (ICP) adalah harga rata-rata 50 jenis minyak mentah Indonesia di pasar dunia yang ditetapkan setiap bulan. ICP digunakan sebagai asumsi harga minyak dalam APBN untuk menentukan besaran penerimaan migas, subsidi energi, dan dana bagi hasil minyak. Berdasarkan Keputusan Menteri Energi dan Sumber Daya Material (ESDM) No. 2704K/12/MEM/2013 tentang Penetapan Formula Harga Minyak Mentah Indonesia, terdapat 52 jenis minyak mentah Indonesia yang masing-masing mempunyai harga yang berbeda yang terbagi dalam dua kelompok yaitu minyak mentah utama Indonesia (*benchmark crude*) dan minyak mentah Indonesia lainnya. Ada delapan jenis minyak mentah utama Indonesia (SLC, Arjuna, Attaka, Cinta,

Duri, Widuri, Belida, dan Senipah *Condensate*) dengan formula harga yang mengacu pada publikasi yang diterbitkan oleh lembaga independen internasional RIM dan PLATT'S dengan formula sebagai berikut (Handoko, 2014).

$$ICP = 50\% \text{ RIM} + 50\% \text{ PLATT'S}$$

dimana:

RIM = Badan Independen berpusat di Tokyo dan Singapura yang menyediakan data harga minyak untuk pasar Asia Pasifik di Timur Tengah.

PLATT'S = Penyedia jasa informasi energi yang berpusat di Singapura.

(Halaman ini sengaja dikosongkan)

BAB III
METODOLOGI PENELITIAN

3.1 Sumber Data dan Variabel Penelitian

Data yang digunakan dalam penelitian ini adalah rata-rata bulanan harga minyak mentah Indonesia atau Indonesian Crude Price (ICP) mulai Januari 1989 sampai dengan Desember 2019 (372 data). Data ini adalah data yang bersifat publik yang sumbernya dari catatan Direktorat Jenderal Minyak Gas dan Bumi, Kementrian Energi dan Sumber Daya Mineral Republik Indonesia. Berikut adalah *link* yang digunakan untuk mengunduh data ICP bulanan:
https://migas.esdm.go.id/post/category/publikasi/informasiberkala/hargaminyakmentah

Variabel yang digunakan dalam penelitian ini adalah rata-rata harga minyak mentah Indonesia atau *Indonesian Crude Price* (ICP) dengan struktur data seperti dalam Tabel 3.1.

Tabel 3.1 Struktur Data Penelitian

Tahun	**Bulan**	t	z_t	y_t
1989	Januari	1	z_1	y_1
1989	Maret	2	z_2	y_2
1989	April	3	z_3	y_3
1989	Mei	4	z_4	y_4
⋮	⋮	⋮	⋮	⋮
2019	Desember	372	z_{372}	y_{372}

Keterangan:

t = periode

z_t = rata-rata harga minyak mentah Indonesia pada periode ke-*t*

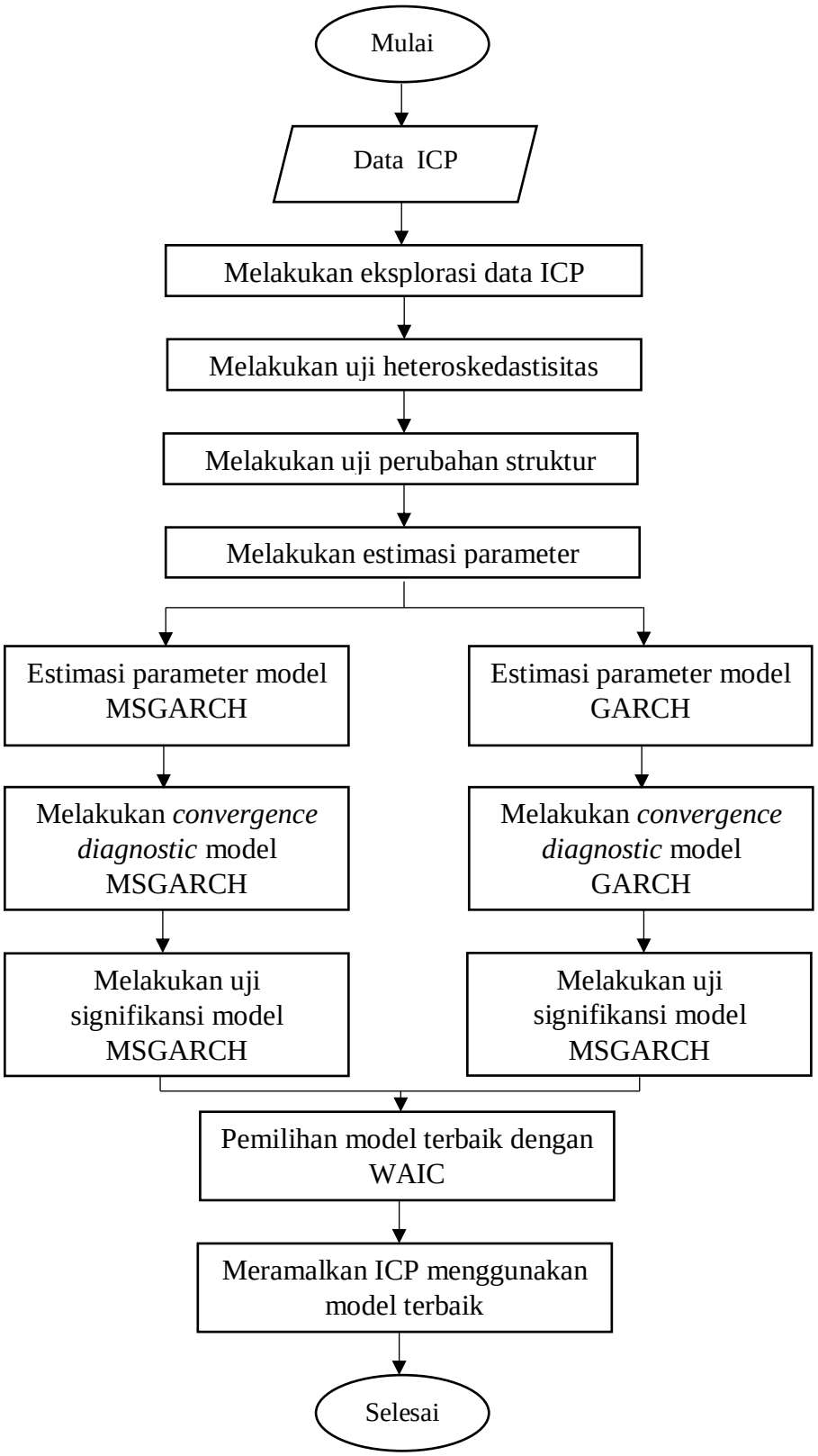
y_t = *differencing lag* 1 dari z_t .

3.2 Tahapan Analisis Data

Berikut ini akan disajikan tahapan analisis data yang akan dilakukan pada penelitian ini.

1. Melakukan eksplorasi data ICP untuk mendapatkan karakteristik data ICP.
   a. Mempersiapkan data, yaitu data *first differencing* ICP.
   b. Membuat plot *time series* pada data.
   c. Membuat histogram untuk mengetahui distribusi data.
   d. Menghitung statistika deskriptif.
   e. Melakukan uji stasioneritas data menggunakan uji *Dickey-Fuller* seperti pada persamaan (2.2).
   f. Melakukan uji heteroskedastisitas data menggunakan uji *Lagrange Multiplier* seperti pada persamaan (2.5).
   g. Melakukan uji perubahan struktur menggunakan uji ICSS seperti pada persamaan (2.8)
2. Melakukan pemodelan ICP.
   a. Pemodelan ICP dengan MSGARCH
      i) Melakukan uji *convergence diagnostic* pada parameter model MSGARCH seperti pada persamaan (2.34).
      ii) Menguji signifikansi parameter MSGARCH menggunakan *credible interval* seperti pada persamaan (2.35).
      iii) Memodelkan ICP dengan MSGARCH.
   b. Pemodelan ICP dengan GARCH
      i) Melakukan uji *convergence diagnostic* pada parameter model GARCH seperti pada persamaan (2.34)
      ii) Menguji signifikansi parameter GARCH menggunakan *credible interval* seperti pada persamaan (2.35).
      iii) Memodelkan ICP dengan GARCH.
   c. Memilih model terbaik menggunakan *Watanabe-Akaike Information Criterion* (WAIC) seperti pada persamaan (2.41).
3. Melakukan peramalan ICP menggunakan model terbaik.

