

TUGAS AKHIR - SS234862

**PENGENDALIAN KUALITAS SISTEM *HEAT RECOVERY
STEAM GENERATOR* DI PLTGU PT PLN NUSANTARA
POWER UNIT PEMBANGKITAN GRESIK
MENGUNAKAN DIAGRAM KONTROL MAX-MEWMA
BERBASIS MODEL *RANDOM FOREST REGRESSOR***

HUSNAN ALI HUSNAIN

NRP 5003211133

Dosen Pembimbing

Prof. Dr. Muhammad Mashuri, M.T.

NIP 19620408 198701 1 001

Program Studi Sarjana Statistika

Departemen Statistika

Fakultas Sains dan Analitika Data

Institut Teknologi Sepuluh Nopember

Surabaya

2025



TUGAS AKHIR - SS234682

**PENGENDALIAN KUALITAS SISTEM *HEAT RECOVERY STEAM GENERATOR* DI PLTGU PT PLN NUSANTARA
POWER UNIT PEMBANGKITAN GRESIK
MENGUNAKAN DIAGRAM KONTROL MAX-MEWMA
BERBASIS MODEL *RANDOM FOREST REGRESSOR***

HUSNAN ALI HUSNAIN

NRP 5003211133

Dosen Pembimbing

Prof. Dr. Muhammad Mashuri, M.T.

NIP 19620408 198701 1 001

Program Studi Sarjana Statistika

Departemen Statistika

Fakultas Sains dan Analitika Data

Institut Teknologi Sepuluh Nopember

Surabaya

2025



FINAL PROJECT - SS234682

**QUALITY CONTROL OF THE HEAT RECOVERY STEAM
GENERATOR AT PLTGU PT PLN NUSANTARA POWER
GRESIK GENERATION UNIT USING A MAX-MEWMA
CONTROL CHART BASED ON A RANDOM FOREST
REGRESSOR MODEL**

HUSNAN ALI HUSNAIN

NRP 5003211133

Advisor

Prof. Dr. Muhammad Mashuri, M.T.

NIP 19620408 198701 1 001

Bachelor Program of Statistics

Department of Statistics

Faculty of Science and Data Analytics

Institut Teknologi Sepuluh Nopember

Surabaya

2025

LEMBAR PENGESAHAN

PENGENDALIAN KUALITAS SISTEM *HEAT RECOVERY STEAM GENERATOR* DI PLTGU PT PLN NUSANTARA POWER UNIT PEMBANGKITAN GRESIK MENGUNAKAN DIAGRAM KONTROL MAX-MEWMA BERBASIS MODEL *RANDOM FOREST REGRESSOR*

TUGAS AKHIR

Diajukan untuk memenuhi salah satu syarat
memperoleh gelar Sarjana Statistika pada
Program Studi Sarjana Statistika
Departemen Statistika
Fakultas Sains dan Analitika Data
Institut Teknologi Sepuluh Nopember

Oleh : **HUSNAN ALI HUSNAIN**
NRP 5003211133

Tanggal Ujian : 18 Juli 2025

Periode Wisuda : September 2025

Disetujui oleh:

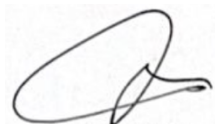
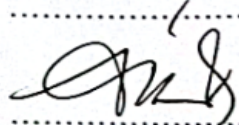
Pembimbing:

1. Prof. Dr. Muhammad Mashuri, M.T.
NIP 19872072 01404 2 001


.....

Penguji:

1. Dr. Muhammad Ahsan, S.Si.
NIP 19900425 202406 1 001
2. Diaz Fitra Aksioma, S.Si, M.Si.
NIP 19870602 201212 2 002


.....

.....

Kepala Departemen Statistika
Fakultas Sains dan Analitika Data
Dr. rer. bot Dedy Dwi Prastyo, S.Si., M.Si.
NIP 19831204 200812 1 002

(halaman ini sengaja dikosongkan)

APPROVAL SHEET

QUALITY CONTROL OF THE HEAT RECOVERY STEAM GENERATOR SYSTEM AT PLTGU PT PLN NUSANTARA POWER GRESIK GENERATION UNIT USING A MAX-MEWMA CONTROL CHART BASED ON A RANDOM FOREST REGRESSOR MODEL

FINAL PROJECT

Submitted to fulfill one of the requirements of
Obtaining a degree Bachelor of Statistics at
Bachelor Program of Statistics
Department of Statistics
Faculty of Science and Data Analytics
Institut Teknologi Sepuluh Nopember

By : **Husnan Ali Husnain**
NRP 5003211133

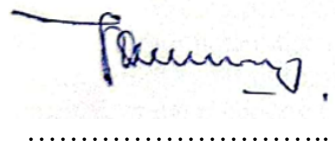
Exam Date : 18 July 2025

Graduation Period : September 2025

Approved by :

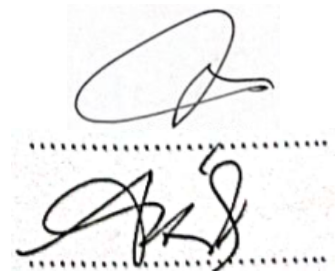
Advisor :

1. Prof. Dr. Muhammad Mashuri, M.T.
NIP 19620408 198701 1 001



Examiners :

1. Dr. Muhammad Ahsan, S.Si.
NIP 19900425 202406 1 001
2. Diaz Fitra Aksioma, S.Si, M.Si.
NIP 19870602 201212 2 002



Head of Statistics Department
Faculty of Science and Data Analytics
Dorcr.ppt Dedy Dwi Prastyo, S.Si., M.Si.
NIP 19831204 200812 1 002

(halaman ini sengaja dikosongkan)

PERNYATAAN ORISINALITAS

Yang bertanda tangan di bawah ini:

Nama Mahasiswa / NRP : Husnan Ali Husnain / 5003211133

Departmen : Statistika

Dosen Pembimbing / NIP : Prof. Dr. Muhammad Mashuri, M.T. / 19620408 198701 1 001

dengan ini menyatakan bahwa Tugas Akhir dengan judul **"Pengendalian Kualitas *Sistem Heat Recovery Steam Generator* di PLTGU PT PLN Nusantara Power Unit Pembangkit Gresik Menggunakan Diagram Kontrol Max-MEWMA Berbasis Model *Random Forest Regressor*"** adalah hasil karya sendiri, bersifat orisinal, dan ditulis dengan mengikuti kaidah penulisan ilmiah.

Bilamana di kemudian hari ditemukan ketidaksesuaian dengan pernyataan ini, maka saya bersedia menerima sanksi sesuai dengan ketentuan yang berlaku di Institut Teknologi Sepuluh Nopember.

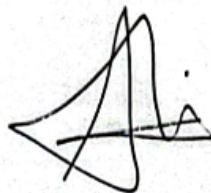
Surabaya, 13 Juni 2025

Mengetahui
Dosen Pembimbing



Prof. Dr. Muhammad Mashuri, M.T
NIP 19620408 198701 1 001

Mahasiswa



Husnan Ali Husnain
NRP 5003211133

(halaman ini sengaja dikosongkan)

STATEMENT OF ORIGINALITY

The undersigned below:

Student's Name / NRP : Husnan Ali Husnain/ 5003211133

Department : Statistics

Advisor / NIP : Prof. Dr. Muhammad Mashuri, M.T. / 19620408 198701 1 001

Hereby declare that the Final Project with the title of **"Quality Control of the Heat Recovery Steam Generator System at PLTGU PT PLN Nusantara Power Gresik Generation Unit Using a Max-MEWMA Control Chart Based on a Random Forest Regressor Model"** is the result of my own work, is original, and is written by following the rules of scientific writing.

If in the future there is a discrepancy with this statement, then I am willing to accept sanctions in accordance with the provisions that apply at Institut Teknologi Sepuluh Nopember.

Surabaya, 13 June 2025

Acknowledged
Advisor



Prof. Dr. Muhammad Mashuri, M.T
NIP 19620408 198701 1 001

Student



Husnan Ali Husnain
NRP 5003211133

(halaman ini sengaja dikosongkan)

ABSTRAK

PENGENDALIAN KUALITAS SISTEM HEAT RECOVERY STEAM GENERATOR DI PLTGU PT PLN NUSANTARA POWER UNIT PEMBANGKITAN GRESIK MENGGUNAKAN DIAGRAM KONTROL MAX-MEWMA BERBASIS MODEL RANDOM FOREST REGRESSOR

Nama Mahasiswa / NRP : Husnan Ali Husnain / 5003211133
Departemen : Statistika FSAD - ITS
Dosen Pembimbing : Prof. Dr. Muhammad Mashuri, M.T.

Abstrak

Sistem *Heat Recovery Steam Generator* (HRSG) pada PLTGU PT. PLN Nusantara Power UP Gresik berperan penting dalam meningkatkan efisiensi energi dan menurunkan emisi karbon. Namun, pengendalian kualitas pada sistem ini menghadapi tantangan signifikan, seperti autokorelasi tinggi dan deteksi pergeseran kecil pada parameter kritis, yaitu *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Gas Temperature* (GT). Penelitian ini menggunakan penerapan diagram kontrol Max-MEWMA berbasis residual dari model *Random Forest Regressor* untuk mengatasi permasalahan tersebut. Model dibangun berdasarkan lag-lag signifikan dari analisis PACF, dan residual yang dihasilkan diuji untuk memenuhi asumsi independensi. Hasil pengujian menunjukkan bahwa residual telah bebas dari autokorelasi, meskipun tidak sepenuhnya memenuhi asumsi normal multivariat. Diagram kontrol Max-MEWMA kemudian diterapkan pada data fase I dan II untuk mendeteksi sinyal *out-of-control* secara simultan. Hasil analisis menunjukkan bahwa proses belum sepenuhnya terkendali secara statistik, dengan karakteristik *Fuel Gas Flow* (FGF) dan *Combustion Pressure* (CP) sebagai penyumbang sinyal *out-of-control* terbanyak. Temuan ini divalidasi melalui diagram Ishikawa, yang mengidentifikasi faktor manusia, mesin, metode, dan pengukuran sebagai penyebab utama. Meskipun indeks kapabilitas multivariat (MC_p dan MC_{pk}) bernilai > 1 , namun proses belum dapat dinyatakan kapabel karena belum memenuhi kestabilan statistik. Penelitian ini menunjukkan bahwa penggabungan metode *Random Forest Regressor* dan Max-MEWMA efektif dalam meningkatkan sensitivitas deteksi dan mendukung pengendalian kualitas sistem HRSG secara *real-time*.

Kata Kunci: *Heat Recovery Steam Generator* (HRSG), *Random Forest Regressor*, Max-MEWMA, Pengendalian Kualitas, Residual, Kapabilitas Proses, *Out-of-Control*.

(halaman ini sengaja dikosongkan)

ABSTRACT

QUALITY CONTROL OF THE HEAT RECOVERY STEAM GENERATOR SYSTEM AT PLTGU PT PLN NUSANTARA POWER GRESIK GENERATION UNIT USING A MAX-MEWMA CONTROL CHART BASED ON A RANDOM FOREST REGRESSOR MODEL

Student Name / NRP : Husnan Ali Husnain / 5003211133
Department : Statistics F-Scientics - ITS
Advisor : Prof. Dr. Muhammad Mashuri, M.T.

Abstract

The Heat Recovery Steam Generator (HRSG) system at PLTGU PT. PLN Nusantara Power UP Gresik plays a critical role in improving energy efficiency and reducing carbon emissions. However, quality control in this system faces significant challenges, such as high autocorrelation and difficulty detecting small shifts in critical parameters, namely Fuel Gas Flow (FGF), Combustion Pressure (CP), and Gas Temperature (GT). This study uses the application of a Max-MEWMA control chart based on residuals from a Random Forest Regressor model to address these issues. The model was constructed using significant lags identified through PACF analysis, and the resulting residuals were tested to satisfy the independence assumption. The tests confirmed that residuals were free from autocorrelation, although not fully normally distributed. The Max-MEWMA control chart was applied to both Phase I and Phase II data to simultaneously detect out-of-control signals. The analysis revealed that the process was not fully in statistical control, with Fuel Gas Flow (FGF) and Combustion Pressure (CP) being the main contributors to the out-of-control signals. These findings were validated using an Ishikawa diagram, which identified human, machine, method, and measurement factors as the primary causes. Although the multivariate process capability indices (MC_p and MC_{pk}) were greater than one, the process could not be deemed capable due to a lack of statistical stability. This research demonstrates that combining the Random Forest Regressor with the Max-MEWMA control chart enhances detection sensitivity and supports real-time quality control of HRSG systems.

Keywords: Heat Recovery Steam Generator (HRSG), Random Forest Regressor, Max-MEWMA, Quality Control, Residual, Process Capability, Out-of-Control.

(halaman ini sengaja dikosongkan)

KATA PENGANTAR

Puji syukur kami panjatkan kehadiran Allah SWT karena dengan rahmat, karunia, serta taufik dan hidayah-Nya penulis dapat menyelesaikan Tugas Akhir yang berjudul “Pengendalian Kualitas Sistem *Heat Recovery Steam* Generator di PLTGU PT PLN Nusantara Power Unit Pembangkit Gresik Menggunakan Diagram Kontrol Max-MEWMA Berbasis Model *Random Forest Regressor*”. Dalam penyusunan tugas akhir ini, penulis mendapat bantuan dan dukungan dari berbagai pihak, oleh sebab itu penulis mengucapkan terima kasih kepada :

1. Mamah, ayah, kakak, dan adek tercinta yang selalu memberikan doa, semangat, kasih sayang dan berbagai dukungan tiada henti baik secara moral maupun material kepada penulis.
2. Dr.rer pol. Dedy Dwi Prastyo, S.Si., M. Si. selaku Kepala Departemen Statistika yang telah memfasilitasi sumber daya dan kebijakan akademis kepada penulis.
3. Ibu Shofi Andari, M.Si., Ph.D. selaku Kepala Program Studi Sarjana Statistika telah membimbing, mengarahkan, serta memberikan dukungan pada penulis.
4. Ibu Dr. Wibawati, S.Si, M.Si. selaku sekretaris Departemen Statistika yang telah memfasilitasi sumber daya dan kebijakan akademis kepada penulis.
5. Prof. Dr. Muhammad Mashuri, M.T. selaku dosen pembimbing yang telah membimbing, mengarahkan, serta memberikan dukungan pada penulis untuk menyelesaikan tugas akhir.
6. Ibu Santi Puteri Rahayu, S.Si., M.Si., Ph.D. selaku dosen wali yang telah memberikan motivasi, saran belajar, dan dukungan selama proses perkuliahan.
7. Bapak Dr. Muhammad Ahsan, S.Si. selaku dosen penguji 1 yang telah memberikan arahan, saran, serta dukungan serta motivasi selama pengerjaan Tugas Akhir.
8. Ibu Diaz Fitra Aksioma, S.Si, M.Si. selaku dosen penguji 2 yang telah memberikan arahan, saran, serta dukungan serta motivasi selama pengerjaan Tugas Akhir.
9. Segenap dosen pengajar dan para staff tenaga pendidik Departemen Statistika Fakultas Sains dan Analitika Data ITS yang telah memberikan bekal ilmu dan membantu penulis selama masa perkuliahan.
10. Sahabat baik Tubagus Kelana Aji yang selalu menemani dalam suka dan duka dari awal Sekolah Menengah Pertama hingga penulis bisa menyelesaikan tugas akhir ini.
11. Risma Dhiva Syabila, yang selalu mendidik penulis untuk tetap selalu kuat dalam segala hal serta memberikan dukungan dan semangat kepada penulis dalam proses pengerjaan Tugas Akhir.
12. Teman seperjuangan Statistika, Hanif, Dafa, Deka, Wily, Rafi, Jehezkiel, Hafizh, Danis, Daniel, Yuri, Farham, Afthon, Alfat, Daniyal, Galih dan Evana. Terima kasih telah menjadikan penulis sebagai teman dan pemberi tawa kepada penulis selama masa perkuliahan.
13. Semua pihak yang telah membantu dalam penyusunan proposal tugas akhir ini, baik secara langsung maupun tidak langsung yang tidak dapat disebutkan satu persatu oleh penulis. Semoga bantuan dan kerja sama yang telah dilakukan mendapat pahala yang setimpal dari Allah Subhanahu wa ta'ala.

Surabaya, 13 Juni 2025

Penulis

halaman ini sengaja dikosongkan)

DAFTAR ISI

LEMBAR PENGESAHAN	i
APPROVAL SHEET	iii
PERNYATAAN ORISINALITAS	v
STATEMENT OF ORIGINALITY	vii
ABSTRAK	ix
ABSTRACT	xi
KATA PENGANTAR	xiii
DAFTAR ISI	xv
DAFTAR GAMBAR	xvii
DAFTAR TABEL	xix
BAB I	1
PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Batasan Masalah	3
1.4 Tujuan	3
1.5 Manfaat	3
BAB II	5
TINJAUAN PUSTAKA	5
2.1 Hasil Penelitian Terdahulu	5
2.2 Energi Listrik	5
2.3 Heat Recovery Steam Generator (HRSG)	6
2.4 Uji Dependensi Antar Variabel	7
2.5 Uji Terasvrita	8
2.6 Autocorrelation Function dan Partial Auto Correlatation Function	9
2.7 Random Forest Regressor	10
2.8 Pengujian Asumsi Residual Independen	13
2.9 Pengujian Asumsi Residual Normal Multivariat	14
BAB III	25
METODOLOGI PENELITIAN	25
3.1 Sumber Data	25
3.2 Variabel Penelitian	25

3.3 Struktur Data.....	25
3.4 Langkah Analisis	26
3.5 Diagram Alir.....	28
BAB IV	29
ANALISIS PEMBAHASAN.....	29
BAB V	51
KESIMPULAN DAN SARAN	51
5.1 Kesimpulan.....	51
5.2 Saran	51
DAFTAR PUSTAKA.....	53
LAMPIRAN	57
Lampiran 1 Data GT 23 Operation 202207010005 atau Blok II pada Pola operasi (1,1,1) PLTGU PT. PLN Nusantara Power UP Gresik pada tanggal 1 Juli – 7 Juli 2022.....	57
Lampiran 2 Syntax R Uji Normalitas.....	58
Lampiran 3 Output Uji Normalitas	58
Lampiran 4 Syntax R Uji Dependensi	59
Lampiran 5 Output Uji Dependensi	59
Lampiran 6 Syntax Deteksi <i>Outlier</i>	60
Lampiran 7 Syntax Random Forest Regressor.....	61
Lampiran 8 Syntax Uji MSE.....	63
Lampiran 9 Output Uji MSE.....	63
Lampiran 10 Data Residual Fase I	64
Lampiran 11 Data Residual Fase II.....	65
Lampiran 12 Uji MCC	66
Lampiran 13 Syntax Diagram Kontrol Max-MEWMA Fase I	66
Lampiran 14 Output Diagram Kontrol Max-MEWMA Fase I	68
Lampiran 15 Output Diagram Kontrol Max-MEWMA $\lambda = 0,2$	71
Lampiran 16 Syntax Pembersihan OOC Diagram Kontrol Fase I $\lambda = 0,2$	72
Lampiran 17 Output Pembersihan OOC Diagram Kontrol Fase I $\lambda = 0,2$	73
Lampiran 18 Syntax Diagram Kontrol Max-MEWMA Fase II.....	75
Lampiran 19 Output Diagram Kontrol Max-MEWMA Fase II $\lambda = 0,2$	76
Lampiran 20 Ouput Diagram Kontrol Max-MEWMA Fase II.....	77
Lampiran 21 Syntax Kombinasi Karakteristik Kualitas	78
Lampiran 22 Output Kombinasi Karakteristik Kualitas	80
BIODATA PENULIS.....	81

DAFTAR GAMBAR

Gambar 2. 1 Struktur Random Forest Regressor multi-output (Parejo, et al., 2021).....	10
Gambar 2. 2 Diagram Ishikiwa	19
Gambar 2. 3 Bagan siklus kerja PLTGU.....	22
Gambar 2. 4 Diagram Alir Proses Produksi Listrik di PLTGU.	23
Gambar 3. 1 Diagram Alir.....	28
Gambar 4. 1 Time Series Plot (a) FGF, (b) CP, (c) GT	30
Gambar 4. 3 Time Series Plot Fase I (a) FGF, (b) CP, (c) GT	32
Gambar 4. 4 Plot ACF (a) FGF, (b) CP, (c) GT	34
Gambar 4. 5 Plot PACF (a) Fuel Gas Flow, (b) Combustion Pressure, dan (c) Gas Temperature.....	35
Gambar 4. 6 Aktual vs Prediksi Fase I (a) Fuel Gas Flow, (b) Combustion Pressure, (c) Gas Temperature.	36
Gambar 4. 7 Aktual vs Prediksi Fase II (a) Fuel Gas Flow, (b) Combustion Pressure, (c) Gas Temperature.	38
Gambar 4. 8 Histogram (a) Fuel Gas Flow, (b) Combustion Pressure, dan (c) Gas Temperature.....	41
Gambar 4. 9 Diagram kendali Max - MEWMA $\lambda=0,2$ fase I.....	44
Gambar 4. 10 Diagram kendali Max - MEWMA $\lambda=0,2$ fase I in control	44
Gambar 4. 11 Diagram kendali Max - MEWMA $\lambda=0,2$ fase II.....	45
Gambar 4.12 Kombinasi (a) Fuel Gas Flow dan Combustion Pressure, (b) Fuel Gas Flow dan Gas Temperatur, dan (c) Combustion Pressure dan Gas Temperature.....	47
Gambar 4. 13 Diagram Ishikawa (a) Fuel Gas Flow dan (b) Combustion Pressure.....	48

(halaman ini sengaja di kosongkan)

DAFTAR TABEL

Tabel 3. 1 Variabel Penelitian	25
Tabel 3. 2 Struktur Data Penelitian	26
Tabel 4. 1 Statistika Deskriptif Fase I	29
Tabel 4. 2 Statistika Deskriptif Fase II	29
Tabel 4. 3 Hasil Uji Barlett	31
Tabel 4. 4 Hasil Uji Terasvirta	32
Tabel 4. 5 MCC Residual Fase I	39
Tabel 4. 6 Hasil Uji Portmanteau Residual Fase I	39
Tabel 4. 7 Uji Normal Multivariat Residual Fase I	39
Tabel 4. 8 Hasil Kombinasi λ UCL Simulasi Bootstrap	42

(halaman ini sengaja dikosongkan)

BAB I

PENDAHULUAN

1.1 Latar Belakang

Energi listrik merupakan sumber daya utama yang mendukung berbagai aspek kehidupan, mulai dari kebutuhan rumah tangga, industri, hingga transportasi. Menurut *International Energy Agency* (2023), konsumsi energi listrik global meningkat rata-rata 6% per tahun sejak tahun 2000. Peningkatan ini didorong oleh pertumbuhan populasi, urbanisasi, dan digitalisasi yang masif. Namun, ketergantungan pada bahan bakar fosil sebagai sumber utama energi listrik masih mendominasi, menyumbang lebih dari 80% total konsumsi energi global. Hal ini berdampak pada peningkatan emisi karbon yang signifikan, sebagaimana dilaporkan oleh *Intergovernmental Panel on Climate Change* (IPCC) (2021), bahwa sektor energi menyumbang hampir 75% dari total emisi gas rumah kaca global.

Di Indonesia, kebutuhan listrik nasional terus meningkat seiring dengan pertumbuhan ekonomi dan perubahan gaya hidup masyarakat. Data Kementerian Energi dan Sumber Daya Mineral (2022) menunjukkan bahwa konsumsi listrik per kapita mengalami kenaikan rata-rata 8% per tahun selama lima tahun terakhir. Peningkatan ini didorong oleh meningkatnya elektrifikasi di wilayah pedesaan dan peningkatan aktivitas industri. Namun, lebih dari 60% kebutuhan listrik di Indonesia masih dipenuhi oleh pembangkit berbasis bahan bakar fosil seperti batu bara dan gas alam. Ketergantungan ini tidak hanya meningkatkan emisi karbon tetapi juga membuat pasokan energi rentan terhadap fluktuasi harga global.

Untuk mengatasi tantangan tersebut, Indonesia telah mulai mengadopsi teknologi ramah lingkungan, salah satunya adalah Pembangkit Listrik Tenaga Gas dan Uap (PLTGU). Teknologi ini menawarkan efisiensi termal hingga 50% lebih tinggi dan pengurangan emisi karbon hingga 30% dibandingkan pembangkit berbasis batu bara (PLN, 2023). PLTGU bekerja dengan prinsip siklus gabungan (*Combined Cycle*), memanfaatkan panas buang dari turbin gas untuk menghasilkan uap yang digunakan dalam turbin uap, sehingga meningkatkan efisiensi dan mengurangi konsumsi bahan bakar tambahan.

Salah satu elemen penting dalam PLTGU adalah sistem *Heat Recovery Steam Generator* (HRSG). HRSG memainkan peran kunci dalam memaksimalkan efisiensi energi dengan menangkap panas buang dari turbin gas dan menggunakannya untuk menghasilkan uap. Sistem ini terdiri dari tiga komponen utama: *economizer*, *evaporator*, dan *superheater*. *Economizer* memanfaatkan panas awal untuk memanaskan air umpan sebelum memasuki *evaporator*. *Evaporator* mengubah air yang dipanaskan menjadi uap jenuh, sedangkan *superheater* meningkatkan suhu uap menjadi *superheated steam* yang siap digunakan untuk menggerakkan turbin uap. Proses ini secara signifikan mengurangi konsumsi bahan bakar tambahan, menekan emisi karbon, dan meningkatkan efisiensi keseluruhan pembangkit listrik.

Namun, seperti teknologi lainnya, HRSG menghadapi tantangan besar terkait pengendalian kualitas operasional. Tantangan utama mencakup menjaga stabilitas parameter kritis seperti *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Exhaust Gas Temperature* (GT). Parameter kritis seperti *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Exhaust Gas Temperature* (GT) harus dijaga stabil untuk memastikan sistem bekerja optimal. Penyimpangan kecil pada parameter ini dapat menyebabkan ketidakseimbangan proses, seperti *overheating*, tekanan uap yang tidak stabil, atau bahkan kerusakan pada komponen penting seperti pipa, drum, dan *heat exchanger*.

Selain itu, data operasional HRSG memiliki sifat *time-series* dengan autokorelasi tinggi, yang menciptakan tantangan tambahan dalam proses pengendalian kualitas. Metode pengendalian tradisional, seperti *Shewhart Control Charts*, sering kali tidak mampu mendeteksi

perubahan kecil pada proses multivariat yang dinamis. Oleh karena itu, pendekatan yang lebih canggih diperlukan untuk memastikan deteksi dini anomali dan menjaga stabilitas proses.

Random Forest Regressor dipilih karena kemampuannya dalam memodelkan hubungan non-linier dan menghasilkan residual yang lebih bebas autokorelasi, sehingga cocok digunakan dalam diagram kontrol modern seperti Max-MEWMA. Karakteristik ini menjadikan residual lebih sensitif dalam mendeteksi perubahan kecil dalam sistem dan lebih andal untuk digunakan dalam pengendalian kualitas. Teknik ini bekerja dengan membangun banyak pohon keputusan dan menggabungkan hasil prediksinya untuk mengurangi *overfitting* serta meningkatkan generalisasi. *Random Forest* tidak memerlukan asumsi distribusi tertentu dan mampu menangani data multivariat yang kompleks. Residual ini bebas dari autokorelasi, sehingga cocok untuk digunakan dalam diagram kontrol. Breiman (2001) menunjukkan bahwa *Random Forest* sangat fleksibel dalam menangkap pola data kompleks tanpa memerlukan asumsi distribusi linier, menjadikannya metode ideal untuk sistem seperti HRSG. Selain itu, pendekatan berbasis *Random Forest* telah banyak digunakan untuk sistem pemantauan kualitas karena kestabilan dan akurasi modelnya dalam kondisi *real-world*.

Penelitian sebelumnya oleh Syahrozhadl (2023) dengan judul "Pengendalian Kualitas Sistem Heat Recovery Steam Generator di PLTGU PT PLN Nusantara Power Unit Pembangkitan Gresik Menggunakan Diagram Kontrol MEWMV dan MEWMA Berbasis Multioutput Least Square Support Vector Regression (MLS-SVR)" menunjukkan bahwa diagram kontrol berbasis residual model MLS-SVR mampu mendeteksi variabilitas proses dan pergeseran rata-rata pada sistem HRSG dengan efektif. Pada fase I, data berhasil dikendalikan secara statistik setelah dilakukan penanganan terhadap titik-titik yang *out-of-control*, namun pada fase II masih ditemukan 126 pengamatan yang berada di luar batas kendali, yang menunjukkan bahwa proses belum sepenuhnya terkendali secara statistik. Selain itu, pembobot yang digunakan yaitu $\lambda = 0,1$ masih menunjukkan hasil yang belum optimal.

Berdasarkan hal tersebut, penelitian ini dilakukan sebagai tindak lanjut untuk memperbaiki dan meningkatkan hasil penelitian sebelumnya dengan menggunakan model yang berbeda, yaitu *Random Forest Regressor*. Model ini diharapkan dapat mengatasi keterbatasan model sebelumnya dengan menghasilkan residual yang lebih stabil dan sensitif, sehingga dapat meningkatkan akurasi deteksi sinyal pada sistem pengendalian kualitas HRSG. Penelitian ini juga mengimplementasikan diagram kontrol Max-MEWMA (*Maximum Multivariate Exponentially Weighted Moving Average*) yang memiliki sensitivitas lebih tinggi dalam mendeteksi pergeseran kecil pada data multivariat.

Dengan kombinasi ini, sistem HRSG dapat dimonitor secara *real-time*, memungkinkan tindakan korektif segera untuk mencegah gangguan besar. Penelitian ini bertujuan untuk mengimplementasikan diagram kontrol Max-MEWMA berbasis residual dari model *Random Forest Regressor* untuk meningkatkan sensitivitas deteksi secara signifikan dan memastikan penerapan yang efektif dalam kondisi operasional nyata. Sensitivitas deteksi yang lebih tinggi sangat penting untuk mendeteksi perubahan kecil pada sistem HRSG secara *real-time*, yang dapat mencegah gangguan besar dan memastikan stabilitas operasional dalam kondisi dinamis. Pendekatan ini dirancang tidak hanya untuk mendeteksi pergeseran kecil, tetapi juga untuk memberikan solusi *real-time* terhadap gangguan sistem, sehingga mampu mengatasi keterbatasan metode sebelumnya

1.2 Rumusan Masalah

1. Bagaimana pengendalian kualitas sistem HRSG di PLTGU PT Nusantara Power Unit Pembangkitan Gresik berdasarkan parameter kritis *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Exhaust Gas Temperature* (GT) menggunakan diagram kontrol Max-MEWMA berbasis model?

2. Apa faktor-faktor penyebab utama variabel *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Exhaust Gas Temperature* (GT) yang menyebabkan proses *out of control* pada sistem HRSG di PLTGU PT PLN Nusantara Power Unit Pembangkitan Gresik?
3. Bagaimana kapabilitas proses pada kualitas HRSG di PLTGU PT Nusantara Power Unit Pembangkitan Gresik berdasarkan analisis statistik?

1.3 Batasan Masalah

Penelitian ini memiliki beberapa batasan untuk memastikan fokus dan pencapaian tujuan secara efektif. Batasan masalah yang diterapkan adalah sebagai berikut.

1. Penelitian hanya dilakukan pada sistem HRSG di PLTGU PT Nusantara Power Unit Pembangkitan Gresik.
2. Parameter yang dianalisis terbatas pada *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Exhaust Gas Temperature* (GT) sebagai parameter kritis yang memengaruhi kualitas operasional HRSG.
3. Metode yang digunakan adalah diagram kontrol Max-MEWMA berbasis model, dengan asumsi data residual telah memenuhi karakteristik bebas autokorelasi.
4. Studi ini hanya mencakup analisis statistik untuk mendeteksi pergeseran rata-rata dan variabilitas sistem HRSG.
5. Data yang digunakan merupakan data operasional historis yang tersedia dari sistem HRSG, tanpa pengumpulan data secara langsung di lapangan.

1.4 Tujuan

Penelitian ini bertujuan untuk.

1. Menerapkan diagram kontrol Max-MEWMA berbasis model untuk memonitor kualitas sistem HRSG di PLTGU PT. PLN Nusantara Power UP Gresik.
2. Mengidentifikasi faktor-faktor penyebab utama variabel *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Exhaust Gas Temperature* (GT) yang menyebabkan proses *out of control* pada sistem HRSG di PLTGU PT. PLN Nusantara Power UP Gresik.
3. Menganalisis kapabilitas proses kualitas HRSG untuk memastikan stabilitas operasional dan efisiensi sistem.