Langkah analisis pada subbab 4.3 dapat disajikan dalam diagram alir pada Gambar 3.1.



Gambar 3.1. Diagram Alir Penelitian

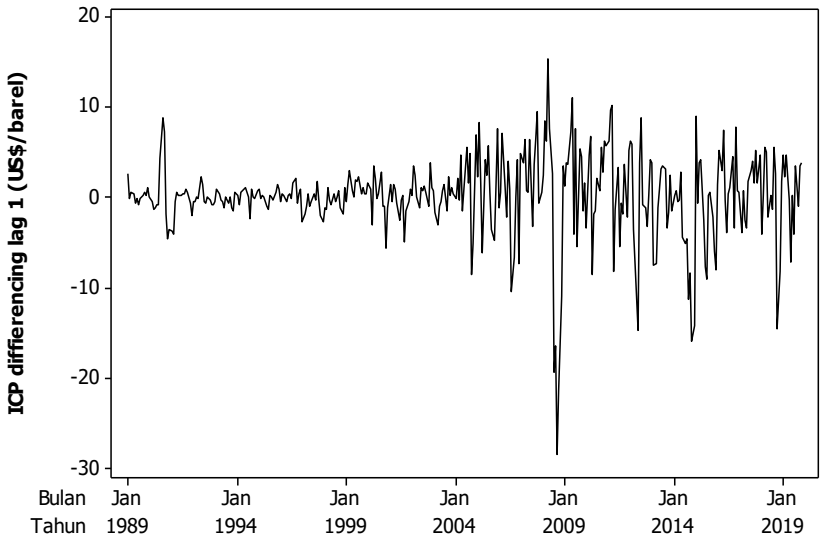
(Halaman ini sengaja dikosongkan)

BAB IV
ANALISIS DAN PEMBAHASAN

Bab ini membahas mengenai hasil analisis dari penelitian ini. Subbab pertama membahas karakteristik data ICP. Selanjutnya pada subbab kedua membahas mengenai pemodelan data ICP yang terdiri dari pemodelan dengan MSGARCH, pemodelan dengan GARCH dan pemilihan model terbaik. Subbab terakhir yaitu subbab ketiga membahas mengenai peramalan ICP.

4.1 Karakteristik Data ICP

Sebelum dilakukan pemodelan terhadap data ICP, langkah awal yang perlu dilakukan adalah dengan mengidentifikasi karakteristik data melalui plot dan statistika deskriptif. Pada penelitian ini data terlebih dahulu dilakukan *differencing* pada *lag* 1, sehingga data ICP yang baru adalah data yang telah dilakukan *differencing* pada *lag* 1 yang selanjutnya disebut sebagai ICP (y_t). Identifikasi karakteristik data melalui plot dapat dilakukan dengan membentuk plot time series pada Gambar 4.1.

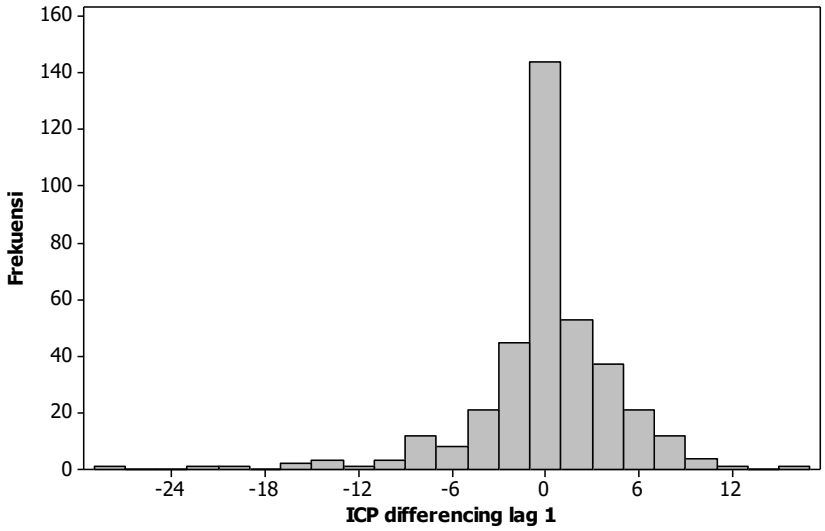


Gambar 4.1. *Time Series* Plot ICP yang di-*differencing* pada *lag* 1

Plot *time series* data ICP pada Gambar 4.1 menunjukkan adanya fluktuasi yang tidak menentu dari waktu ke waktu. Berdasarkan hasil uji Dickey Fuller pada Lampiran 2a, nilai statistik uji Dickey Fuller sebesar -7,82 dengan *p-value* kurang

dari 0,05. Hal ini menunjukkan bahwa data ICP yang telah di-*differencing* pada *lag* 1 memenuhi asumsi stasioneritas dalam *mean*. Namun dapat dilihat bahwa fluktuasi dari waktu ke waktu tidak konstan, terutama volatilitasnya. Pada periode tertentu data berfluktuasi di dekat *mean*, namun ada beberapa periode lain dimana fluktuasinya cukup tajam. Kondisi tersebut menunjukkan bahwa data telah stasioner dalam *mean* namun ada indikasi tidak stasioner dalam variansi.

Untuk mengetahui pola distribusi dari *differencing* ICP, maka langkah selanjutnya adalah membuat histogram. Histogram dari ICP disajikan pada Gambar 4.2.



Gambar 4.2. Histogram ICP yang di-*differencing* pada *lag* 1

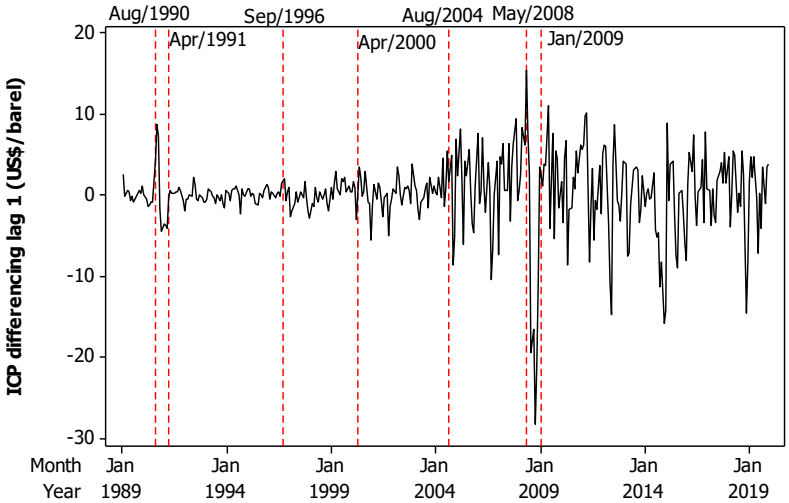
Histogram pada Gambar 4.2 menunjukkan pola leptokurtik dan berekor tebal. Hal ini sesuai dengan hasil statistik deskriptif pada data ICP pada Tabel 4.1. Pada Tabel 1 dapat dilihat bahwa rata-rata ICP adalah sebesar 0,14 dengan standar deviasi 4,58. Nilai *skewness* sebesar -1.51 menunjukkan bahwa distribusi data menceng ke kiri. Nilai kurtosis sebesar 6,99 menunjukkan keruncingan dari pola distribusi. Ketika nilai kurtosis melebihi 3, maka pola distribusi tersebut disebut leptokurtik. Hal ini menunjukkan bahwa pemodelan dengan uni modal tidak cukup mewakili. Hal tersebut juga merupakan salah satu alasan digunakannya model *Markov switching* dalam memodelkan ICP.