1.5 Manfaat

Penelitian ini diharapkan memberikan manfaat sebagai berikut.

1. Menyediakan pendekatan yang inovatif dalam pengendalian kualitas sistem HRSG di PLTGU menggunakan diagram kontrol Max-MEWMA berbasis model, yang dapat menjadi acuan bagi penelitian di masa depan.
2. Membantu PLTGU PT. PLN Nusantara Power UP Gresik dalam memantau dan menjaga stabilitas operasional sistem HRSG secara *real-time*, sehingga meningkatkan efisiensi operasional dan umur peralatan.
3. Mendukung upaya transisi energi bersih di Indonesia melalui pengurangan emisi karbon dengan meningkatkan efisiensi sistem pembangkit listrik tenaga gas dan uap (PLTGU).
4. Memberikan kontribusi terhadap pengembangan teknologi pengendalian kualitas proses multivariat, yang relevan untuk diterapkan dalam berbagai sektor industri lainnya.

(halaman ini sengaja dikosongkan)

BAB II

TINJAUAN PUSTAKA

2.1 Hasil Penelitian Terdahulu

Penelitian sebelumnya terkait pengendalian kualitas pada sistem *Heat Recovery Steam Generator* (HRSG) telah dilakukan oleh beberapa peneliti. Erika Adhisty Iskandar (2020) melakukan penelitian dengan judul “Pengendalian Kualitas Sistem Heat Recovery Steam Generator (HRSG) PLTGU di PT X dengan Diagram Kontrol Maximum Multivariate Cumulative Sum”. Penelitian ini menunjukkan bahwa rata-rata variabel gas flow dan gas temperature berada dalam batas spesifikasi perusahaan. Namun, nilai minimum variabel gas flow dan gas temperature masih berada di luar batas spesifikasi yang ditentukan.

Selain itu, penelitian terkait telah dilakukan oleh Wildy Fachrizal (2020) dalam penelitian berjudul "Pengendalian Kualitas Garam Konsumsi Beryodium Di PT Garam (Persero) Menggunakan Diagram Maximum Multivariate Exponential Weighted Moving Average (Max-MEWMA)". Hasil penelitian menunjukkan bahwa produksi garam konsumsi beryodium mengalami ketidakterkendalian proses akibat variasi kadar H_2O dan $NaCl$, serta gangguan teknis pada mesin. Analisis kapabilitas proses menunjukkan bahwa meskipun presisi cukup baik, akurasi masih rendah, sehingga proses produksi belum kapabel secara keseluruhan.

Selanjutnya, Tita Riza Aulia (2021) melakukan penelitian pada kualitas proses produksi air di PDAM Tirtanadi Sumatera Utara menggunakan diagram kontrol MEWMA berbasis residual model Multioutput Least Square Support Vector Regression (MLS-SVR). Penelitian ini menunjukkan bahwa diagram kontrol mampu mengatasi data autokorelasi sehingga menghasilkan diagram yang terkendali secara statistik.

Pendekatan serupa juga digunakan Annisa Yunika (2023) dalam penelitiannya yang berjudul “Pengendalian Kualitas Proses Produksi Gula Kristal Putih di Pabrik Gula Djatiroto PT Perkebunan Nusantara XI Menggunakan Diagram Kontrol Max-MCUSUM Berbasis Residual Model MLS-SVR” menemukan bahwa diagram kontrol Max-MCUSUM efektif dalam mengatasi data autokorelasi dan mampu memonitor rata-rata serta variabilitas proses secara simultan.

Penelitian sebelumnya dilakukan oleh Khusna, Mashuri, Suhartono, Prastyo, & Ahsan (2019) yang mengusulkan diagram kontrol Max-MCUSUM berbasis MLS-SVR untuk memonitor rata-rata dan variabilitas proses pada data berautokorelasi multivariat secara simultan. Terakhir, Nur Laili Syahrozhadl (2023) memperluas penelitian terkait HRSG dengan pendekatan berbasis residual MLS-SVR. Dalam studinya yang berjudul “Pengendalian Kualitas Sistem Heat Recovery Steam Generator (HRSG) di PLTGU PT PLN Nusantara Power Unit Pembangkitan Gresik Menggunakan Diagram Kontrol MEWMV dan MEWMA Berbasis Residual Model Multioutput Least Square Support Vector Regression (MLS-SVR)”, ia menemukan bahwa diagram kontrol ini mampu mengidentifikasi variabel penyebab out of control secara akurat, dengan variabel laju aliran bahan bakar dan tekanan pembakaran menjadi penyebab utama.

2.2 Energi Listrik

Energi listrik adalah salah satu bentuk energi yang dihasilkan oleh pergerakan partikel bermuatan, seperti elektron, dalam suatu rangkaian listrik. Energi ini memiliki peran vital dalam kehidupan modern karena digunakan untuk berbagai kebutuhan, mulai dari penerangan, penggerak mesin, hingga pengoperasian perangkat teknologi canggih (Kumparan, 2023). Energi listrik dapat diperoleh dari berbagai sumber daya alam, termasuk pembangkit listrik tenaga air, tenaga surya, tenaga angin, panas bumi, dan bahan bakar fosil (Sainstekno, 2024).

Proses produksi energi listrik melibatkan tahapan kompleks, seperti eksplorasi sumber energi, konversi energi primer menjadi energi listrik melalui generator, hingga distribusi ke pengguna akhir melalui jaringan transmisi dan distribusi (Kimia, 2023).

Dalam proses pembangkitan, berbagai teknologi digunakan sesuai dengan sumber daya yang tersedia. Sebagai contoh, pembangkit listrik tenaga air menggunakan aliran air untuk memutar turbin, sedangkan pembangkit listrik tenaga gas dan uap (PLTGU) memanfaatkan panas dari pembakaran bahan bakar fosil untuk menghasilkan uap yang memutar turbin. Energi listrik yang dihasilkan kemudian ditransmisikan menggunakan tegangan tinggi untuk mengurangi kerugian energi selama perjalanan ke pengguna akhir (Kumparan, 2023).

Manfaat energi listrik sangat signifikan, baik dalam meningkatkan kualitas hidup masyarakat, mendorong pertumbuhan ekonomi, maupun mendukung perkembangan teknologi modern (Sainstekno, 2024). Kehadiran energi listrik memungkinkan pengoperasian alat elektronik, penerangan yang efisien, serta mendorong inovasi di bidang industri dan transportasi. Namun, pengelolaan energi listrik juga menghadapi tantangan, seperti keterbatasan sumber energi primer, efisiensi proses pembangkitan, dan dampak lingkungan (Kimia, 2023). Oleh karena itu, penggunaan teknologi ramah lingkungan seperti pembangkit energi terbarukan dan Heat Recovery Steam Generator (HRSG) menjadi solusi untuk meningkatkan efisiensi sekaligus mengurangi dampak negatif terhadap lingkungan (Sainstekno, 2024).

Sumber energi listrik dapat dibagi menjadi dua kategori utama, yaitu sumber energi primer seperti air, angin, dan bahan bakar fosil, serta sumber energi sekunder yang dihasilkan dari konversi energi primer. Dengan meningkatnya kebutuhan global terhadap energi listrik, pengembangan teknologi pembangkit yang efisien dan berkelanjutan terus menjadi prioritas utama dalam industri energi listrik (Kumparan, 2023).

2.3 Heat Recovery Steam Generator (HRSG)

Heat Recovery Steam Generator (HRSG) merupakan salah satu komponen utama dalam Pembangkit Listrik Tenaga Gas dan Uap (PLTGU) yang berfungsi untuk memanfaatkan energi panas dari gas buang turbin gas. Gas buang ini, yang masih memiliki suhu tinggi, digunakan untuk memanaskan air dalam pipa-pipa penukar panas hingga berubah menjadi uap. Uap yang dihasilkan kemudian dimanfaatkan untuk menggerakkan turbin uap, sehingga meningkatkan efisiensi pembangkit listrik secara keseluruhan. Teknologi HRSG dirancang untuk memaksimalkan pemanfaatan energi yang biasanya terbuang dalam sistem pembangkitan konvensional, menjadikannya elemen penting dalam meningkatkan efisiensi dan menurunkan emisi karbon di sektor energi (Polines, 2024).

Dalam PLTGU PT. PLN Nusantara Power UP Gresik, performa HRSG sangat dipengaruhi oleh stabilitas parameter kritis seperti laju aliran bahan bakar (*Fuel Gas Flow*), tekanan pembakaran (*Combustion Pressure*), dan suhu gas buang (*Exhaust Gas Temperature*). Ketiga parameter ini harus dikendalikan secara optimal untuk memastikan sistem bekerja dalam batas toleransi yang telah ditentukan. Namun, kompleksitas operasional HRSG sering kali menimbulkan tantangan, terutama ketika data proses menunjukkan fluktuasi atau hubungan antarparameter yang saling memengaruhi. Kondisi ini dapat menyebabkan sistem berada dalam keadaan *out of control*, yang pada akhirnya berdampak pada penurunan efisiensi dan kualitas proses pembangkitan.

Untuk mengatasi tantangan ini, penelitian dilakukan dengan mengimplementasikan metode pengendalian kualitas berbasis diagram kontrol MAX-MEWMA menggunakan

residual model *Random Forest Regressor*. Pendekatan ini menawarkan kemampuan untuk memonitor parameter kritis HRSK secara simultan dengan sensitivitas yang lebih tinggi terhadap pergeseran rata-rata dan variabilitas proses. Dengan menghilangkan autokorelasi dalam data proses, metode ini mampu mendeteksi ketidaksesuaian proses lebih dini dan memberikan analisis yang lebih akurat mengenai penyebab utama kondisi *out of control*.

Melalui pendekatan ini, penelitian bertujuan untuk meningkatkan efisiensi operasional HRSK di PLTGU PT. PLN Nusantara Power UP Gresik. Dengan memahami faktor-faktor yang menyebabkan ketidaksesuaian proses dan menerapkan metode pengendalian kualitas yang canggih, diharapkan hasil penelitian ini tidak hanya mampu meningkatkan stabilitas operasional tetapi juga mendukung upaya pengurangan emisi karbon yang menjadi salah satu target utama dalam pengembangan energi berkelanjutan di Indonesia.

2.4 Uji Dependensi Antar Variabel

Uji dependensi antar variabel, seperti yang dijelaskan dalam literatur statistik, bertujuan untuk mengukur sejauh mana variabel-variabel dalam suatu dataset saling berhubungan atau tidak. Salah satu metode yang sering digunakan untuk tujuan ini adalah Uji *Bartlett's Sphericity*, yang dirancang untuk menguji apakah terdapat hubungan antar variabel dalam analisis multivariat. Jika variabel $X_1, X_2, \dots, X_p$ bersifat independen, maka matriks korelasi antar variabel akan sama dengan matriks identitas (**I**). Hipotesis yang digunakan dalam uji ini adalah sebagai berikut.

H_0 : Tidak terdapat korelasi antar variabel ($\rho = \mathbf{I}$)

H_1 : Terdapat korelasi antar variabel ($\rho \neq \mathbf{I}$)

Statistik uji:

$$X_{hitung}^2 = - \left[(n-1) - \frac{2p+5}{6} \right] \ln |\mathbf{R}| \quad (2.1)$$

Keterangan

n = jumlah pengamatan

p = jumlah karakteristik kualitas

$\mathbf{R}$ = matriks korelasi dari masing-masing karakteristik kualitas

$|\mathbf{R}|$ = determinan dari matriks korelasi

Daerah penolakan ditentukan dengan menolak H_0 jika $X_{hitung}^2 > X_{tabel}^2 \left(\alpha, \frac{1}{2} p(p-1) \right)$.

Hal ini menunjukkan bahwa terdapat korelasi di antara variabel-variabel tersebut. Matriks korelasi pada sampel dapat dihubungkan dengan matriks korelasi berdasarkan matriks kovarian (Morrison, 2005).

$$\mathbf{R} = \mathbf{r} \begin{pmatrix} 1 & r_{12} & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & r_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p1} & r_{p2} & \cdots & 1 \end{pmatrix} \quad (2.2)$$

$$r_{jh} = \frac{\sum_{i=1}^n \sum_{j=1}^p \sum_{h=1}^p (x_{ij} - \bar{x}_j)(x_{ih} - \bar{x}_h)}{\sqrt{\sum_{i=1}^n \sum_{j=1}^p \sum_{h=1}^p (x_{ij} - \bar{x}_j)^2 (x_{ih} - \bar{x}_h)^2}} \quad (2.3)$$

Keterangan

r_{ij} = nilai korelasi antara karakteristik kualitas ke- j dan karakteristik kualitas ke- h

x_{ij} = nilai observasi ke- i karakteristik kualitas ke- j , ($j=1, 2, \dots, p$)

x_{ij} = nilai observasi ke- i karakteristik kualitas ke- h , ($j=1, 2, \dots, p$), $j \neq h$

$\bar{x}_j$ = rata-rata karakteristik kualitas ke- j , ($j=1, 2, \dots, p$)

$\bar{x}_h$ = rata-rata karakteristik kualitas ke- h , ($h=1, 2, \dots, p$)

i = observasi ke- i ($i=1, 2, \dots, n$)

2.5 Uji Terasvirta

Uji Terasvirta merupakan metode untuk mendeteksi keberadaan non-linieritas dalam suatu model. Meskipun memiliki tujuan serupa dengan *white test*, keduanya memiliki pendekatan yang berbeda. Uji Terasvirta memanfaatkan model jaringan saraf (*neural network*) yang dikembangkan melalui pendekatan ekspansi Taylor dalam menentukan parameter modelnya. Sementara itu, *white test* menggunakan parameter yang ditentukan secara acak. Hal ini membuat Terasvirta lebih terarah dalam mengidentifikasi struktur non-linier dalam data (Prabowo, Suhartono, & Prastyo, 2020). Berikut merupakan contoh model yang bersifat non-linier.

$$y_t = \varphi(\gamma'w_t) + \beta'w_t + u_t, \quad (2.4)$$

dimana

- $\beta'w$: Komponen linier
- $\varphi(\gamma'w_t)$: Komponen non-linier
- γ' : Bobot dari model *neural network* dari lapisan *input* ke lapisan tersembunyi (*hidden layer*) untuk komponen non-linier
- β' : Bobot dari model *neural network* dari lapisan *input* ke lapisan *output* untuk komponen linier
- φ : Fungsi aktivasi sigmoid

persamaan (2.4) dapat ditulis ulang sebagai berikut.

$$y_t = \beta'w_t + \sum_{h=1}^q \theta_{0h} \left\{ \varphi(\gamma'w_t) - \frac{1}{2} \right\} + u_t. \quad (2.5)$$

Dimana θ_{0h} merupakan bobot dari model *neural network* dari lapisan tersembunyi (*hidden layer*) ke lapisan *output* untuk komponen non-linier. Apabila komponen non-linier adalah 0, maka data memiliki hubungan linier. Oleh karena itu, hipotesis nol (H_0) dirumuskan sebagai berikut.

$$H_0 : \theta_{01} = \theta_{02} = \dots = \theta_{0q} = 0 \text{ (model linier)}$$

$$H_1 : \text{Setidaknya ada satu } \theta_{0h} \neq 0, \text{ dimana } h = 1, 2, \dots, q \text{ (model non linierr)}$$

Nilai parameter *neural network* dalam uji terasvirta menggunakan ekspansi Taylor sehingga model baru diperoleh sebagai berikut.

$$y_t = \beta'w_t + \sum_{i=1}^q \sum_{j=1}^q \delta_{ij}y_{t-i}y_{t-j} + \sum_{i=1}^q \sum_{j=1}^q \sum_{k=1}^q \delta_{ijk}y_{t-i}y_{t-j}y_{t-k} + u_t. \quad (2.6)$$

Apabila komponen kuadrat (δ_{ij}) dan kubik (δ_{ijk}) adalah 0, maka gagal tolak H_0 sehingga didapatkan keputusan yaitu model linier. Selain itu, dengan menggunakan pendekatan distribusi *chi-square*, tolak H_0 ketika $\chi_{hitung}^2 > \chi_{(\alpha;m)}^2$ dengan m adalah jumlah prediktor tambahan yang ditambahkan ke model untuk menguji keberadaan hubungan nonlinier dimana m prediktor ini adalah komponen kuadrat dan kubik dari ekspansi Taylor. Dengan demikian, tolak H_0 ketika $\chi_{hitung}^2 > \chi_{(\alpha;m)}^2$ atau $p\text{-value} < \alpha$ sehingga didapatkan kesimpulan bahwa model nonlinier (Prabowo, Suhartono, & Prastyo, 2020).

2.6 Autocorrelation Function dan Partial Auto Correlatation Function

Dalam penelitian yang melibatkan analisis deret waktu pada sistem *Heat Recovery Steam Generator* (HRSG), seperti variabel *Fuel Gas Flow*, *Combustion Pressure*, dan *Exhaust Gas Temperature*, penting untuk memahami hubungan antara nilai-nilai observasi pada berbagai lag waktu. Analisis *Autocorrelation Function* (ACF) dan *Partial Autocorrelation Function* (PACF) digunakan untuk mengidentifikasi pola autokorelasi yang dapat memengaruhi performa model kontrol statistik multivariat, seperti Max-MEWMA.

Fungsi autokorelasi atau ACF merupakan salah satu cara untuk mengetahui kestasioneran data dalam mean, selain itu ACF juga digunakan untuk melihat hubungan linieritas yang terjadi antara pengamatan y_i dengan y_{i+k} . Korelasi antara y_i dengan y_{i+k} dinyatakan dalam persamaan berikut (Wei, 2006).

$$\hat{p}_k = \frac{\sum_{i=k+1}^n (y_i - \bar{y})(y_{i-k} - \bar{y})}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.7)$$

Keterangan

$\hat{p}_k$ = koefisien autokorelasi sampe pada lag ke-k

y_i = pengamatan pada waktu ke-i

$\bar{y}$ = rata rata pengamatan

y_{i-k} = pengamatan waktu ke i-k

Fungsi PACF sangat berguna dalam menentukan lag optimal yang akan digunakan sebagai input dalam pemodelan regresi deret waktu. Berdasarkan penelitian terdahulu, PACF digunakan untuk mengidentifikasi berapa banyak lag yang memiliki pengaruh signifikan terhadap nilai saat ini, sehingga dapat digunakan dalam pemodelan yang lebih akurat (Khusna, Mashuri, Suhartono, Prastyo, & Ahsan, 2018).

Partial Autocorrelation Function (PACF) digunakan untuk menghitung tingkat hubungan langsung antara nilai pengamatan y_i dan y_{i+k} dengan mengeliminasi pengaruh linier dari variabel-variabel sebelumnya, seperti $y_{i+1}, y_{i+2}, \dots, y_{i+k-1}$. Setelah pengaruh variabel-variabel tersebut dihilangkan, korelasi dihitung berdasarkan persamaan berikut.

$$\bar{\phi}_{k+1,k+1} = \frac{\hat{p}_{k+1} - \sum_{l=1}^k \hat{\phi}_{kl} \hat{p}_{k+1-l}}{1 - \sum_{l=1}^k \hat{\phi}_{kl} \hat{p}_l} \quad (2.8)$$

Dan

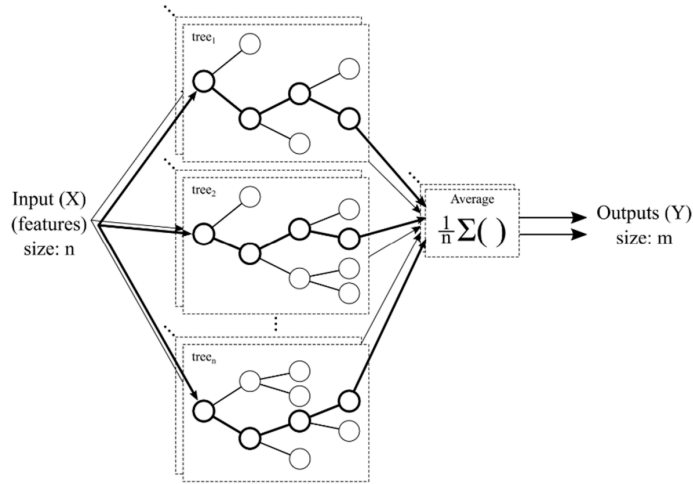
$$\bar{\phi}_{k+1,l} = \hat{\phi}_{kl} - \hat{\phi}_{k+1,k-1} \hat{\phi}_{k,k+1} \quad (2.9)$$

Dengan $l = 1, 2, \dots, k$.

Pada penelitian ini, PACF digunakan untuk menentukan lag yang digunakan sebagai variabel input dalam model *Random Forest Regressor*. Dengan memilih lag yang signifikan berdasarkan plot PACF, model dapat menangkap pola hubungan dalam data secara lebih akurat tanpa memasukkan terlalu banyak variabel yang tidak relevan

2.7 Random Forest Regressor

Random Forest (RF) merupakan metode ensemble learning berbasis decision tree yang dikembangkan oleh Breiman (2001), dan telah berkembang menjadi salah satu algoritma yang paling banyak digunakan untuk pemodelan regresi maupun klasifikasi, termasuk pada data multivariat dan time series. *Random Forest* membangun sejumlah pohon keputusan (*decision tree*) pada data hasil *bootstrap* sampling, kemudian merata-ratakan output pohon untuk regresi (*bagging*). Setiap pohon menggunakan subset acak dari fitur yang tersedia pada setiap proses *split node* (*random feature subsampling*), yang meningkatkan keragaman model dan kekebalan terhadap *overfitting* (Ho, 1998).



Gambar 2. 1 Struktur *Random Forest Regressor* multi-output (Parejo, et al., 2021).

Pada Gambar 2.1 tersebut, terlihat bahwa sejumlah pohon regresi dilatih pada data input dan output model diperoleh dengan cara melakukan rata-rata prediksi dari seluruh pohon. Pendekatan ini dapat diimplementasikan untuk prediksi satu maupun beberapa variabel target secara simultan. Dalam konteks pemantauan proses industri, pendekatan modern seperti residual-based monitoring telah banyak dikembangkan untuk mengatasi keterbatasan metode kontrol statistik klasik. Menurut ulasan oleh Wang, Qin, & Ma (2022), *Random Forest* dapat digunakan sebagai model prediktif non-linier untuk menghasilkan residual, yaitu selisih antara

nilai aktual dengan hasil prediksi model. Residual ini kemudian dimonitor menggunakan teknik kontrol statistik seperti diagram kontrol Max-MEWMA.

Pada regresi multivariat, seperti diuraikan oleh Linusson (2012) dan Segal (1992), fungsi *split* dalam pohon regresi dapat dirancang untuk mempertimbangkan lebih dari satu output target. Fungsi split tradisional (*univariate*) menggunakan *sum of squared errors*:

$$SS(t) = \sum_{i \in t} (y_i - \bar{y}(t))^2 \quad (2.10)$$

Namun untuk multivariat, digunakan *covariance-weighted sum of squared errors*:

$$SS(t) = \sum_{i \in t} (y_i - \bar{y}(t))' V^{-1}(t, \eta) (y_i - \bar{y}(t)) \quad (2.11)$$

dengan $V(t, \eta)$ adalah matriks kovariansi pada node t , dan parameter t mewakili struktur kovariansi antar output. Tujuan dari fungsi split adalah meminimalkan:

$$\phi(s, t) = (s, t) = SS(t) - SS(t_L) - SS(t_R) \quad (2.12)$$

dimana $SS(t_L)$ dan $SS(t_R)$ adalah *child nodes* hasil split. Dengan meminimalkan $\phi(s, t)$, pohon membentuk *node-node* yang homogen terhadap seluruh target sekaligus.

Pada penerapan *multi-output Random Forest*, pembobotan informasi dilakukan menggunakan *joint entropy* atau *information gain* pada setiap *node*, sebagaimana diadopsi dari Glocker, Pauly, Konukoglu, & Criminisi (2012) dan Linusson, Rudenwall, & Olausson (2012). Untuk regresi, *impurity node* diukur dengan *differential entropy*:

$$H_{r,j}(t) = - \int_{r \in \mathbb{R}^{n_{jy}}} p(r|x) \log p(r|x) dr \quad (2.13)$$

dan *entropy* ini dapat dinormalisasi terhadap *entropy* pada *root node*:

$$H_{norm,i}(t) = \frac{H_i(t)}{H_i(t_0)} \quad (2.14)$$

Sehingga *normalized joint entropy* seluruh target dapat dihitung sebagai rata-rata:

$$H(t) = \frac{1}{N} \sum_{i=1}^N H_{norm,i}(t) \quad (2.15)$$

dan *joint information gain*:

$$IG_i(t; f, x) = H_{norm,i}(t) - H_{norm,i}(t|f, x) \quad (2.16)$$

Split terbaik pada node adalah yang menghasilkan *information gain* maksimum terhadap seluruh target. Dalam penerapannya pada data time series, input model dapat berupa kombinasi lag signifikan dari masing-masing karakteristik kualitas. Model MRF kemudian dibentuk dari

data latih (X_f, Y_f) , residual dihitung dari data kontrol (X_c, Y_c) , dan pemantauan dilakukan terhadap data uji . Skema ini identik dengan tahapan pada proses pemodelan residual dalam penelitian ini. Pendekatan *residual-based monitoring* dapat digambarkan dalam bentuk berikut:

$$\hat{y}_i = f(x_i) \Rightarrow e_i = y_i - \hat{y}_i \quad (2.17)$$

Dengan monitoring residual e_i menggunakan diagram kontrol.

Kombinasi rentang parameter yang digunakan dalam proses tuning *Random Forest* pada penelitian ini, seperti jumlah pohon ($n_estimators$), kedalaman maksimum pohon (max_depth), minimum jumlah sampel pada *node split* ($min_samples_split$), dan minimum jumlah sampel pada daun ($min_samples_leaf$), diadopsi dari referensi utama Breiman (2001), Linusson (2012), serta rekomendasi praktik yang terdapat pada dokumentasi *scikit-learn* Pedregosa *et al.*, (2011). Rentang parameter yang dicoba umumnya meliputi $n_estimators$ antara 50 hingga 500, max_depth antara 3 hingga *None*, $min_samples_split$ antara 2 hingga 10, dan $min_samples_leaf$ antara 1 hingga 7. Penetapan rentang parameter ini juga mempertimbangkan hasil uji coba pendahuluan pada data *time series* multivariat yang digunakan dalam penelitian, sehingga diperoleh konfigurasi parameter yang dapat mengimbangi kompleksitas pola data, menjaga kemampuan generalisasi model, serta meminimalisasi risiko *overfitting*.

Nilai dari parameter-parameter tersebut ditentukan melalui proses pencarian kombinasi terbaik yang menghasilkan nilai *Mean Squared Error* (MSE) terkecil pada data pelatihan. Nilai MSE dihitung berdasarkan rumus berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.18)$$

Dimana adalah nilai aktual, adalah nilai prediksi, dan adalah jumlah total observasi.

Pemilihan parameter optimal pada *Random Forest* dilakukan melalui proses pencarian kombinasi terbaik yang menghasilkan nilai *Mean Squared Error* (MSE) terkecil pada data pelatihan. Proses pencarian ini dikenal sebagai *grid search*, di mana model dievaluasi pada setiap kombinasi parameter menggunakan validasi silang (*cross-validation*) khusus *time series*. Secara matematis, parameter optimal θ dapat didefinisikan sebagai berikut:

$$\theta = \arg \min_{\theta \in \Theta} \frac{1}{K} \sum_{k=1}^K MSE_{val,k}(\theta) \quad (2.19)$$

Setelah seluruh kombinasi diuji, parameter θ^* yang menghasilkan rata-rata MSE validasi terendah dipilih sebagai parameter optimal. Model *Random Forest* dengan parameter ini kemudian digunakan pada seluruh data pelatihan untuk tahap prediksi dan perhitungan residual selanjutnya. Dengan pemilihan parameter yang optimal, model mampu menghasilkan residual yang lebih representatif untuk digunakan dalam pemantauan proses menggunakan diagram kontrol Max-MEWMA.

Dengan pendekatan ini, *Random Forest* tidak hanya berfungsi sebagai alat prediksi, tetapi juga sebagai komponen penting dalam sistem pengendalian kualitas berbasis residual. Model ini memisahkan komponen deterministik dari proses, dan memungkinkan kontrol chart mendeteksi variasi acak yang tidak dapat dijelaskan oleh model.

Pendekatan ini memiliki beberapa keunggulan. Pertama, *Random Forest* mampu menangani data multivariat dan *time series* dengan kompleksitas tinggi. Kedua, metode ini tidak memerlukan asumsi kenormalan pada data asli, sehingga cocok digunakan untuk proses industri yang dinamis dan tidak terstruktur secara statistik. Ketiga, *Random Forest* menghasilkan residual yang dapat dimonitor secara statistik, khususnya dalam kerangka diagram kontrol multivariat. Keempat, model ini memungkinkan integrasi langsung dengan metode kontrol statistik seperti Max-MEWMA yang digunakan dalam penelitian ini.

Meskipun memiliki keunggulan dalam fleksibilitas dan akurasi, *Random Forest* tidak memberikan model parametris eksplisit, sehingga interpretasi terhadap hubungan antar variabel tidak semudah model regresi linier tradisional. Oleh karena itu, penggunaan RF lebih difokuskan pada efektivitas deteksi ketidakterkendalian dalam sistem produksi daripada interpretasi koefisien model.

Dengan demikian, penggunaan *Random Forest Regressor* multivariat dalam penelitian ini didasarkan pada rekomendasi kuat dari literatur baik secara praktis maupun teoritis, sebagai metode yang efektif dan fleksibel dalam monitoring proses industri modern.

2.8 Pengujian Asumsi Residual Independen

Pengujian asumsi residual independen dilakukan dengan mengecek adanya autokorelasi. Autokorelasi merupakan korelasi antara anggota serangkaian observasi yang diurutkan menurut waktu (seperti data deret waktu). Suatu data deret waktu dikatakan stasioner dalam *mean* jika tidak menunjukkan tren naik atau turun yang signifikan sehingga fluktuasinya cenderung berada di sekitar rata-rata yang stabil. Stasioneritas data dalam *mean* secara visual dapat dilihat menggunakan plot *Matrix Cross Correlation* (MCC) (Wei W. W., 2006). MCC memiliki kegunaan untuk mengidentifikasi stasioneritas data deret waktu multivariat dan membantu dalam menentukan urutan waktu pada *Vector Moving Average* (VMA). Vektor deret waktu dengan observasi $\epsilon_1, \epsilon_2, \epsilon_3, \dots, \epsilon_n$ mempunyai matriks korelasi yang ditunjukkan pada persamaan berikut ini.

$$\hat{\rho}(k) = [\hat{\rho}_{jl}(k)], \quad (2.20)$$

$\hat{\rho}_{jl}(k)$ adalah sampel *cross correlation* antara residual dari variabel j dan residual dari variabel l dengan lag k yang diperoleh dari persamaan berikut ini.

$$[\hat{\rho}_{jl}(k)] = \frac{\sum_{t=1}^{n-k} (\epsilon_{j,t} - \bar{\epsilon}_j)(\epsilon_{l,t+k} - \bar{\epsilon}_l)}{[\sum_{t=1}^n (\epsilon_{j,t} - \bar{\epsilon}_j)^2 \sum_{t=1}^n (\epsilon_{l,t} - \bar{\epsilon}_l)^2]^{\frac{1}{2}}}. \quad (2.21)$$

Dimana

- $\epsilon_{j,t}$: Residual ke- j ($j = 1, 2, \dots, p$) pada waktu ke- t ($t = 1, 2, \dots, n$)
- $\bar{\epsilon}_j$: Rata-rata residual ke- j
- $\epsilon_{l,t+k}$: Residual ke- l ($l = 1, 2, \dots, p$) dengan k -lag pada waktu ke- t
- $\bar{\epsilon}_l$: Rata-rata residual ke- l

Matrix cross-correlation function digunakan untuk mengidentifikasi orde model *Vector Moving Average*. Akan tetapi, pada penerapannya bentuk matriks dan grafik akan semakin kompleks karena meningkatnya dimensi vector sehingga untuk mengatasi hal tersebut Tiao dan

Box dalam Wei (2006) memperkenalkan metode yang mudah untuk merangkum korelasi sampel yaitu menggunakan koefisien dengan tanda (+), (-), dan (.) pada posisi ke- (j,l) dari matriks *cross-correlation*. Berikut ini merupakan penjelasan terkait koefisien tanda tersebut.

1. Tanda (+) mengartikan bahwa nilai korelasi sampel ($\hat{\rho}_{jl}(k)$) yang lebih besar dari 2 kali estimasi standar error. Artinya terdapat hubungan korelasi positif antara kualitas ke- j dan l .
2. Tanda (-) mengartikan bahwa nilai korelasi sampel ($\hat{\rho}_{jl}(k)$) yang lebih kecil dari 2 kali estimasi standar error. Artinya terdapat hubungan korelasi negatif antara kualitas ke- j dan l .
3. Tanda (.) mengartikan bahwa nilai korelasi sampel ($\hat{\rho}_{jl}(k)$) terletak diantara ± 2 kali estimasi standar error. Artinya tidak terdapat hubungan korelasi antara kualitas ke- j dan l .