Tabel 4.1 Statistik Deskriptif ICP

Variabel	Nilai
Mean	0,14
Deviasi standar	4,58
Skewness	1,51
Kurtosis	6,99

Sebelum menerapkan model MSGARCH pada data ICP terlebih dahulu dilakukan pendeteksian adanya heteroskedastisitas dalam data ICP. Pendeteksian dilakukan dengan metode *Lagrange Multiplier* sebagaimana dijelaskan pada subbab 2.2.2 sebelumnya. Berdasarkan uji efek heteroskedastisitas pada Lampiran 2b, diketahui bahwa statistik uji *LM* sampai *lag*-12 untuk data ICP menghasilkan *p-value* kurang dari 0,05. Dengan demikian dapat disimpulkan bahwa data ICP mengandung heteroskedastisitas.

Selain uji heteroskedastisitas, pengujian perubahan struktur perlu dilakukan untuk mengetahui ada tidaknya perubahan struktur pada data ICP. Uji perubahan struktur dilakukan dengan uji ICSS. Berdasarkan uji ICSS pada Lampiran 2c, diketahui bahwa terdapat tujuh *break point*. Titik-titik terjadinya *break point* dapat dilihat melalui Gambar 4.3.



Gambar 4.3. *Structural Break Point* ICP

Pada Gambar 4.3 dapat dilihat bahwa perubahan struktur terjadi pada tujuh titik. Titik-titik tersebut adalah pada bulan Agustus 1990, April 1991, September 1996, April 2000, Agustus 2004, Mei 2008 dan Januari 2009.

Adanya efek heteroskedastisitas pada data dapat ditangkap melalui model GARCH dan terjadinya perubahan struktur pada data dapat ditangkap melalui model *Markov switching*. Dengan demikian pada penelitian ini diterapkan gabungan antara model GARCH dan *Markov switching* yang disebut sebagai model *Markov switching* GARCH.

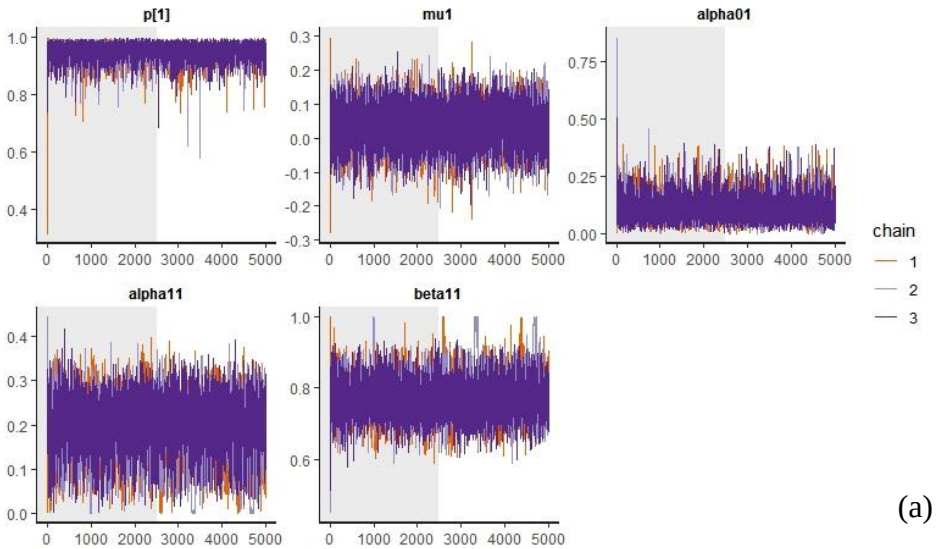
4.2 Pemodelan ICP

Estimasi parameter dalam penelitian ini menggunakan metode *Bayesian*. Bahasa pemrograman yang digunakan adalah Stan dimana di dalamnya menerapkan algoritma Hamiltonian Monte Carlo (HMC) dengan No-U Turn Sample (NUTS). Pemrograman Stan dapat dilakukan melalui *software* R dengan package "rstan". Pada sub bab estimasi parameter dibahas mengenai hasil estimasi parameter model MSGARCH dan GARCH.

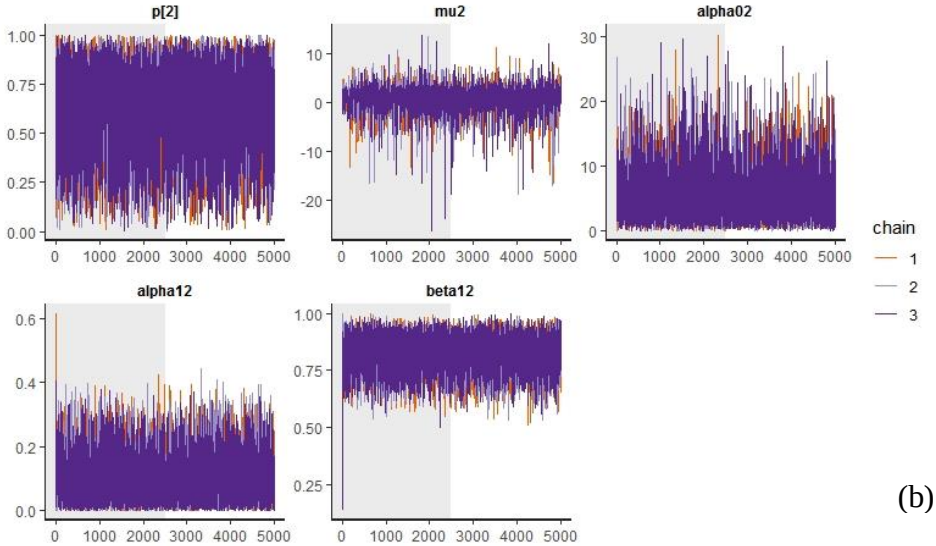
4.2.1 Pemodelan ICP dengan MSGARCH

Estimasi parameter dalam *Bayesian* dilakukan dengan metode iteratif, oleh karena itu sebelum menentukan signifikansi parameter dalam model, hal pertama yang harus dilakukan adalah melakukan *convergence diagnostic* untuk mengetahui apakah iterasi telah konvergen untuk setiap parameter. *Convergence diagnostic* dapat diketahui melalui beberapa indikator yaitu melalui plot simulasi historis dan $\hat{R}$.

Secara visual, *convergence diagnostic* dapat dilihat melalui plot simulasi historis dari posterior. plot simulasi historis dari parameter MS-GARCH dapat dilihat melalui Gambar 4.4 dan Gambar 4.5.



(a)



(b)

Gambar 4.5. Simulasi Historis Parameter MSGARCH: (a) *State* 1 (b) *State* 2.

Gambar 4.4 dan Gambar 4.5 menunjukkan simulasi historis dari parameter MSGARCH dengan 5000 iterasi. Plot dengan latar belakang berwarna abu-abu menunjukkan *warm-up* atau yang biasa disebut *burn-in*, yang berarti estimasi parameter didapatkan setelah kondisi *burn-in*. *Warmup* yang digunakan pada estimasi ini adalah setengah dari jumlah iterasi total, yakni 2500. Gambar 4.4 dan Gambar 4.5 yang terarsir menunjukkan rentang periode pembangkitan parameter model hingga *burn*-in. Simulasi historis dengan tiga *chain* yang berbeda pada Gambar 4.4 dan Gambar 4.5 menunjukkan bahwa proses estimasi parameter

MSGARCH baik pada *state* 1 maupun *state* 2 cenderung mencapai kondisi konvergen karena plot berfluktuasi di sekitar *mean* yang sama.

Selain melalui plot simulasi historis, *convergence diagnostic* dapat dilihat melalui nilai $\hat{R}$. Nilai dari $\hat{R}$ masing-masing dapat dilihat pada Tabel 4.2. Tabel 4.2 menunjukkan bahwa semua parameter MSGARCH mempunyai nilai $\hat{R}$ yang mendekati satu. Berdasarkan indikator $\hat{R}$, dapat disimpulkan bahwa proses estimasi parameter MSGARCH telah mencapai kondisi konvergen.