Selain dilakukan pengecekan dengan plot MCCF, pengujian asumsi residual independen juga bisa dilakukan dengan pengujian *Portmanteau* untuk menguji asumsi *white noise*. Uji *Portmanteau* merupakan generalisasi dari uji Ljung Box untuk kasus multivariat dengan hipotesis sebagai berikut.

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

$$H_1 : \text{Minimal terdapat satu } \rho_l \neq 0, l = 1, 2, 3, \dots, s$$

Statistik uji yang digunakan adalah sebagai berikut.

$$Q_s = n^2 \sum_{i=1}^s \frac{1}{n-i} \text{tr}(\hat{\mathbf{r}}_t \hat{\mathbf{r}}_0^{-1} \hat{\mathbf{r}}_t' \hat{\mathbf{r}}_0^{-1}). \quad (2.22)$$

Dimana

- s : Jumlah lag yang digunakan dalam pengujian
- n : Banyaknya pengamatan
- $\hat{\mathbf{r}}_t'$: Matriks kovarians silang pada waktu t
- $\hat{\mathbf{r}}_0$: Matriks kovarians silang pada waktu 0

Untuk kriteria keputusannya yaitu gagal tolak H_0 apabila statistik uji $Q_s < \chi_{k^2(s-p)}^2$ atau $p\text{-value} > \alpha$ yang memiliki arti bahwa residual memenuhi asumsi *white noise* (Lutkepohl, 2005).

2.9 Pengujian Asumsi Residual Normal Multivariat

Uji asumsi residual normalitas multivariat bertujuan untuk mengevaluasi apakah data dalam dimensi yang lebih tinggi (≥ 2 variabel) mengikuti distribusi normal multivariat. Uji normalitas multivariat sangat penting terutama dalam penerapan diagram kendali multivariat karena beberapa diagram kendali multivariat bergantung pada asumsi normalitas data yang digunakan. Suatu residual $\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \boldsymbol{\varepsilon}_3, \dots, \boldsymbol{\varepsilon}_p$ dikatakan berdistribusi normal multivariat ketika memiliki fungsi densitas sebagai berikut (Johnson & Wichern, 2007).

$$f(\boldsymbol{\varepsilon}_1, \boldsymbol{\varepsilon}_2, \dots, \boldsymbol{\varepsilon}_p) = \frac{1}{\sqrt{(2\pi)^p |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2} [(\boldsymbol{\varepsilon}_j - \boldsymbol{\mu})' \boldsymbol{\Sigma}^{-1} (\boldsymbol{\varepsilon}_j - \boldsymbol{\mu})]\right). \quad (2.23)$$

Dimana

ϵ_j = Vektor residual karakteristik kualitas ke- j , $j = 1, 2, 3, \dots, p$

μ = Vektor rata-rata residual untuk semua karakteristik kualitas

Σ^{-1} = Invers dari matriks kovarian sampel

p = Banyaknya karakteristik kualitas

Salah satu metode yang sering digunakan untuk melakukan pengujian distribusi normal multivariat adalah uji *Shapiro-Wilk* dengan pengujian hipotesis sebagai berikut.

H_0 : Data pengamatan berdistribusi normal multivariat

H_1 : Data pengamatan tidak berdistribusi normal multivariat

Statistik uji pada uji *Shapiro-Wilk* dapat dilihat pada persamaan 2.6

$$W^* = \frac{1}{p} \sum_{j=1}^p W_j \quad (2.24)$$

dimana

$$W_j = \frac{(\sum_{i=1}^n a_i y_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (2.25)$$

Dengan nilai a merupakan koefisien tabel *normality test*.

Dengan taraf signifikansi α , Gagal tolak H_0 atau dapat disimpulkan data pengamatan berdistribusi normal multivariat apabila nilai $W^* > C_{\alpha;n;p}$ atau $p - value < \alpha$ dengan $C_{\alpha;n;p}$ merupakan nilai quantil statistik *Shapiro Wilk* (Shapiro & Wilk, 1965).

2.10 Maximum Exponentially Weighted Moving Average Chart

Dalam metode ini, asumsi bahwa sebuah proses terdiri dari p karakteristik kualitas dilambangkan dengan X , dengan $X \sim N_k(\mu, \Sigma)$ dan $X_{i2}, \dots, X_{in}$, $i = 1, 2, \dots, n$; i adalah ukuran sampel n yang dibuat dari proses. Misalkan μ_0 dan Σ_0 menjadi vektor rata-rata dan matriks kovarian yang diinginkan masing-masing. Asumsikan bahwa vector acak X_{ij} tidak saling tergantung satu sama lain, baik di dalam sampel dan di antara sampel.

Pertama, untuk memantau vektor rata-rata proses. Diketahui $\bar{X}_i = n^{-1} \sum_{j=1}^n X_{ij}$ menjadi vektor rata-rata sampel per subgroup. Dengan demikian, nilai statistik Z_i pada diagram pengendalian MEWMA didefinisikan sebagai berikut.

$$Z_i = (1 - \lambda)Z_{i-1} + \lambda(\bar{X}_i - \mu_0) \quad (2.26)$$

Di mana Z_0 merupakan titik awal, dan μ_0 adalah vektor rata-rata keseluruhan. Parameter λ adalah smoothing parameter dengan rentang nilai $0 \leq \lambda \leq 1$.

Karena

$$E(Z_i) = \mu - \mu_0 \quad (2.27)$$

Dan

$$Cov(\mathbf{Z}_i) = \frac{\lambda[1 - (1 - \lambda)^{2i}]}{n(2 - \lambda)} \Sigma \quad (2.28)$$

Dan

$$\mathbf{Z}_i = \frac{\lambda[1 - (1 - \lambda)^{2i}]}{n(2 - \lambda)} \Sigma \quad (2.29)$$

Karena itu,

$$\mathbf{Z}_i \sim N_k(\boldsymbol{\mu} - \boldsymbol{\mu}_0, Cov(\mathbf{Z}_i)) \quad (2.30)$$

Dan

$$T_i = \frac{n(2 - \lambda)}{\lambda[1 - (1 - \lambda)^{2i}]} \mathbf{Z}_i' \Sigma^{-1} \mathbf{Z}_i \sim X_{k, \delta^2}^2 \quad (2.31)$$

Di mana $Cov(.)$ merepresentasikan fungsi kovarians dari distribusi multivariat, Σ^{-1} adalah invers dari Σ , sedangkan p dan δ^2 masing-masing menunjukkan derajat kebebasan dan parameter *non-centrality* dari distribusi *chi-square non-sentral*. Nilai δ^2 dihitung sebagai $\left\{ \frac{n}{2-\lambda} [\lambda[1 - (1 - \lambda)^{2i}]] \right\} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)' \Sigma^{-1} (\boldsymbol{\mu} - \boldsymbol{\mu}_0)$.

Statistik pada diagram pengendalian MEWMA yang digunakan untuk memantau proses vektor rata-rata didefinisikan sebagai berikut.

$$U_i = \Phi^{-1} \left[H_p \left\{ \frac{n(2 - \lambda)}{\lambda[1 - (1 - \lambda)^{2i}]} \mathbf{Z}_i' \Sigma_0^{-1} \mathbf{Z}_i \right\} \right] \quad (2.32)$$

Di mana $H_p(.)$ merupakan fungsi distribusi chi-square dengan p sebagai derajat kebebasan. $\Phi(.)$ adalah fungsi distribusi normal standar, dan Φ^{-1} merupakan inversi dari $\Phi(.)$. Untuk memantau diagram pengendalian MEWMA pada proses variabilitas, didefinisikan sebagai berikut.

$$W_i = \Sigma_{j=1}^n (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i)' \Sigma_0^{-1} (\mathbf{X}_{ij} - \bar{\mathbf{X}}_i) \quad (2.33)$$

Dengan demikian, $W_i \sim X_k^2, (n - 1)$ ketika $\Sigma = \Sigma_0$ dan nilai statistik Y_i pada diagram pengendalian MEWMA dirumuskan sebagai berikut.

$$Y_i = (1 - \lambda)Y_{i-1} + \lambda \Phi^{-1} \{ H_{p, (n-1)}(W_i) \} \quad (2.35)$$

Di mana Y_0 merupakan titik awal yang bernilai nol. Parameter λ adalah smoothing parameter dengan rentang nilai . Ketika variabilitas proses berada dalam kondisi terkendali ($\Sigma = \Sigma_0$) dan $\mathbf{Y}_0 = \mathbf{0}$, maka hasilnya dapat dirumuskan sebagai berikut.

$$E(Y_i) = 0 \quad (2.36)$$

Dan

$$Var(Y_i) = \frac{\lambda[1 - (1 - \lambda)^{2k}]}{2 - \lambda} \quad (2.37)$$

Statistik pada diagram pengendalian MEWMA yang digunakan untuk memantau variabilitas proses dirumuskan sebagai berikut.

$$V_i = \sqrt{\frac{2 - \lambda}{\lambda[1 - (1 - \lambda)^{2i}]}} \cdot Y_i \quad (2.38)$$

Diketahui bahwa statistik U_i dan V_i bersifat independen satu sama lain. Ketika $\mu = \mu_0$, Z_0 , dan Y_0 , maka kedua statistik U_i dan V_i mengikuti distribusi normal standar. Oleh karena itu, berdasarkan U_i dan V_i , sebuah statistik baru untuk *single chart* multivariat dirumuskan sebagai berikut.

$$M_i = \max \{|U_i|, |V_i|\} \quad (2.39)$$

Karena M_i merupakan nilai maksimum dari U_i dan V_i , yang didasarkan pada nilai absolut dari dua statistik *Multivariate Weighted Exponential Statistic Moving Average* (MEWMA), maka wajar jika grafik ini disebut sebagai grafik Max-MEWMA (Chen, Cheng, & Xie, 2005). Berdasarkan statistik U_i dan V_i , yang mengikuti distribusi normal standar, sesuai dengan referensi (Montgomery, 2020), MEWMA memiliki kekuatan untuk tetap *robust* terhadap asumsi distribusi normal. Nilai M_i yang besar pada grafik Max-MEWMA menunjukkan bahwa vektor rata-rata dan/atau matriks kovarians dari proses telah menyimpang dari μ_0 . Sebaliknya, nilai M_i yang kecil menunjukkan bahwa vektor rata-rata dan variabilitas proses tetap mendekati nilai nominalnya.

Karena M_i bernilai non-negatif, diagram kontrol Max-MEWMA hanya bergantung pada batas kendali atas (UCL) untuk kondisi awal. Jika $M_i > UCL$, maka diagram kontrol akan memberikan sinyal alarm bahwa proses berada di luar kendali.

i. Batas Kendali dari Max-MEWMA

Karena U_i dan V_i bersifat independen dan mengikuti distribusi normal standar, dengan asumsi $\mu = \mu_0$ dan $\Sigma = \Sigma_0$, maka *Cumulative Distribution Function* (CDF) dalam kondisi *in-control* dari M_i dirumuskan sebagai berikut.

$$\begin{aligned} f(y, \mu_0, \Sigma_0) &= P(M_i \leq y | \mu = \mu_0, \Sigma = \Sigma_0) \\ &= P(|U_i| \leq y, |V_i| \leq y) \\ &= P(|U_i| \leq y)P(|V_i| \leq y) \\ &= [2\Phi(y) - 1]^2, \quad y \leq 0 \end{aligned} \quad (2.40)$$

Probability density function (PDF) dari M_i dalam kondisi terkendali adalah sebagai berikut.

$$f(y, \mu_0, \Sigma_0) = 4\phi(y)[2\Phi(y) - 1]^2 \quad (2.41)$$

Dengan $\phi(\cdot)$ sebagai fungsi distribusi kumulatif normal standar. Rumus $\phi(\cdot)$, $E(M_i)$, $Var(M_i)$ adalah sebagai berikut.

$$\begin{aligned}\phi &= F(y, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) = P(X \leq x) \\ &= \int_{-\infty}^x f(x, \mu_x, \sigma_x) dt\end{aligned}\tag{2.42}$$

$$\begin{aligned}&= \int_{-\infty}^x \frac{1}{\sigma_x \sqrt{2\pi}} \exp\left(\frac{(t - \mu_x)^2}{2\sigma_x^2}\right) dt \\ E(M_i) &= \int_0^{\infty} f(x, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) dy \\ &= \int_0^{\infty} y^2 \cdot 4\phi(y) [2\Phi(y) - 1]^2 dy\end{aligned}\tag{2.43}$$

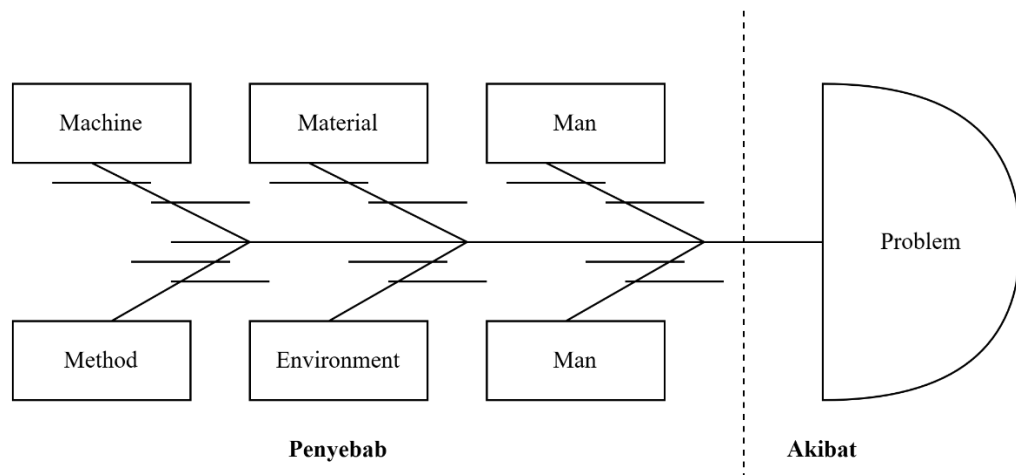
$$\begin{aligned}Var(M_i) &= \int_0^{\infty} y^2 f(y, \boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0) dy \\ &= \int_0^{\infty} y^2 \cdot 4\phi(y) [2\Phi(y) - 1]^2 dy\end{aligned}\tag{2.44}$$

Diagram kontrol Max-MEWMA hanya memerlukan batas kendali atas atau biasanya dikenal dengan *upper control limit (UCL)*. Dalam penelitian ini, penentuan UCL dilakukan menggunakan simulasi *bootstrap*. Penetapan UCL berbasis bootstrap menggunakan target ARL sebesar 370 dan α sebesar 0,0027. Langkah pertama dalam menentukan UCL adalah menghitung nilai statistik M_i dengan mengambil sampel data sesuai persamaan (2.39). Selanjutnya, dilakukan iterasi sebanyak N kali ($l = 1, 2, \dots, N$, dengan $n = 100$ iterasi melibatkan *resampling* statistik M_i sebanyak B kali (jumlah B sama dengan jumlah statistik M_i yang telah dihitung di langkah pertama). Dari hasil tersebut, dicari nilai *percentile* ke-100(1- α)% dari distribusi statistik M_i . Langkah terakhir, nilai UCL dihitung sebagai rata-rata dari *percentile* yang didapatkan di langkah kedua, yaitu:

$$UCL = \frac{1}{N} \sum_{l=1}^N M_{(100(1-\alpha))}^l\tag{2.45}$$

2.11 Diagram Ishikawa

Diagram *Ishikawa*, atau yang dikenal juga sebagai diagram sebab-akibat, adalah alat analisis mutu yang bertujuan untuk memahami hubungan menyeluruh antara kecacatan dan penyebabnya. Diagram ini unik karena menggunakan data verbal (non-numerik) atau data kualitatif dalam penyajiannya. Alat ini memberikan gambaran tentang kondisi "penyimpangan mutu" yang dipengaruhi oleh berbagai penyebab yang saling terkait (Montgomery, 2020).



Gambar 2. 2 Diagram *Ishikawa*

Diagram *Ishikawa* mencakup beberapa kategori utama, yaitu Manusia (*Man*), Mesin (*Machine*), Material (*Material*), Metode (*Method*), Lingkungan (*Environment*), dan Pengukuran (*Measurement*). Penjelasan masing-masing kategori adalah sebagai berikut.

- Manusia (*Man*), karyawan atau operator memiliki peran besar dalam pelaksanaan proses produksi karena mereka secara langsung berinteraksi dengan mesin dan bahan baku. Tingkat kedisiplinan dan keterampilan operator menjadi faktor penting yang harus diperhatikan.
- Mesin (*Machine*), mesin merupakan elemen penting yang terhubung langsung dengan bahan baku yang diolah. Proses pengaturan mesin juga menjadi aspek signifikan yang perlu diperhatikan untuk memastikan kelancaran produksi.
- Bahan baku (*Material*), faktor material memiliki pengaruh besar terhadap hasil akhir produk. Kualitas bahan baku yang buruk dapat menyebabkan cacat pada produk.
- Metode (*Method*), kesalahan dalam metode kerja dapat menghasilkan produk yang cacat atau tidak sesuai dengan standar yang ditetapkan.
- Lingkungan (*Environment*), kondisi lingkungan berpotensi memengaruhi bahan baku yang diproses maupun kondisi karyawan selama proses produksi, sehingga dapat berdampak pada hasil akhir.
- Pengukuran (*Measurement*), akurasi dalam pengukuran menjadi faktor krusial, karena kesalahan pada tahap ini dapat menyebabkan produk cacat atau tidak memenuhi spesifikasi yang diinginkan.

2.12 Analisis Kapabilitas Proses

Kapabilitas proses mengacu pada kemampuan suatu proses dalam beroperasi sesuai dengan standar yang telah ditetapkan sebelumnya (Montgomery D., 2020). Analisis kapabilitas proses digunakan untuk mengevaluasi sejauh mana suatu proses dapat memenuhi batas toleransi yang telah ditentukan. Sebuah produk dianggap memiliki kapabilitas jika memenuhi beberapa kriteria berikut:

- Proses berada dalam kondisi yang terkendali secara statistik, yang dapat diamati melalui diagram kendali yang dibuat berdasarkan checksheet. Suatu proses dikatakan terkendali apabila seluruh data pengamatan berada dalam batas kendali (*in control*).
- Proses sesuai dengan batas spesifikasi yang telah ditentukan oleh perusahaan, baik berdasarkan standar alat, kebijakan internal, maupun permintaan pelanggan.

3. Proses memiliki tingkat presisi yang tinggi, yaitu nilai pengamatan yang konsisten dan tidak menyebar jauh satu sama lain.
4. Proses memiliki akurasi yang tinggi, yang berarti hasil pengamatan mendekati target atau nilai yang diharapkan.

Dalam analisis kapabilitas proses, terdapat beberapa indeks yang digunakan, yaitu C_p , C_{pk} , P_p dan P_{pk} . Perbedaan utama antara indeks C dan P terletak pada cakupan data yang digunakan. Indeks C berfokus pada sampel data, sedangkan indeks P mempertimbangkan keseluruhan populasi data. Selain itu, indeks C digunakan untuk mengevaluasi kapabilitas proses dalam jangka panjang, sementara indeks P lebih relevan untuk analisis dalam jangka pendek.

Menurut *Automotive Industry Action Group* (AIAG), indeks C_p dan C_{pk} sebaiknya digunakan ketika proses telah terkendali secara statistik (*in control*). Sebaliknya, jika proses belum dalam kondisi terkendali, maka indeks yang lebih sesuai adalah P_p dan P_{pk} (Montgomery, 2020). Perhitungan indeks C_p dan C_{pk} dapat dirujuk pada persamaan 2.38 dan 2.39.

$$C_p = \frac{USL - LSL}{6\sigma} \quad (2.46)$$

Dimana USL adalah batas spesifikasi atas, LSL adalah batas spesifikasi bawah, dan σ adalah standar deviasi proses. Indeks C_{pk} dihitung menggunakan rumus.

$$C_{pk} = \min(C_{pu}, C_{pl}) \quad (2.47)$$

Dengan

$$C_{pu} = \frac{USL - \mu}{3\sigma}, C_{pl} = \frac{\mu - LSL}{3\sigma} \quad (2.48)$$

Menurut Montgomery (2020) hasil analisis kapabilitas dapat dibagi menjadi tiga kondisi:

1. Jika $C_{pk} < 1$, proses tidak kapabel karena tingkat presisi rendah, dan banyak produk tidak sesuai spesifikasi.
2. Jika $C_{pk} \geq 1$, proses dianggap kapabel karena menghasilkan produk sesuai spesifikasi dengan tingkat akurasi dan presisi yang memadai.
3. Jika proses berada di luar kendali statistik, maka evaluasi diperlukan untuk mengidentifikasi penyebab dan memperbaiki kapabilitas proses.

Karena penelitian ini melibatkan lebih dari satu karakteristik kualitas, analisis kapabilitas proses dilakukan secara multivariat. Perhitungan indeks kapabilitas dalam konteks ini dapat dilakukan dengan merujuk pada persamaan 2.41 dan 2.42 (Raissi, 2009).

$$MC_p = \sum_{k=1}^n D_k C_p(Y_k) \quad (2.49)$$

$$MC_{pk} = \sum_{k=1}^n D_k C_{pk}(Y_k) \quad (2.50)$$

MC_p dan MC_{pk} merupakan versi multivariat dari indeks C_p dan C_{pk} , dengan D_k sebagai bobot untuk setiap karakteristik kualitas yang ditentukan oleh perusahaan. Jika bobot spesifik untuk masing-masing karakteristik kualitas tidak tersedia, maka bobot dianggap bernilai sama.

2.13 Analisis Efisiensi *Heat Recovery Steam Generator* (HRSG)

Dalam suatu sistem, analisis difokuskan pada area di mana materi dan energi bergerak melaluinya. Persamaan kesetimbangan laju energi dalam sistem tersebut dapat dinyatakan sebagai berikut:

$$\Sigma_i \dot{m}_i = \Sigma_o \dot{m}_o \quad (2.51)$$

Dengan

$\Sigma_i \dot{m}_i$: jumlah laju aliran massa masuk ke sistem, (kg/s)

$\Sigma_o \dot{m}_o$: jumlah laju aliran massa keluar ke sistem, (kg/s)

Gas buang merupakan gas hasil dari proses pembakaran yang memiliki suhu lebih tinggi dibandingkan suhu atmosfer. Pada proses pembakaran, bahan bakar bereaksi dengan udara dan menghasilkan produk berupa gas buang yang mengandung berbagai senyawa, seperti H_2O , CO_2 , dan N_2 , serta O_2 sebagai tambahan jika suplai udara berlebih. Energi panas yang terkandung dalam gas buang yang disalurkan ke HRSG ($\dot{Q}_{eg}$) dapat dihitung menggunakan persamaan berikut.

$$\dot{Q}_{eg} = \dot{m}_{eg} C_{p_{eg}} (T_i - T_o) \quad (2.52)$$

Dengan

T_i : temperatur gas buang (K).

T_o : temperature lingkungan (K).

$\dot{m}_{eg}$: laju aliran massa gas buang (kg/dt).

$C_{p_{eg}}$: panas spesifik gas buang (kJ/kg.K).

n, m : jumlah mol konstituen.

laju aliran massa udara yang diperlukan dapat diketahui dengan persamaan berikut :

$$\dot{m}_o = AFR \times \dot{m}_f \quad (2.53)$$

Dengan

AFR : perbandingan udara dan bahan bakar

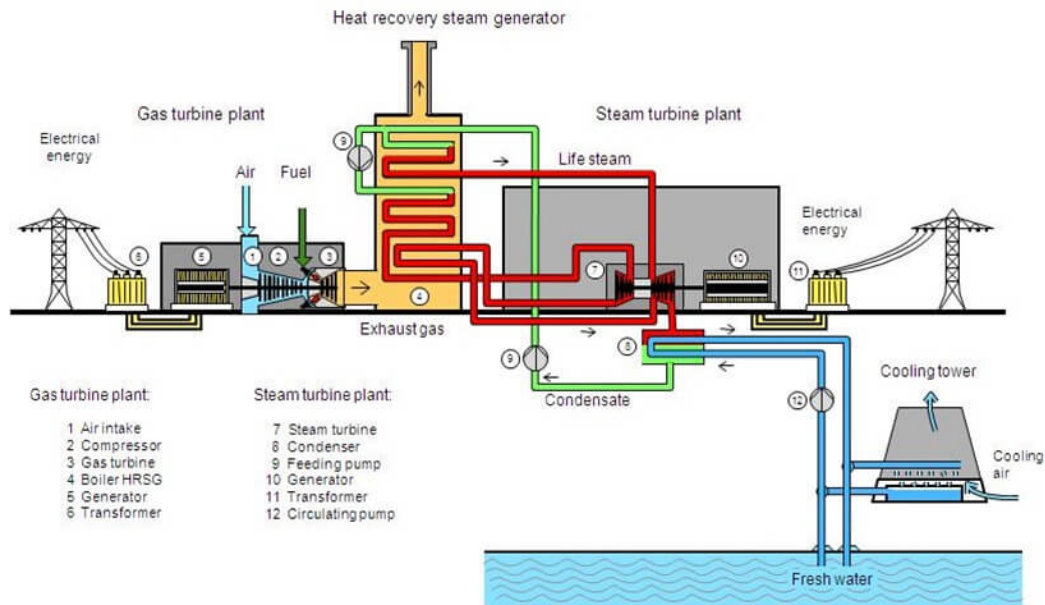
$\dot{m}_f$: laju aliran massa bahan bakar (kg/dt)

laju aliran massa gas buang dapat diketahui dengan persamaan berikut ini :

$$\dot{m}_{eg} = \Sigma(\dot{m}_f + \dot{m}_o) \quad (2.54)$$

Besarnya efisiensi thermal HRSG didefinisikan sebagai perbandingan antara laju energi yang dibutuhkan air menjadi uap panas lanjut (*superheated*) dengan laju aliran energi gas buang didalam HRSG tersebut

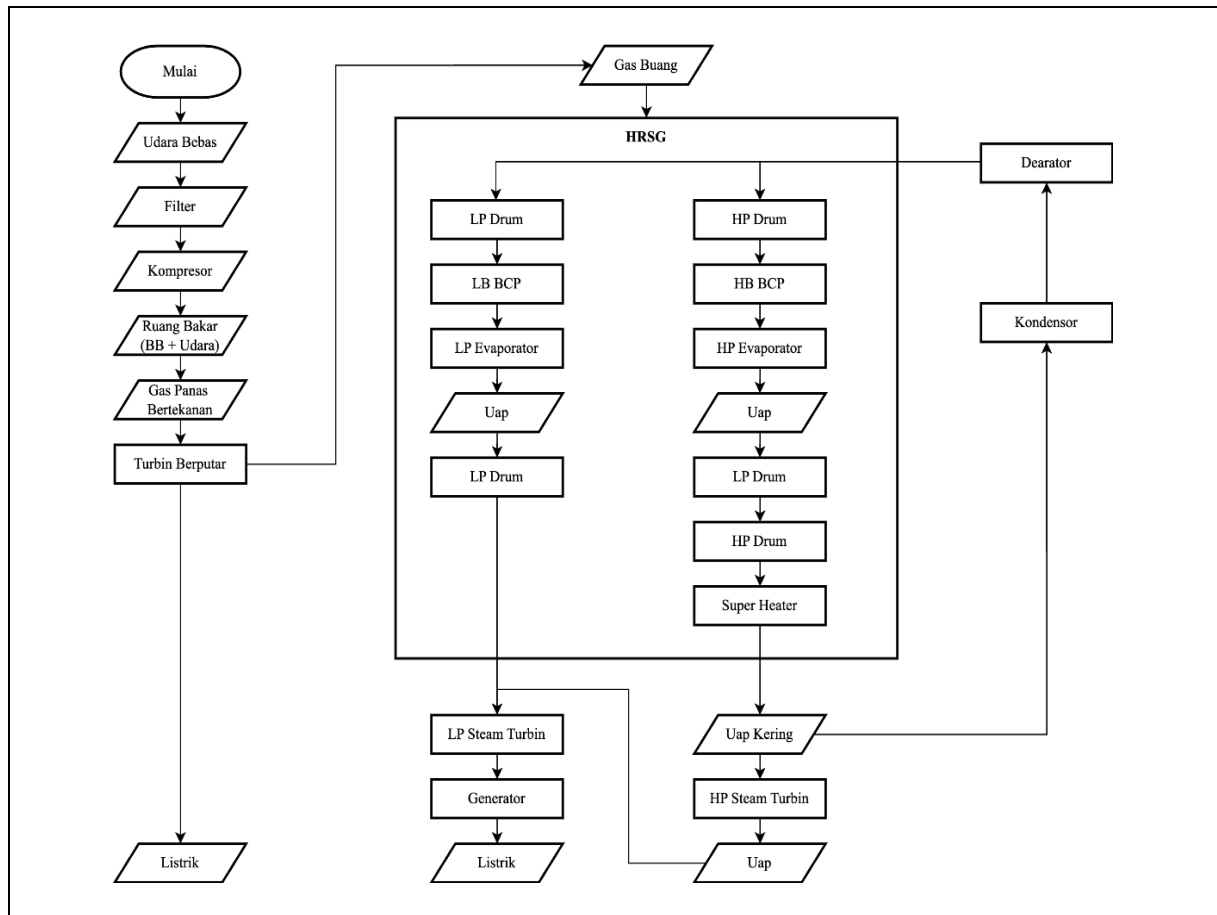
2.14 Sistem Kerja *Heat Recovery Steam Generator* (HRSG)



Gambar 2. 3 Bagan siklus kerja PLTGU.

PLTGU adalah pembangkit listrik yang menggabungkan Turbin Gas dan Turbin Uap. Kombinasi ini memanfaatkan suhu tinggi dari gas buang hasil pembakaran Turbin Gas. Panas dari gas buang tersebut dipindahkan sepenuhnya melalui proses konveksi ke air di dalam pipa-pipa HRSG. HRSG, atau *Heat Recovery Steam Generator*, merupakan komponen yang menghubungkan turbin gas dengan turbin uap dalam sistem *combine cycle power plant*. Fungsi utama HRSG adalah memanfaatkan panas dari gas buang turbin gas untuk memanaskan air dalam pipa-pipa di HRSG hingga menghasilkan *superheated steam* yang mampu menggerakkan turbin uap.

Laju temperatur gas buang di HRSG terus berubah seiring waktu. Hal ini disebabkan oleh variasi laju aliran massa udara dan bahan bakar yang masuk ke turbin gas, yang mengakibatkan perbedaan laju aliran massa dan suhu gas buang. Akibatnya, energi panas dari gas buang yang diserap oleh air untuk menghasilkan *superheated steam* juga mengalami fluktuasi. Kondisi ini memengaruhi efisiensi HRSG, yang berubah-ubah sesuai dengan kondisi operasional saat itu. Perubahan ini tidak hanya berdampak pada performa perpindahan panas, tetapi juga dapat menimbulkan ketidakterkendalian pada proses produksi uap. Jika tidak dimonitor dengan cermat, fluktuasi ini berpotensi menyebabkan deviasi yang signifikan terhadap target output energi maupun tekanan operasi. Oleh karena itu, penting untuk melakukan pemantauan kontinu terhadap parameter-parameter kritis guna menjaga stabilitas dan keandalan sistem HRSG dalam jangka panjang.



Gambar 2. 4 Diagram Alir Proses Produksi Listrik di PLTGU.

Pada Gambar 2.4, ditampilkan siklus kerja HRSG yang dapat dijelaskan sebagai berikut:

1. Udara bebas masuk dan disaring menggunakan filter untuk menghindari partikel kecil sebelum masuk ke kompresor.
2. Kompresor menghisap udara, kemudian udara bertekanan tersebut dialirkan ke ruang bakar.
3. Udara bertekanan bercampur dengan gas alam di ruang bakar, yang kemudian dibakar sehingga menghasilkan gas bertekanan tinggi yang diarahkan ke sudu-sudu turbin.
4. Turbin berputar akibat dorongan gas panas yang diarahkan ke sudu-sudu turbin, sehingga menggerakkan generator untuk menghasilkan energi listrik.
5. Gas buang dari turbin gas yang masih memiliki suhu tinggi dialirkan ke HRSG untuk memanaskan air di dalam pipa-pipa pemanas, sebelum akhirnya gas buang tersebut dilepaskan ke atmosfer melalui cerobong dengan suhu yang lebih rendah.
6. Uap kering (*superheated steam*) yang dihasilkan dari HRSG diarahkan untuk memutar turbin uap.
7. Generator yang digerakkan oleh turbin uap menghasilkan energi listrik.
8. Uap kering dari turbin uap kemudian didinginkan di kondensor menggunakan air laut, menghasilkan air yang kemudian disalurkan kembali ke drum untuk disirkulasikan ke HRSG.