Tabel 4.2. *Convergence Diagnostic* Parameter MSGARCH

Parameter	$\hat{R}$
p_{11}	0,9999
p_{22}	1,0051
μ_1	0,9999
α_{01}	1,0019
α_{11}	1,0085
β_{11}	1,0087
μ_2	1,0027
α_{02}	1,0009
α_{12}	1,0037
β_{12}	1,0034

Setelah dilakukan *convergence diagnostic* dan didapatkan kesimpulan bahwa iterasi telah konvergen, maka selanjutnya adalah mendapatkan *mean posterior* dari setiap parameter beserta *credible interval*-nya. Hasil estimasi parameter model MSGARCH disajikan dalam Tabel 4.3.

Tabel 4.3. Estimasi Parameter MSGARCH

Parameter	***Mean***	**Deviasi standar**	**2,5%**	**97,5%**
p_{11}	0,9564	0,0286	0,8908	0,9913
p_{22}	0,6319	0,2203	0,1649	0,9720

Tabel 4.3. Estimasi Parameter MSGARCH (lanjutan)

Parameter	*Mean*	Deviasi standar	2,5%	97,5%
μ_1	0,0336	0,0608	-0,0853	0,1527
α_{01}	0,1045	0,0549	0,0176	0,2368
α_{11}	0,1880	0,0708	0,0034	0,3079
β_{11}	0,7836	0,0630	0,6727	0,9812
μ_2	0,9810	2,4363	-5,3197	4,3123
α_{02}	4,9431	3,6946	0,6181	14,4552
α_{12}	0,0997	0,0811	0,0026	0,2892
β_{12}	0,8236	0,0786	0,6529	0,9525

Kolom ke-dua pada Tabel 4.3 merupakan hasil *mean* posterior atau hasil estimasi titik dengan standard deviasi ditunjukkan pada kolom ke-tiga, sedangkan estimasi interval ditunjukkan oleh kolom *credible interval.* Pengujian signifikansi parameter juga dapat dilihat melalui kolom *credible interval*. Parameter μ tidak signifikan karena *credible interval* memuat angka nol. Sedangkan parameter lainnya signifikan karena *credible interval* tidak memuat angka nol.

Berdasarkan Tabel 4.3 maka persamaan untuk model MSGARCH dapat ditulis pada persamaan (4.1) dan (4.2).

Persamaan untuk *state* 1:

$$\begin{aligned} y_{1,t} &= 0,0336 + \varepsilon_{1,t} \\ \varepsilon_{1,t} &= \sigma_{1,t} u_t \\ \sigma_{1,t}^2 &= 0,1045 + 0,1880\varepsilon_{1,t-1}^2 + 0,7836\sigma_{1,t-1}^2 \end{aligned} \tag{4.1}$$

Persamaan untuk *state* 2:

$$\begin{aligned} y_{2,t} &= 0,9810 + \varepsilon_{2,t} \\ \varepsilon_{2,t} &= \sigma_{2,t} u_t \\ \sigma_{2,t}^2 &= 4,9431 + 0,0997\varepsilon_{2,t-1}^2 + 0,8236\sigma_{2,t-1}^2 \end{aligned} \tag{4.2}$$

Berdasarkan persamaan (4.1) dan persamaan (4.2) dapat dilihat bahwa persamaan (4.1) mewakili *state* dengan volatilitas rendah dan persamaan (4.2) mewakili *state* dengan volatilitas tinggi. Antara persamaan (4.1) dan persamaan (4.2) dihubungkan oleh sebuah matriks probabilitas transisi yang menunjukkan

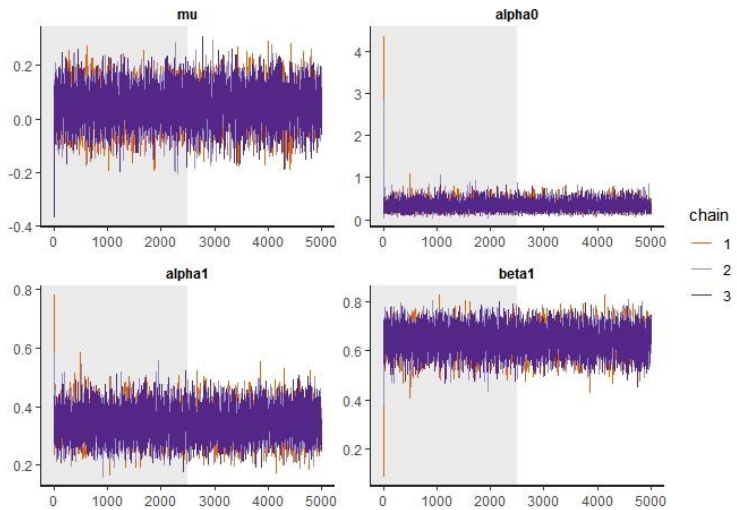
probabilitas suatu *state* akan bertransisi ke *state* yang lain atau bertahan pada *state* yang sama. Matriks probabilitas transisi ditunjukkan oleh persamaan (4.3).

$$\mathbf{P} = \begin{pmatrix} 0,9564 & 1-0,9654 \\ 1-0,6319 & 0,6319 \end{pmatrix} = \begin{pmatrix} 0,9564 & 0,0436 \\ 0,3681 & 0,6319 \end{pmatrix} \tag{4.3}$$

Berdasarkan matriks transisi pada Persamaan 4.3 diketahui bahwa peluang bertransisi dari volatilitas rendah ke volatilitas tinggi sangat kecil yaitu sebesar 4%, sedangkan ketika berada pada volatilitas tinggi, maka akan segera berpindah ke volatilitas rendah dengan probabilitas sebesar 37%. Probabilitas data selalu berada pada volatilitas rendah adalah sebesar 96% sedangkan berada pada volatilitas tinggi dengan probabilitas 63%. Dengan demikian data akan lebih sering berada di volatilitas rendah daripada berada di volatilitas tinggi. Hasil prediksi *state* dapat dilihat pada Lampiran 4.

4.2.2 Pemodelan ICP dengan GARCH

Sebagai pembanding, dalam penelitian ini juga dilakukan pemodelan dengan metode *single regime* GARCH. Metode estimasi parameter yang digunakan sama seperti sebelumnya yaitu dengan metode *Bayesian*, sehingga perlu dilakukan *convergence diagnostic* terlebih dahulu untuk memastikan bahwa iterasi telah konvergen. Plot simulasi historis dari parameter GARCH dapat dilihat melalui Gambar 4.6.



Gambar 4.6. Simulasi Historis Parameter GARCH

Gambar 4.6 yang terarsir menunjukkan rentang periode pembangkitan parameter model hingga *burn*-in. Simulasi historis dengan tiga *chain* yang berbeda pada Gambar 4.6 menunjukkan bahwa proses estimasi parameter GARCH baik cenderung mencapai kondisi konvergen karena plot berfluktuasi di sekitar *mean* yang sama. Selain melalui plot simulasi historis, juga perlu dilihat nilai $\hat{R}$. Nilai $\hat{R}$ dari parameter GARCH disajikan pada Tabel 4.4.

Tabel 4.4. *Convergence Diagnostic* GARCH

Parameter	$\hat{R}$
μ	0,9997
α_0	1,0006
α_1	1,0002
β_1	1,0004

Berdasarkan Tabel 4.4 dapat dilihat bahwa semua parameter GARCH memiliki nilai $\hat{R}$ yang mendekati satu. Dengan demikian, jika dievaluasi melalui kriteria $\hat{R}$ estimasi parameter GARCH telah mencapai konvergen. Setelah diketahui bahwa semua parameter telah konvergen, maka dapat diuji apakah parameter yang dihasilkan signifikan.. Hasil estimasi parameter model GARCH disajikan dalam Tabel 4.5.

Tabel 4.5 Estimasi Parameter GARCH

Parameter	***Mean***	**Standard Deviasi**	**2,5%**	**97,5%**
μ	0,0448	0,0682	-0,0918	0,1760
α_0	0,3318	0,1151	0,1506	0,59`0
α_1	0,3397	0,0526	0,2477	0,4470
β_1	0,6477	0,0527	0,5402	0,7462

Pengujian signifikansi parameter dapat dilihat melalui kolom *credible interval*. Parameter μ tidak signifikan karena *credible interval* melalui nol.