(halaman ini sengaja dikosongkan)

BAB III METODOLOGI PENELITIAN

3.1 Sumber Data

Penelitian ini menggunakan data sekunder yang diperoleh dari Departemen Operasional PLTGU PT PLN Nusantara Power UP Gresik. Data tersebut menggunakan pola operasi (1,1,1), yaitu 1 turbin gas, 1 HRSG, dan 1 turbin uap, di PLTGU PT PLN Nusantara Power UP Gresik. Unit penelitian adalah PLTGU, dengan fokus pada parameter fuel gas flow, combustion pressure, dan gas temperature. Pengamatan terhadap ketiga karakteristik kualitas tersebut dilakukan setiap menit. Data yang digunakan dalam penelitian ini dibagi menjadi dua fase. Fase I yaitu pada tanggal 1 Juli – 6 Juli 2022 dan Fase II pada tanggal 6 Juli 2022 – 7 Juli 2022.

3.2 Variabel Penelitian

Variabel yang dikaji dalam penelitian ini merupakan faktor-faktor material yang memengaruhi kestabilan sistem HRSG. Adapun variabel – variabelnya sebagai berikut :

Tabel 3. 1 Variabel Penelitian

Variabel	Karakteristik Kualitas	Satuan	Batas Spesifikasi
Y_1	<i>Fuel Gas Flow</i>	KNM/h	20,2 – 32,73
Y_2	<i>Combustion Pressure</i>	Kg/cm	8,7 – 12,5
Y_3	<i>Gas Temperature</i>	Celcius	481 – 523

Penjelasan untuk masing-masing variable penelitian pada Tabel 3.1 adalah sebagai berikut.

1. *Fuel Gas Flow* (Y_1)
Laju aliran bahan bakar gas dalam sistem pembakaran turbin gas yang memengaruhi temperatur gas buang, digunakan untuk pemanasan HRSG.
2. *Combustion Pressure* (Y_2)
Tekanan dalam sistem pembakaran turbin gas yang berpengaruh pada hasil temperatur gas buang untuk pemanasan HRSG.
3. *Gas Temperature* (Y_3)
Temperatur gas buang yang digunakan untuk memanaskan air menjadi uap, yang harus berada dalam batas spesifikasi tertentu agar sistem HRSG tetap stabil.

3.3 Struktur Data

Struktur data yang digunakan dalam penelitian ini disajikan dalam Tabel 3.2. Data terdiri dari pengamatan terhadap karakteristik kualitas yang mencakup Y_1 (*Fuel Gas Flow*), Y_2 (*Combustion Pressure*), dan Y_3 (*Gas Temperature*) pada setiap pengamatan, sebagai berikut:

Tabel 3. 2 Struktur Data Penelitian

Pengamatan	Karakteristik Kualitas		
	y_1	y_2	y_3
1	$y_{1.1}$	$y_{1.2}$	$y_{1.3}$
2	$y_{2.1}$	$y_{2.2}$	$y_{2.3}$
3	$y_{3.1}$	$y_{3.2}$	$y_{3.3}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$y_{n.1}$	$y_{n.2}$	$y_{n.3}$

Keterangan

n = jumlah observasi

$y_{n.m}$ = jumlah pengamatan ke- n dan karakteristik kualitas ke- m

3.4 Langkah Analisis

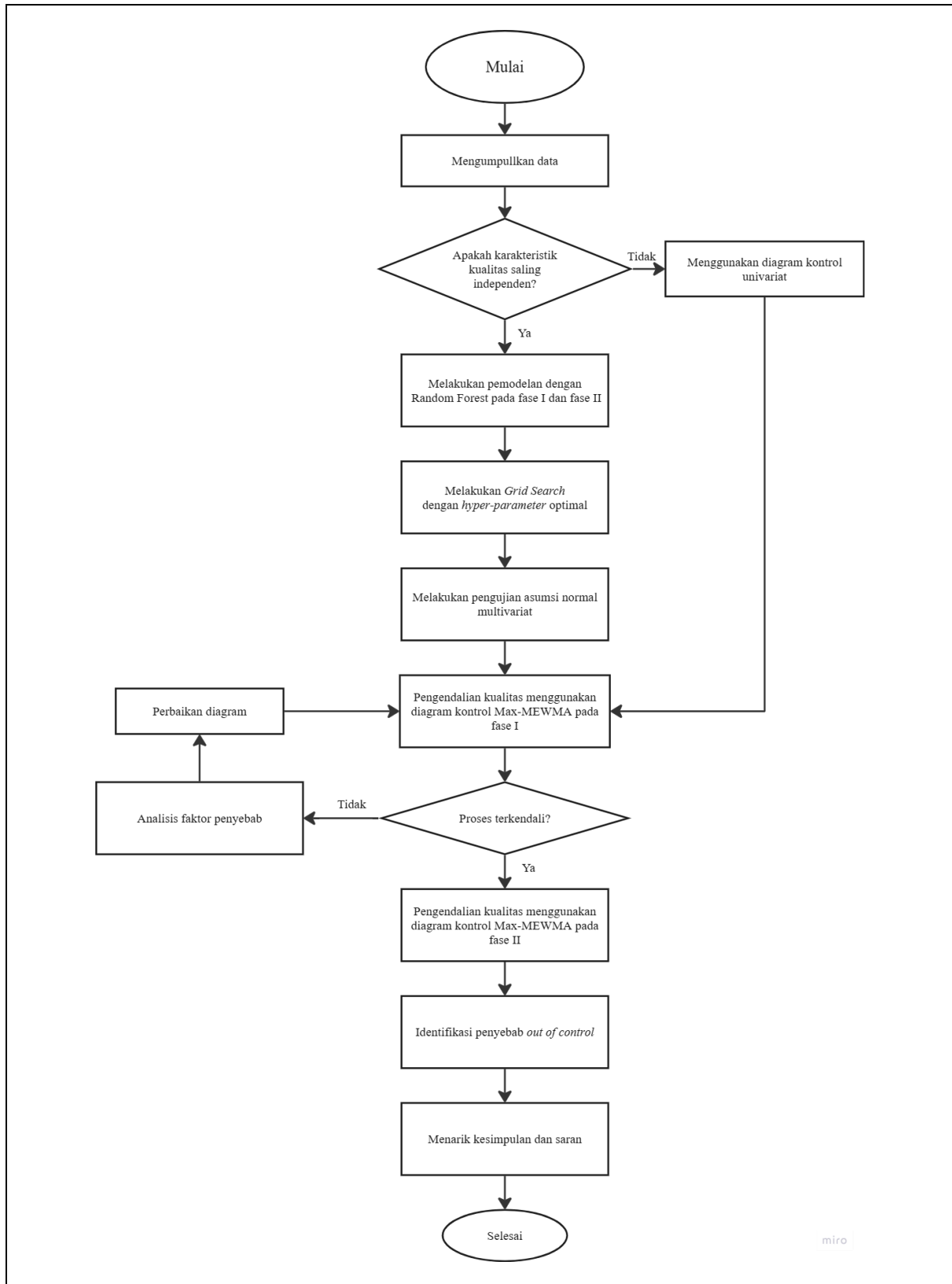
Berikut adalah langkah-langkah analisis yang dilakukan dalam penelitian ini:

1. Mengumpulkan data sekunder dari laporan kualitas.
2. Melakukan eksplorasi terhadap setiap karakteristik kualitas.
3. Memeriksa signifikansi plot *Autocorrelation Function* (ACF) pada karakteristik kualitas di PLTGU PT PLN Nusantara Power UP Gresik.
4. Memeriksa signifikansi plot *Partial Autocorrelation Function* (PACF) untuk menentukan lag yang digunakan sebagai input pada model *Random Forest Regressor*.
5. Menentukan karakteristik kualitas yang akan digunakan sebagai output model, yaitu karakteristik yang menunjukkan ACF signifikan. Setiap karakteristik kualitas dimodelkan secara simultan dalam kerangka multivariat (multioutput).
6. Membangun model *Random Forest Regressor* dengan input berupa variabel-variabel lag yang signifikan (hasil PACF) dengan output berupa vektor multivariat dari karakteristik kualitas.
 - a. Dilakukan pemilihan *split* terbaik dicari dengan meminimalkan fungsi pada persamaan (2.11).
 - b. Menghitung pembobotan pemilihan split dilakukan menggunakan *joint entropy* atau *information gain* dengan menggunakan persamaan (2.16).
7. Melakukan *tuning hyperparameter* model *Random Forest* (seperti jumlah pohon, kedalaman pohon, jumlah fitur maksimum, dan minimum jumlah sampel) melalui *grid search* dengan validasi silang *TimeSeriesSplit*. Model diukur dengan kriteria Mean Squared Error (MSE) minimum seperti pada persamaan (2.18) hingga diperoleh model terbaik, dimana parameter optimal θ didapatkan dengan persamaan (2.19) yang nantinya akan digunakan untuk memodelkan data fase II.

8. Menghitung residual dari hasil prediksi pada data fase I dengan cara mengurangkan nilai aktual dengan hasil prediksi dari model *Random Forest*.
9. Melakukan uji normalitas multivariat pada data residual fase I menggunakan uji *Shapiro Wilk* dan pengecekan autokorelasi pada data residual fase I menggunakan plot MCC dan uji *Pormanteau*.
10. Melakukan pemantauan kualitas produk dengan membuat diagram pengendali Max-MEWMA untuk mengendalikan variabilitas proses.
 - a. Melakukan pemantauan kualitas Max-MEWMA pada fase I. Menentukan parameter μ_0 (nilai dari vektor rata-rata), dan Σ_0 (nilai dari matriks kovarians). Estimasi nilai μ_0 dan Σ_0 didapatkan dari data fase I keseluruhan menggunakan peta pengendali Max-MEWMA untuk mendapatkan data yang *in control*.
 - b. Untuk masing-masing sampel, hitung Z_0 dengan $Z_0 = 0$ sebagai nilai awal dan hitung Y_i dengan $Y_0 = 0$ sebagai nilai awal.
 - c. Menghitung nilai U_i menggunakan persamaan (2.21).
 - d. Menghitung nilai V_i menggunakan persamaan (2.25).
 - e. Menghitung nilai M_i menggunakan persamaan (2.26).
 - f. Membuat diagram kendali dengan menggunakan nilai M_i yang diperoleh dari nilai $|U_i|$ atau $|V_i|$ kemudian memplotkan pada diagram dengan batas kendali atas (UCL). Diagram kendali Max-MEWMA menggunakan UCL dari persamaan (2.32)
 - g. Apabila nilai $|U_i| \geq UCL$ dan $|V_i| \geq UCL$ hal ini mengindikasikan bahwa telah terjadi pergeseran proses pada vektor rata-rata dan matriks kovarians.
8. Melakukan pengendalian kualitas fase II dengan menggunakan batas kendali atas fase I dan digunakan parameter vektor mean dan matriks varian kovarian dari data fase I yang sudah terkendali secara statistik
9. Mengidentifikasi penyebab *out of control*.
10. Mengidentifikasi faktor melalui diagram Ishikawa untuk mengetahui tidak terkendalinya sistem HRSG.
11. Analisis kapabilitas proses.
12. Membuat kesimpulan

3.5 Diagram Alir

Diagram alir berdasarkan langkah analisis yang digunakan dalam penelitian ini ditunjukkan pada Gambar 3.1 sebagai berikut.



Gambar 3. 1 Diagram Alir.

BAB IV ANALISIS PEMBAHASAN

4.1 Analisis Karakteristik Data

Kinerja dari *Heat Recovery Steam Generator* dapat dianggap baik apabila ketiga parameter kualitas berada dalam rentang spesifikasi yang telah ditetapkan. Untuk itu, dilakukan analisis statistik deskriptif sebagai langkah awal guna memahami karakteristik data hasil pengamatan. Sebelum dilakukan analisis ini, data akan dibagi ke dalam dua fase, yaitu Fase I dan Fase II, dengan tujuan agar proses yang dianalisis di Fase I mencerminkan kondisi yang sudah terkendali. Berikut ini adalah hasil statistik deskriptif untuk masing-masing fase. Tabel 4.1 menyajikan statistik deskriptif untuk data pengamatan pada Fase I.

Tabel 4. 1 Statistika Deskriptif Fase I

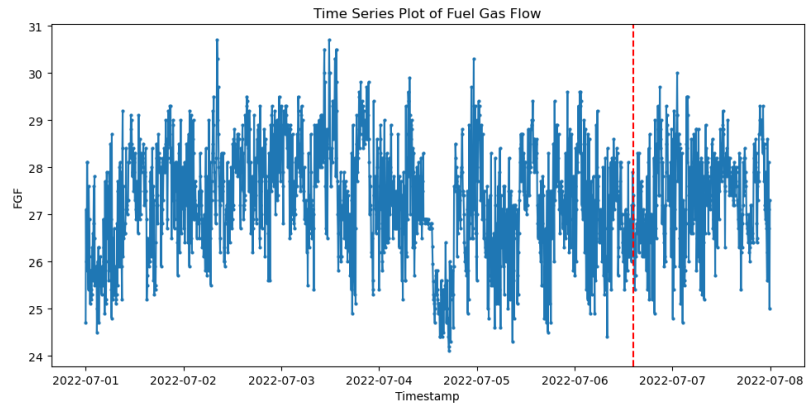
Variabel	Mean	Var	Min	Max
FGF	27,278	1,3750	24,1	30,7
CP	10,763	0,7254	9,1	12,1
GT	506,50	256,43	482	549,8

Tabel 4.1 menunjukkan bahwa nilai rata-rata dari ketiga karakteristik kualitas yaitu FGF, CP, dan GT berada dalam batas spesifikasi yang telah ditetapkan oleh PT PLN Nusantara Power. Berdasarkan nilai minimum, FGF (*Fuel Gas Flow*) memiliki nilai terendah sebesar 24,1 yang masih berada dalam batas kontrol proses. Nilai varians dari FGF dan CP masing-masing sebesar 1,3750 dan 0,7254, relatif kecil dibandingkan dengan nilai rata - ratanya, yang mengindikasikan bahwa kedua karakteristik tersebut cenderung homogen dan memiliki persebaran data yang sempit. Hal ini diperkuat dengan jarak antara nilai minimum dan maksimum yang tidak terlalu jauh, yaitu 24,1 hingga 30,7 untuk FGF dan 9,1 hingga 12,1 untuk CP. Sebaliknya, karakteristik GT (*Gas Temperature*) memiliki varians sebesar 256,43, jauh lebih tinggi dibanding dua karakteristik lainnya, yang menunjukkan bahwa data GT memiliki penyebaran yang lebih besar dan fluktuasi yang lebih tinggi sepanjang proses.

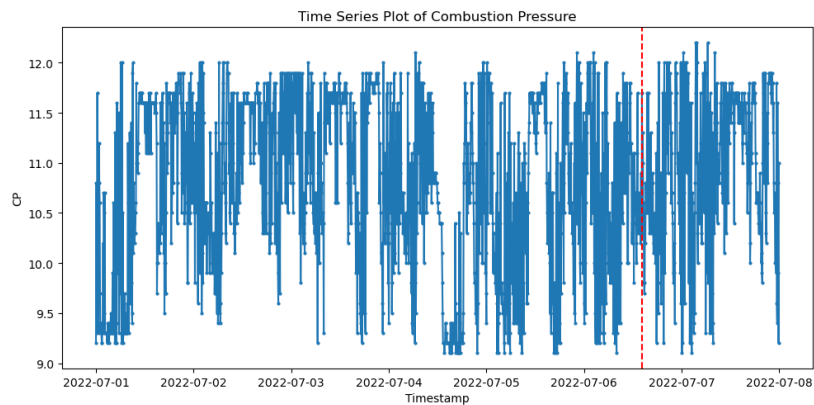
Tabel 4. 2 Statistika Deskriptif Fase II

Variabel	Mean	Var	Min	Max
FGF	27,394	1	24,7	30
CP	10,958	0,565	9,1	12,2
GT	501,093	137,175	483,4	545,2

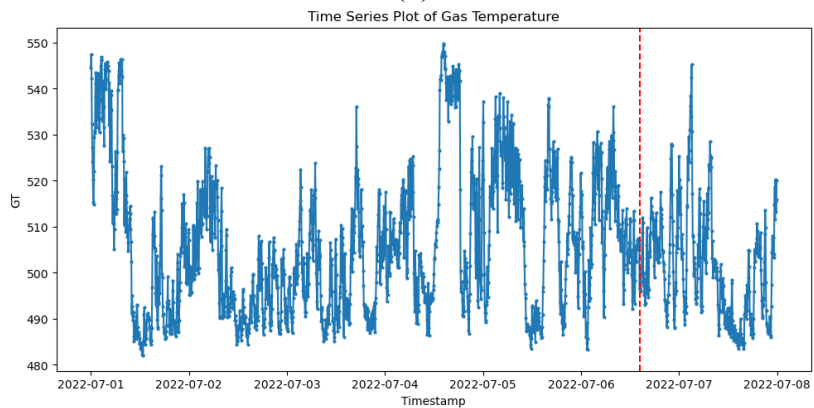
Tabel 4.2 menunjukkan bahwa nilai rata-rata dari ketiga karakteristik kualitas, yaitu FGF, CP, dan GT, berada dalam rentang batas spesifikasi yang telah ditetapkan oleh PT PLN Nusantara Power. Nilai varians FGF sebesar 1 dan CP sebesar 0,565 mengindikasikan persebaran data yang relatif kecil terhadap nilai rata-ratanya, sehingga menunjukkan tingkat homogenitas yang baik. Sementara itu, varians GT yang lebih besar yaitu 137,175 adalah hal yang wajar mengingat skala pengukuran suhu yang jauh lebih tinggi dibandingkan dua variabel lainnya.