Sedangkan parameter lainnya yaitu α_0 , α_1 , dan β_1 signifikan karena *credible interval* tidak melalui nol.

Model GARCH yang dihasilkan dapat dituliskan dalam persamaan (4.4):

$$\begin{aligned} y_t &= 0,0448 + \varepsilon_t \\ \varepsilon_t &= \sigma_t u_t \\ \sigma_t^2 &= 0,3318 + 0,3397\varepsilon_{t-1}^2 + 0,6477\sigma_{t-1}^2 \end{aligned} \tag{4.4}$$

4.2.3 Pemilihan Model Terbaik

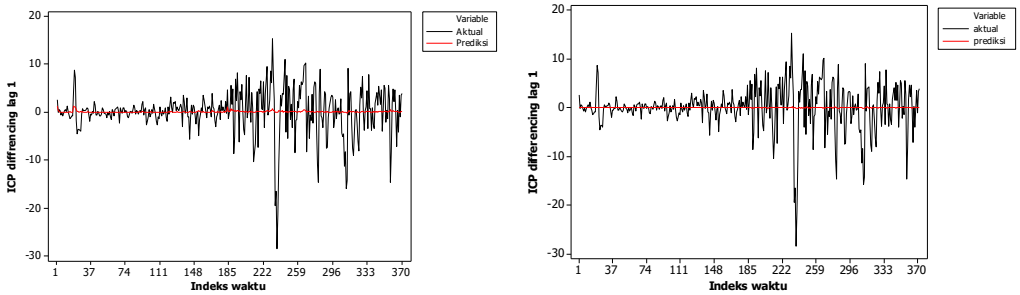
Tahap terakhir dalam pemodelan adalah menguji kebaikan model. Pada penelitian ini model yang dibandingkan adalah model GARCH dan MS-GARCH. Kriteria kebaikan model yang digunakan adalah WAIC dimana model dengan WAIC terkecil merupakan model terbaik. Berikut adalah Tabel 4.6 yang berisi nilai WAIC dari model GARCH dan model MS-GARCH.

Tabel 4.6 Perbandingan Kebaikan Model MSGARCH dan GARCH

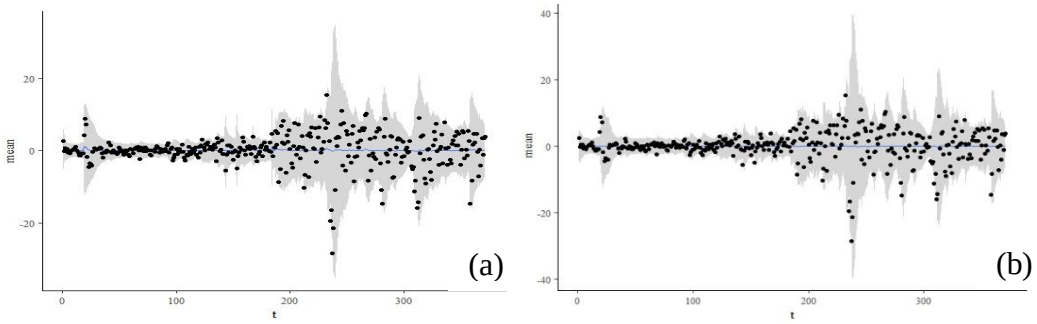
Model	WAIC
MSGARCH	1869,42
GARCH	1876,27

Nilai WAIC dari model MSGARCH lebih kecil daripada model pembandingnya yaitu GARCH. Dengan demikian dapat disimpulkan bahwa model MSGARCH memiliki kinerja yang lebih baik daripada model GARCH dalam memodelkan data ICP.

Secara visual, hasil prediksi ICP menggunakan model MSGARCH dan GARCH masing-masing dapat dilihat melalui Gambar 4.7 dan Gambar 4.8.



Gambar 4.7 Prediksi *Mean* ICP : (a) MSGARCH (b) GARCH



Gambar 4.8 Prediksi *Mean* dan Variansi ICP : (a) MSGARCH (b) GARCH

Gambar 4.7 menunjukkan plot data aktual dan prediksi dari model *mean* MSGARCH dan GARCH, dimana garis hitam menunjukkan data aktual dan garis merah menunjukkan hasil prediksi. Berdasarkan Gambar 4.8, titik hitam menunjukkan *differencing* ICP, garis biru menunjukkan *mean* dari prediksi ICP, dan warna abu-abu menunjukkan variansi dari prediksi ICP.

4.3 Peramalan ICP

Berdasarkan penentuan model terbaik menggunakan WAIC, model terbaik yang dihasilkan adalah model MSGARCH dengan demikian peramalan dilakukan menggunakan model MSGARCH. Peramalan pada penelitian ini adalah *one step ahead prediction* atau prediksi satu langkah ke depan (Schnatter, 2006). Pada MSGARCH, peramalan dilakukan melalui dua tahap yaitu peramalan *state* dan peramalan volatilitas. Algoritma peramalan menggunakan model MSGARCH disajikan dalam Algoritma 4.1.

Algoritma 4.1. Peramalan dengan Model MSGARCH

1. Menentukan banyak *state*.
2. Mendapatkan model pada setiap *state* dan matriks probabilitas transisinya.
3. Mendapatkan label *state* pada data periode ke T.
4. Menggunakan matriks probabilitas transisi untuk menghitung probabilitas terjadinya *state* tertentu pada periode $T+1$.
5. Mendapatkan *state* yang akan terjadi pada periode $T+1$.
6. Mendapatkan ramalan *mean* dan volatilitas menggunakan model yang bersesuaian dengan hasil ramalan *state* pada tahap 5.

Berdasarkan Algoritma 4.1, pada penelitian ini telah ditetapkan banyak *state* adalah dua *state*. Model yang didapatkan untuk setiap *state* telah disajikan pada persamaan (4.1) dan persamaan (4.2) dan matriks probabilitas transisi disajikan pada persamaan (4.3). Seperti yang telah dijelaskan sebelumnya bahwa persamaan (4.1) merupakan persamaan volatilitas rendah dan persamaan (4.2) merupakan persamaan dengan volatilitas tinggi serta kedua model tersebut dihubungkan oleh sebuah matriks probabilitas transisi pada persamaan (4.3).

Dari hasil pemodelan MSGARCH, didapatkan bahwa *state* pada periode terakhir (T = 372) adalah *state* 1. Hal ini ditunjukkan dari nilai probabilitas observasi ke 372 berada di *state* 1 $\left(\xi_{1,372}\right)$ lebih besar daripada probabilitas observasi 372 berada di *state* 2 $\left(\xi_{2,372}\right)$, sehingga *state* pada periode terakhir (T = 372) adalah *state* 1. Nilai probabilitas ini dapat dilihat pada Lampiran 9. Kemudian, berdasarkan matriks probabilitas transisi pada persamaan (4.3), jika diketahui pada periode ke (T = 372) terjadi *state* 1, maka pada satu periode ke depan (T+1 = 373) akan terjadi *state* 1 dengan probabilitas sebesar 0,9564. Sementara itu, jika diketahui pada periode ke 372 terjadi *state* 1, maka pada satu periode ke depan 373 akan terjadi *state* 2 dengan probabilitas sebesar 0,0436. Probabilitas terjadinya *state* 1 pada periode ke 373 lebih besar daripada *state* 2. Dengan demikian, ramalan *state* yang akan terjadi pada periode 373 adalah *state* 1.

Setelah didapatkan ramalan *state*, selanjutnya adalah meramalkan ICP. Pada peramalan ICP pertama kali yang dilakukan adalah meramalkan variansi/volatilitas Berdasarkan hasil ramalan, pada periode 373 atau pada Januari 2020 tidak akan terjadi perubahan *state* dari sebelumnya yakni tetap berada pada *state* 1 atau masih berada pada kondisi volatilitas rendah. Berdasarkan hasil perhitungan pada Lampiran 10, ramalan harga ICP pada bulan Januari 2020 mengalami peningkatan dari bulan sebelumnya yaitu dari US$ 67,18/barel menjadi US$ 67,21/barel dengan besarnya ketidakpastian ditunjukkan oleh hasil ramalan variansi sebesar 13,24.

Hasil ramalan variansi ICP ini dapat dijadikan acuan pemerintah untuk penetapan asumsi ICP pada tahun 2020. Namun perlu diketahui bahwa variansi cukup besar yang akan beresiko pada kebijakan pemerintah dalam penetapan asumsi ICP pada tahun 2020. Variansi yang besar menunjukkan ketidakpastian

yang besar pula. Pemerintah dalam hal ini perlu melakukan antisipasi dana subsidi migas ketika pada realisasinya ICP berada di atas asumsi yang ditetapkan.

(Halaman ini sengaja dikosongkan)

BAB V
KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis dan pembahasan yang dilakukan, maka dapat disimpulkan bahwa:

1. Data ICP memenuhi asumsi stasioneritas dalam *mean*, namun mengandung heteroskedastisitas dan mengalami perubahan struktur pada tujuh titik.
2. MSGARCH dan GARCH menghasilkan estimasi parameter yang konvergen dilihat dari nilai $\hat{R}$ yang mendekati satu. Semua parameter dalam model MSGARCH dan GARCH signifikan kecuali parameter μ . Model terbaik untuk memodelkan data ICP adalah MSGARCH dengan nilai WAIC sebesar 1869.42. Persamaan model MSGARCH yang dihasilkan adalah model volatilitas rendah dan model volatilitas tinggi sebagai berikut.
   a. Model volatilitas rendah
   $$y_{1,t} = 0,0336 + \varepsilon_{1,t}$$
   $$\varepsilon_{1,t} = \sigma_{1,t} u_t$$
   $$\sigma_{1,t}^2 = 0,1045 + 0,1880\varepsilon_{1,t-1}^2 + 0,7836\sigma_{1,t-1}^2$$
   b. Model volatilitas tinggi
   $$y_{2,t} = 0,9810 + \varepsilon_{2,t}$$
   $$\varepsilon_{2,t} = \sigma_{2,t} u_t$$
   $$\sigma_{2,t}^2 = 4,9431 + 0,0997\varepsilon_{2,t-1}^2 + 0,8236\sigma_{2,t-1}^2$$
3. Hasil ramalan satu langkah ke depan menunjukkan bahwa pada Januari 2019 *state* yang akan terjadi adalah *state* 1 yaitu *state* dengan volatilitas rendah. ICP pada bulan Januari 2019 diduga akan mengalami kenaikan yaitu US$ 67,18/barel menjadi US$ 67,21/barel dan ramalan variansi sebesar 13,24.

5.2 Saran

Setelah dilakukan analisis hingga didapatkan kesimpulan, maka terdapat beberapa hal yang dapat digunakan sebagai bahan pertimbangan untuk penelitian selanjutnya.

1. Menambahkan model *mean* yang mengikuti proses ARIMA sehingga model *mean* tidak tunggal.

2. Melakukan evaluasi kebaikan model melalui *out-sample* menggunakan konsep *one step ahead predictions leave future out* (LFO) sebanyak rata-rata panjang *state*. Hal ini tidak diterapkan dalam penelitian ini karena untuk melakukan *refitting* model sebanyak rata-rata panjang *state* membutuhkan waktu komputasi yang sangat lama.

DAFTAR PUSTAKA

Ardia, D., Bluteau, K., Boudt, K. & Catania, L., 2018. Forecasting Risk With Markov Switching GARCH Models: A large-scale Performance Study. *International Journal of Forecasting,* pp. 733-747.

Bauwens, L., Preminger, A. & Rombouts, J. V., 2010. Theory and inference for a Amrkov Switching GARCH Model. *Econometric Journal,* Volume Vol. 13, pp. pp 218-244.

Betancourt, M., 2017. A Conceptual Introduction to Hamiltonian Monte Carlo. *arXiv:1701.02434v2.*

Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics,* Volume 31, pp. 307-327.

Box, G. E. & Jenkins, G. M., 1976. *Time Series Analysis Forecasting and Control.* California: Holden Day, Inc.

Carpenter, B. et al., 2017. Stan: A Probabilistic Programming Language. *Journal of Statistical Software,* Volume 76.

Das, D., 2004. A Bayesian MCMC Algorithm for Markov Switching GARCH models. *Journal of Econometrc Society.*

Engle, R. F., 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Journal of The Economic Society,* Volume 50, p. 987–1008.

Gelman, A. et al., 2014. *Bayesian Data Analysis, Third Edition.* Boca Raton: CRC Press.

Gray, S. F., 1996. Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics,* Volume 42, pp. 27-62.

Gunay, S., 2015. markov Regime Switching Generalized Autoregressive Conditional Heteroscedastic Model and Volatility Modeling Oil Retuns. *International Journal of Energy Economics and Policy,* Volume 5, pp. 979-985.

Hadiyat, M. A., Iriawan, N. & Mulyono, S., 2007. *Regime Switching GARCH.* Yogyakarta, Gajahmada University.

Hamilton, J. D., 1989. A new approach to the economic analysis of nonstationary time series and the business cycle. *Journal of Econometrics,* Volume 57, pp. 357-384.

Hamilton, J. D., 2005. Regime Switching Model. *Unpublish Working Paper.*

Hamilton, J. & Susmel, R., 1994. Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics,* Volume 64, pp. 307-333.

Handoko, R., 2014. Kajian Ekonomi Keuangan (KEK): Model Proyeksi Harga Minyak Mentah Indonesia (ICP) Bulanan dengan Metode ARIMA. Maret, Volume Vol. 18, pp. 1-84.

Hanke, J. E. & Wichern, D. W., 2014. *Business Forecasting Ninth Edition.* United States of of America: Pearson Education.

Henneke, J. S., Rachev, S. T., Fabbozy, F. J. & Nikolov, M., 2009. MCMC Based Estimatiomm of Markov Switching ARMA-GARCH Models. *Journal of Applied Economic,* pp. 259-271.

Herliyasari, R. R., Iriawan, N. & Kuswanto, H., 2020. Modeling Indonesian Crude Price Using Markov Switching Generalized Autoregressive Conditional Heteroscedasticity. *International Journal of Advance Science and Technology,* Volume 29, pp. 3272-3279.

Hoffman, M. D. & Gelman, A., 2014. The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo. *Journal of Machine Learning Research,* Volume 15, pp. 1351-1381.

Inclan, C. & Tiao, G. C., 1994. Use of Cumulative Sums of Squares for Retrospective Detection of Changes of Variance. *Journal of the American Statistical Association,* Volume 89, pp. 913-923.

Jiang, Y., Ma, C. Q., Yang, X. G. & Ren, Y. S., 2018. Time-Varying Volatility Feedback of Energy Prices: Evidence from Crude Oil, Petroleum Products, and Natural Gas Using a TVP-SVM Model. *Sustainability.*

King, R., Morgan, B. J. T., Gimenez, O. & Brooks, S. P., 2010. *Bayesian Analysis for Population Ecology.* Boca Raton: CRC Press.

Klassen, F., 2002. Improving GARCH volatility forecasts with regime switching GARCH. *In Advances in Markov-switching models,* pp. 223-254.

Marcucci, J., 2005. Forecasting stock market volatility with regime switching GARCH models. *Studies in Nonlinear Dynamics & Econometrics,* Volume 9, p. 1558–3708.

Neal, R. M., 2011. MCMC using Hamiltonian Monte Carlo. *arXiv:1206.1901v1.*

Nishio, M. & Arakawa, A., 2019. Performance of Hamiltonian Monte Carlo and No-U-Turn Sampler for estimating genetic parameters and breeding values. *Genetic Selection Evolution,* pp. 51-73.

Runfang, Y., Jiangze, D. & Xiaotao, L., 2017. *Improved Forecast Ability of Oil Market Volatility Based on Combined Markov Switching and GARCH-class Model.* Omaha, Elsevier B.V, pp. 415-422.

Schnatter, S. F., 2006. *Finite Mixture and Markov Switching Models.* Spring Street: Springer.

Sopipan, N. P., Sattahayam, P. & Premanode, B., 2012. Forecasting Volatility of Gold Using Markov Regime Switching and Trading Strategy. *Journal of Mathematcal Finance,* Volume 2, pp. 121-131.

Vehtari, A., Gelman, A. & Gabry, J., 2017. Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC. *Statistics and Computing,* Volume 27, p. 1413–1432.

Wei, W. W., 2006. *Time Series Analysis Univariate and Multivariate Method.* USA: Pearson Education, Inc.

(Halaman ini sengaja dikosongkan)

LAMPIRAN

Lampiran 1. Data Pengamatan

No	Tahun	Bulan	ICP
1	1989	Jan	14.81
2	1989	Feb	17.4
3	1989	Mar	17.3
4	1989	Apr	17.91
5	1989	Mei	18.34
6	1989	Jun	17.77
7	1989	Jul	17.65
8	1989	Agt	16.86
9	1989	Sep	16.69
10	1989	Okt	17.02
11	1989	Nop	17.56
12	1989	Des	17.78
13	1990	Jan	18.96
14	1990	Feb	19.11
15	1990	Mar	18.59
16	1990	Apr	17.22
17	1990	Mei	16.02
18	1990	Jun	15.24
⋮	⋮	⋮	⋮
361	2019	Jan	56.55
362	2019	Feb	61.31
363	2019	Mar	63.6
364	2019	Apr	68.31
365	2019	Mei	68.07
366	2019	Jun	61
367	2019	Jul	61.32
368	2019	Agt	57.27
369	2019	Sep	60.84
370	2019	Okt	59.82
371	2019	Nop	63.26
372	2019	Des	67.18

Lampiran 2. Uji Asumsi

a. Uji Stasioneritas

```
> library(tseries)
> adf.test(y)

        Augmented Dickey-Fuller Test

data:  y
Dickey-Fuller = -7.8061, Lag order = 7, p-value = 0.01
alternative hypothesis: stationary
```

b. Uji Heteroskedastisitas

```
> library(fDMA)
> archtest(y, lag = 1)
        Engle's LM ARCH Test
data:  y
statistic = 133.33, lag = 1, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 1 are present
> archtest(y, lag = 2)
        Engle's LM ARCH Test
data:  y
statistic = 134.87, lag = 2, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 2 are present
> archtest(y, lag = 3)
        Engle's LM ARCH Test
data:  y
statistic = 136.69, lag = 3, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 3 are present
> archtest(y, lag = 4)
        Engle's LM ARCH Test
data:  y
statistic = 136.69, lag = 4, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 4 are present
> archtest(y, lag = 5)
        Engle's LM ARCH Test
data:  y
statistic = 146.42, lag = 5, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 5 are present
> archtest(y, lag = 6)
        Engle's LM ARCH Test
data:  y
statistic = 146.22, lag = 6, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 6 are present
> archtest(y, lag = 7)
        Engle's LM ARCH Test
data:  y
statistic = 145.77, lag = 7, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 7 are present
> archtest(y, lag = 8)
        Engle's LM ARCH Test
data:  y
statistic = 145.56, lag = 8, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 8 are present
> archtest(y, lag = 9)
        Engle's LM ARCH Test
data:  y
statistic = 145.55, lag = 9, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 9 are present
```

Lampiran 2. Uji Asumsi (lanjutan)

```
> archtest(y, lag = 10)
        Engle's LM ARCH Test
data:  y
statistic = 145.78, lag = 10, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 10 are present
> archtest(y, lag = 11)
        Engle's LM ARCH Test
data:  y
statistic = 145.43, lag = 11, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 11 are present
> archtest(y, lag = 12)        Engle's LM ARCH Test
data:  y
statistic = 144.98, lag = 12, p-value < 2.2e-16
alternative hypothesis: ARCH effects of order 12 are present
```

b. Uji Structural Breaks

```
> library(ICSS)
> ICSS(y)
[1]  19  27  92 135 187 232 240
```

Lampiran 3. *Stan Code*

a. Model GARCH

```
data {
int<lower=0> T;
real y[T];
}

parameters {
real mu;
real<lower=0> alpha0;
real<lower=0,upper=1> alpha1;
real<lower=0,upper=(1-alpha1)> beta1;
real<lower=0> sigma1;
}

transformed parameters {
real<lower=0> sigma[T];
sigma[1] = sigma1;
for (t in 2:T)
sigma[t] = sqrt(alpha0 + alpha1 * pow(y[t-1] - mu, 2) + beta1 * pow(sigma[t-1],
2));
}

model {
//prior
mu ~ normal(0, 0.1);
alpha0 ~ normal(0, 1);
alpha1 ~ beta(1, 1);
beta1 ~ beta(1, 1);

//likelihood
y ~ normal(mu, sigma);
}

generated quantities{
vector[T] log_lik;
for (t in 1:T){
log_lik[t] = normal_lpdf(y[t]|mu,sigma[t]);
}
```

b. Model MSGARCH

```
data {
int T;
vector[T] y;
}

parameters {
vector<lower = 0, upper = 1>[2] p;
real mu1;
real mu2;
real <lower = 0> alpha01;
real <lower = 0> alpha02;
```

Lampiran 3. *Stan Code* (lanjutan 1)

```
real<lower=0, upper=1> alpha11;
real<lower=0, upper=1> alpha12;
real<lower=0, upper=1-alpha11> beta11;
real<lower=0, upper=1-alpha12> beta12;
real<lower=0> sigma11;
real<lower=0> sigma12;
real<lower = 0, upper = 1> xi1_init;
real y_tm1_init;
real y_tm2_init;
}

transformed parameters{
matrix[T, 2] sigma_t;
for(t in 2:T){
sigma_t[1, 1] = sigma11;
sigma_t[1, 2] = sigma12;
sigma_t[t, 1] = sqrt(alpha01 + alpha11 * pow(y[t-1], 2) + beta11 *
pow(sigma_t[t-1, 1], 2));
sigma_t[t, 2] = sqrt(alpha02 + alpha12 * pow(y[t-1], 2) + beta12 *
pow(sigma_t[t-1, 2], 2));
}
}

model {
matrix[T, 2] eta;
matrix[T, 2] xi;
vector[T] f;

// priors
p ~ beta(1, 1);
mu1 ~ normal(0, 0.1);
mu1 ~ normal(0, 0.1);
alpha01 ~ normal(0, 10);
alpha02 ~ normal(0, 10);
alpha11 ~ beta(1,1);
alpha12 ~ beta(1,1);
beta11 ~ normal(1,0.1);
beta12 ~ normal(1,0.1);
sigma11~ cauchy(0,1);
sigma12~ cauchy(0,1);
xi1_init ~ beta(2, 2);
y_tm1_init ~ normal(0, .1);
y_tm2_init ~ normal(0, .1);
// menghitung matriks kontribusi likelihood pada state 1 dan state 2
for(t in 1:T) {
eta[t,1] = exp(normal_lpdf(y[t]| mu1, sigma_t[t,1]));
eta[t,2] = exp(normal_lpdf(y[t]| mu2, sigma_t[t,2]));
}
// menghitung vektor kontribusi likelihood
for(t in 1:T) {
if(t==1) {
f[t] = p[1]*xi1_init*eta[t,1] + // stay in state 1
(1 - p[1])*xi1_init*eta[t,2] + // transition from 1 to 2
p[2]*(1 - xi1_init)*eta[t,2] + // stay in state 2
```

Lampiran 3. *Stan Code* (lanjutan 2)

```
(1 - p[2])*(1 - xi1_init)*eta[t,1]; // transition from 2 to 1

xi[t,1] = (p[1]*xi1_init*eta[t,1] +(1 - p[2])*(1 - xi1_init)*eta[t,1])/f[t];
xi[t,2] = 1.0 - xi[t,1];

} else {
f[t] = p[1]*xi[t-1,1]*eta[t,1] + // stay in state 1
(1 - p[1])*xi[t-1,1]*eta[t,2] + // transition from 1 to 2
p[2]*xi[t-1,2]*eta[t,2] + // stay in state 2
(1 - p[2])*xi[t-1,2]*eta[t,1]; // transition from 2 to 1

// menghitung matriks probabilitas berada di state 1 dan state 2
xi[t,1] = (p[1]*xi[t-1,1]*eta[t,1] +(1 - p[2])*xi[t-1,2]*eta[t,1])/f[t];
xi[t,2] = 1.0 - xi[t,1];
}
}
// likelihood
target += sum(log(f));
}

generated quantities{
matrix[T, 2] eta_p;
matrix[T, 2] xi_p;
vector[T] f_p;
vector[T] log_lik;
for(t in 1:T) {
eta_p[t,1] = exp(normal_lpdf(y[t]| mu1, sigma_t[t,1]));
eta_p[t,2] = exp(normal_lpdf(y[t]| mu2, sigma_t[t,2]));
}
for(t in 1:T) {
if(t==1) {
f_p[t] = p[1]*xi1_init*eta_p[t,1] + // stay in state 1
(1 - p[1])*xi1_init*eta_p[t,2] + // transition from 1 to 2
p[2]*(1 - xi1_init)*eta_p[t,2] + // stay in state 2
(1 - p[2])*(1 - xi1_init)*eta_p[t,1]; // transition from 2 to 1
xi_p[t,1] = (p[1]*xi1_init*eta_p[t,1] +(1 - p[2])*(1 -
xi1_init)*eta_p[t,1])/f_p[t];
xi_p[t,2] = 1.0 - xi_p[t,1];
} else {
f_p[t] = p[1]*xi_p[t-1,1]*eta_p[t,1] + // stay in state 1
(1 - p[1])*xi_p[t-1,1]*eta_p[t,2] + // transition from 1 to 2
p[2]*xi_p[t-1,2]*eta_p[t,2] + // stay in state 2
(1 - p[2])*xi_p[t-1,2]*eta_p[t,1]; // transition from 2 to 1

xi_p[t,1] = (p[1]*xi_p[t-1,1]*eta_p[t,1] +(1 - p[2])*xi_p[t-
1,2]*eta_p[t,1])/f_p[t];
xi_p[t,2] = 1.0 - xi_p[t,1];
}
}

// likelihood
log_lik=log(f_p);
}
```

Lampiran 3. *Stan Code* (lanjutan 3)

c. WAIC

```
library(loo)
library(resample)
waic <- function(stanfit){
  log_lik <- extract (stanfit, "log_lik")$log_lik
  dim(log_lik) <- if (length(dim(log_lik))==1) c(length(log_lik),1) else
    c(dim(log_lik)[1], prod(dim(log_lik)[2:length(dim(log_lik))]))
  S <- nrow(log_lik)
  n <- ncol(log_lik)
  lpd <- log(colMeans(exp(log_lik)))
  p_waic <- colVars(log_lik)
  elpd_waic <- lpd - p_waic
  waic <- -2*elpd_waic
  loo_weights_raw <- 1/exp(log_lik-max(log_lik))
  loo_weights_normalized <- loo_weights_raw/
    matrix(colMeans(loo_weights_raw),nrow=S,ncol=n,byrow=TRUE)
  loo_weights_regularized <- pmin (loo_weights_normalized, sqrt(S))
  elpd_loo <- log(colMeans(exp(log_lik)*loo_weights_regularized)/
                    colMeans(loo_weights_regularized))
  p_loo <- lpd - elpd_loo
  pointwise <- cbind(waic,lpd,p_waic,elpd_waic,p_loo,elpd_loo)
  total <- colSums(pointwise)
  se <- sqrt(n*colVars(pointwise))
  return(list(waic=total["waic"], elpd_waic=total["elpd_waic"],
              p_waic=total["p_waic"], elpd_loo=total["elpd_loo"],
p_loo=total["p_loo"],
              pointwise=pointwise, total=total, se=se))
}
```

Lampiran 4. Probabilitas observasi berada pada *state* 1 dan *state* 2

t	**Probabilitas *State* 1** $\left(\xi_{1,t}\right)$	**Probablitas *State* 2** $\left(\xi_{2,t}\right)$	***State***
1	0.0001	0.9999	2
2	0.7638	0.2362	1
3	0.9313	0.0687	1
4	0.9718	0.0282	1
5	0.9842	0.0158	1
6	0.9891	0.0109	1
7	0.9882	0.0118	1
8	0.9911	0.0089	1
9	0.9916	0.0084	1
10	0.9909	0.0091	1
11	0.9924	0.0076	1
12	0.9814	0.0186	1
13	0.9909	0.0091	1
14	0.9915	0.0085	1
15	0.9793	0.0207	1
16	0.9829	0.0171	1
17	0.9879	0.0121	1
18	0.9889	0.0111	1
19	0.0306	0.9694	2
20	0.0016	0.9984	2
⋮	⋮	⋮	⋮
367	0.9227	0.0773	1
368	0.9389	0.0611	1
369	0.9241	0.0759	1
370	0.9457	0.0543	1
371	0.9318	0.0682	1
372	0.9183	0.0817	1

Lampiran 5. Perhitungan Ramalan ICP

Perhitungan ICP (*differencing*)

$$\hat{y}_{t+1} = \mu_{s_t}$$
$$\hat{y}_{372+1} = \mu_1$$
$$\hat{y}_{373} = \mu_1$$
$$= 0,03$$

Perhitungan ICP

$$\hat{z}_{t+1} = \hat{y}_{t+1}$$
$$\hat{z}_{372+1} = \hat{y}_{372+1} + z_{372}$$
$$\hat{z}_{373} = \hat{y}_{373} + z_{372}$$
$$= 0,03 + 67,18$$
$$= 67,21$$

Perhitungan variansi

$$\hat{\sigma}^2_{t+1} = \alpha_{0,s_t} + \alpha_{1,s_t,}\varepsilon_t^2 + \beta_{1,s_t}\sigma_t^2$$
$$\hat{\sigma}^2_{372+1} = \alpha_{01} + \alpha_{11}\varepsilon^2_{372} + \beta_{11}\sigma^2_{372}$$
$$= 0,10 + 0,19(15,13) + 0,78(13,16)$$
$$= 13,24$$

(Halaman ini sengaja dikosongkan)

BIODATA PENULIS

Penulis dengan nama lengkap Ria Retna Herliyasari lahir di Ponorogo, 26 Maret 1996. Merupakan anak pertama dari Heriyanto dan Suherlin. Penulis memulai pendidikan pada tahun 2000 di Taman Kanak-Kanak Dharma Wanita Tugu 1 dan berijazah pada tahun 2002. Setelah itu, penulis melanjutkan pendidikan di Sekolah Dasar Negeri 3 Tugu, Mlarak, Ponorogo pada tahun 2002 hingga 2008. Kemudian pada tahun 2008 melanjutkan pendidikan di Sekolah Menengah Pertama Negeri 1 Mlarak Ponorogo dan berijazah pada tahun 2011. Pada tahun 2011 penulis melanjutkan Pendidikan di Sekolah Menengah Atas Negeri 1 Ponorogo hingga lulus di tahun 2014. Selanjutnya pada tahun 2014 hingga tahun 2018 penulis melanjutkan pendidikan untuk mengambil gelar sarjana (S1) di departemen Statistika Institut Teknologi Sepuluh Nopember. Setelah lulus S1, penulis melanjutkan studi S2 di departemen Statistika Institut Teknologi Sepuluh Nopember (ITS) pada tahun 2018 hingga 2020 melalui beasiswa *fresh graduate* yang diberikan oleh Institut Teknologi Sepuluh Nopember kepada para alumninya. Bagi pembaca yang ingin berdiskusi, memberikan saran dan kritik tentang Tesis ini dapat disampaikan melalui email riaretna0@gmail.com.

(Halaman ini sengaja dikosongkan)