(a)



(b)



(c)

Gambar 4.1 Time Series Plot (a) FGF, (b) CP, (c) GT

Pada Gambar 4.1 dapat kita lihat time series plot dari ketiga karakteristik kualitas tiap selang 5 menit pada tanggal 1 Juli 2022 – 7 Juli 2022. Karakteristik kualitas laju aliran bahan bakar (Gambar 4.1(a)) cenderung lebih stabil akan tetapi pada hari terakhir dia mengalami penurunan dan kenaikan yang cukup signifikan, lalu untuk karakteristik kedua yaitu pada tekanan pembakaran (Gambar 4.1(b)) juga memiliki pola yang sama, pada hari pertama pola cenderung stabil akan tetapi saat hari terakhir terdapat beberapa plot yang naik dan turun secara signifikan. Pada karakteristik terakhir temperature gas buang (Gambar 4.1(c)) memiliki pola yang tidak stabil dari hari pertama hingga hari terakhir tetapi pada hari terakhir terdapat beberapa pola yang naik dan turun secara signifikan. Ketiga karakteristik kualitas ini sangat

berpengaruh terhadap efisiensi sistem *Heat Recovery Steam Generator* (HRSG). Dimana persamaan (2.42) menjelaskan bahwa temperatur gas buang akan berdampak terhadap HRSG. HRSG memiliki beberapa parameter operasional yang kritis dan perlu diperhatikan agar efisiensinya tetap terjaga selama beroperasi. Alasan ketiga parameter ini kritis adalah karena ketiga hal tersebut mempengaruhi terhadap besar kecil nya energi panas yang di lepas ke sistem HRSG untuk memanaskan air menjadi superheated steam. Jika energi panas yang terlepas semakin besar akan berpengaruh terhadap kekuatan struktur material HRSG, yang mana hal ini bisa menyebabkan kegagalan material akibat suhu yang diterima terlalu tinggi. Jika energi panas yang terlepas terlalu kecil maka akan berpengaruh terhadap waktu pembentukan air menjadi superheated steam karena suhu rendah maka waktu yang diperlukan semakin lama dan berakibat pada pemborosan bahan bakar.

4.2 Pemeriksaan Dependensi Antar Variabel

Sebelum dilakukan pengendalian kualitas terhadap sistem pembangkit pada PLTGU PT. PLN Nusantara Power UP Gresik menggunakan diagram kontrol Max-MEWMA, perlu dilakukan pemeriksaan terhadap asumsi dependensi antar variabel. Hal ini penting karena diagram kontrol multivariat mensyaratkan adanya keterkaitan antar karakteristik kualitas yang dimodelkan. Dalam penelitian ini, karakteristik kualitas proses yang dianalisis meliputi Fuel Gas Flow (FGF), Combustion Pressure (CP), dan Gas Temperature (GT). Untuk memastikan adanya hubungan antar ketiga variabel tersebut, dilakukan uji Bartlett. Hasil pengujian ditunjukkan pada Tabel 4.3 berikut.

Tabel 4. 3 Hasil Uji Barlett

Data	Chi-square	df	p-value
Fase I	4289,82	3	0
Fase II	882,6329	3	5,176288e-191

Berdasarkan Tabel 4.3, diperoleh nilai statistik uji Bartlett untuk data fase I sebesar 4289,82 dan untuk data fase II sebesar 882,63 dengan derajat bebas (df) sebesar 3. Nilai ini jauh lebih besar dari nilai kritis *Chi-square* pada taraf signifikansi 5%, sehingga hipotesis nol yang menyatakan tidak adanya hubungan antar variabel dapat ditolak. Selain itu, jika ditinjau dari nilai *p-value*, diperoleh hasil sebesar 0 untuk fase I dan $5,18 \times 10^{-191}$ untuk fase II, yang keduanya jauh lebih kecil dari batas signifikansi 0,05.

Dengan demikian, dapat disimpulkan bahwa ketiga variabel proses, yaitu *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Gas Temperature* (GT), memiliki hubungan yang signifikan dan bersifat dependen secara multivariat, baik pada fase I maupun fase II. Temuan ini memperkuat dasar bahwa pendekatan pengendalian berbasis multivariat, seperti diagram kontrol Max-MEWMA, sangat relevan untuk digunakan dalam menganalisis dan memonitor proses pembangkitan listrik di PLTGU PT. PLN Nusantara Power UP Gresik.

4.3 Uji Nonlinieritas

Setelah dilakukan pemeriksaan terhadap asumsi dependensi antar variabel karakteristik kualitas proses, tahap selanjutnya yang perlu dilakukan adalah menguji apakah hubungan antara variabel prediktor dan variabel respon bersifat linier atau terdapat indikasi non-linier. Pengujian ini penting dilakukan untuk menentukan model regresi yang sesuai sebelum digunakan dalam pemodelan residual sebagai dasar penerapan diagram kontrol Max-MEWMA. Dalam penelitian ini, pengujian nonlinieritas dilakukan terhadap masing-masing variabel respon, yaitu *Fuel Gas*

Flow (FGF), *Combustion Pressure* (CP), dan *Gas Temperature* (GT), dengan Uji nonlinieritas yang digunakan yaitu uji terasvirta dan didapatkan hasil sebagai berikut.

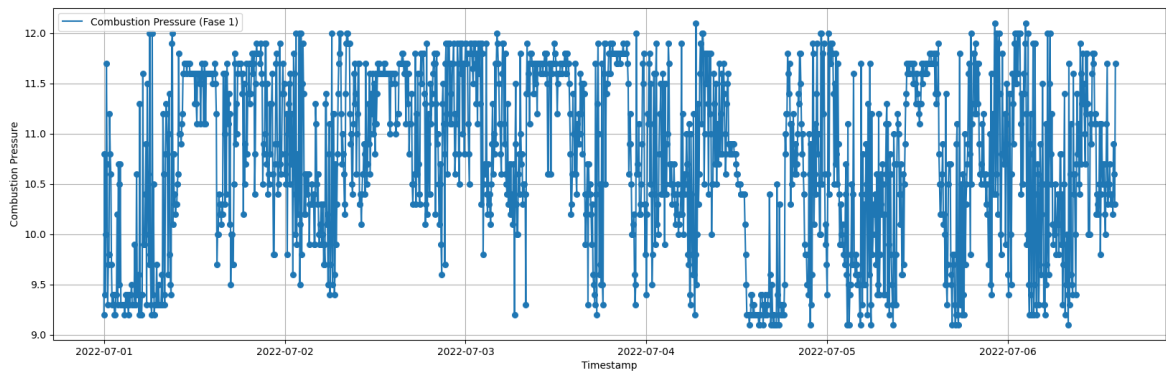
Tabel 4. 4 Hasil Uji Terasvirta

Variabel Respon	Chi-square	df	P-value
Fuel Gas Flow	42,813	2	6,185808e-19
Combustion Pressure	37,076	2	1,54082e-16
Gas Temperature	54,469	2	2,372794e-23

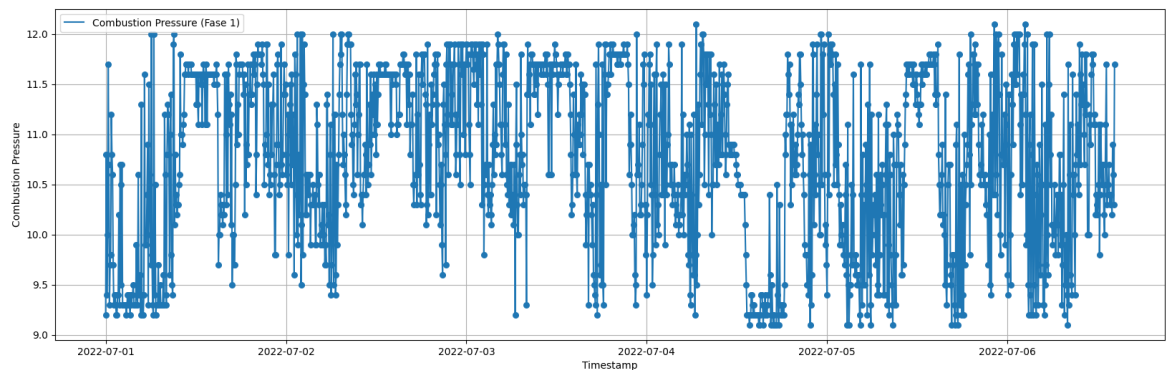
Dengan menggunakan nilai $\alpha = 5\%$, dapat diketahui bahwa nilai $\chi^2_{\text{tabel}} = 5,991$ untuk derajat bebas (df) = 2. Berdasarkan hasil uji Terasvirta pada Tabel 4.4, diperoleh bahwa seluruh nilai χ^2 hitung untuk ketiga variabel (*Fuel Gas Flow*, *Combustion Pressure*, dan *Gas Temperature*) lebih besar dari χ^2 tabel, yaitu masing-masing sebesar 42,813; 37,076; dan 53,469. Selain itu, seluruh *p-value* yang diperoleh juga lebih kecil dari $\alpha = 5\%$, yaitu $< 2.2 \times 10^{-16}$. Oleh karena itu, keputusan yang diambil adalah menolak H_0 . Hal ini memperkuat kesimpulan bahwa hubungan antar variabel prediktor dan respon bersifat nonlinier, sehingga pemodelan menggunakan algoritma regresi nonlinier seperti *Random Forest Regressor* menjadi lebih relevan.

4.4 Pengecekan Autokorelasi

Sebelum dilakukan uji autokorelasi, data terlebih dahulu dibagi menjadi bagian *training* dan *testing*. Pembagian ini didasarkan pada hasil pengamatan terhadap *time series* plot masing-masing variabel, di mana grafik menunjukkan adanya pola kenaikan atau penurunan yang signifikan. Berdasarkan Gambar 4.1, ditampilkan visualisasi deret waktu dari setiap karakteristik kualitas untuk mendukung proses peninjauan tersebut.



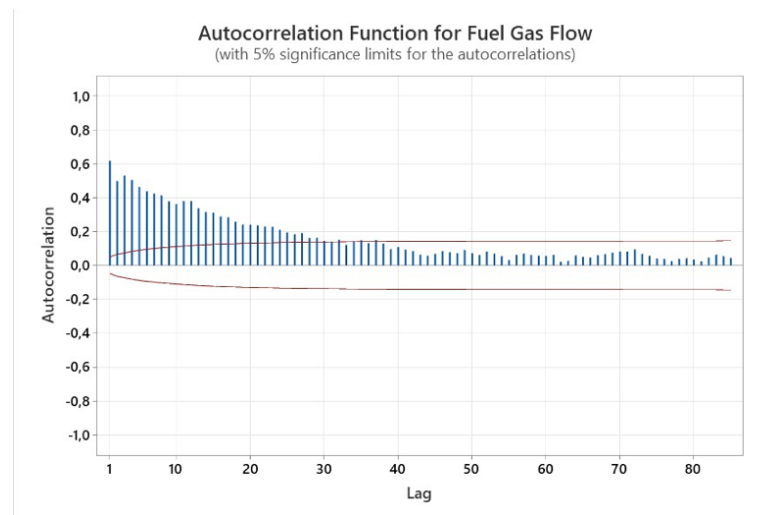
(b)



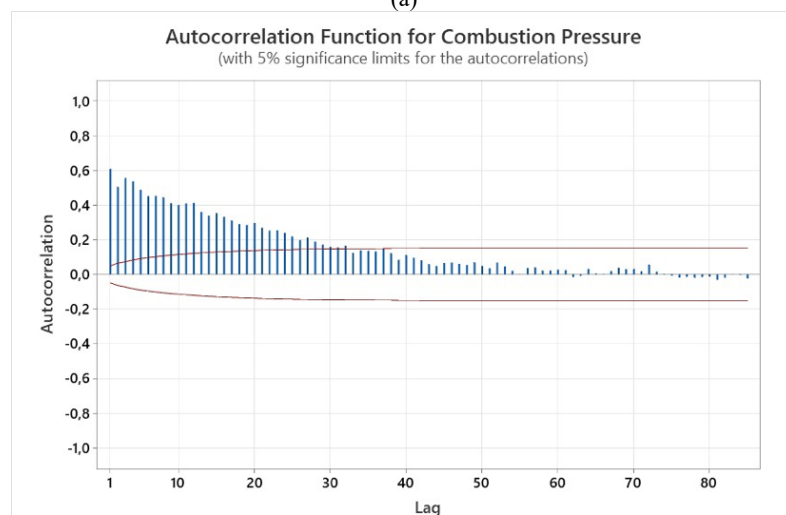
(c)

Gambar 4. 2 Time Series Plot Fase I (a) FGF, (b) CP, (c) GT

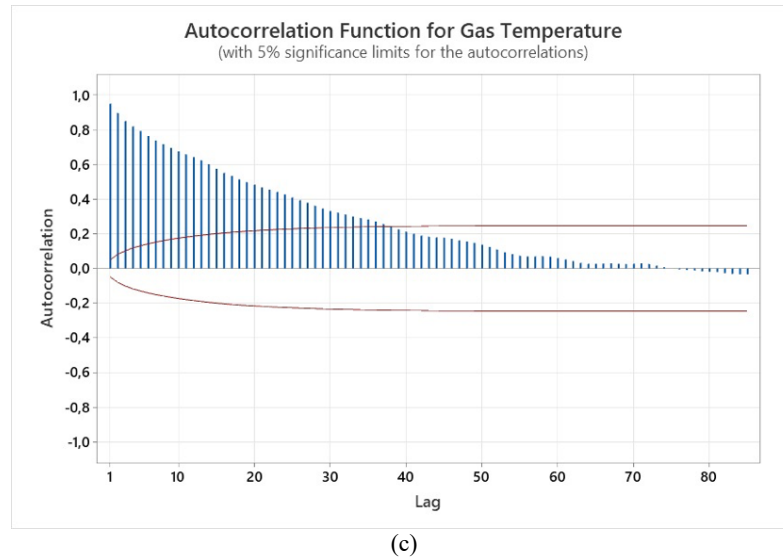
Gambar 4.2 menunjukkan *time series* plot dari ketiga karakteristik kualitas, yaitu laju aliran bahan bakar (FGF), tekanan pembakaran (CP), dan temperatur gas buang (GT), selama periode 1 Juli hingga 7 Juli 2022. Pada karakteristik FGF, terlihat fluktuasi nilai yang cukup tajam dalam periode waktu yang singkat, menandakan adanya ketidakstabilan di akhir minggu. Pola serupa juga tampak pada karakteristik CP, yang menunjukkan kestabilan relatif di awal minggu namun mengalami kenaikan dan penurunan signifikan menjelang akhir minggu. Sementara itu, GT tampak menunjukkan fluktuasi yang tidak teratur sejak awal, dengan pola naik-turun yang lebih ekstrem dibandingkan dua variabel lainnya. Berdasarkan pola deret waktu tersebut, data Fase I digunakan sebagai data *training* sebanyak 1600 observasi, sementara sisanya digunakan sebagai data Fase II untuk *testing*. Selanjutnya, data ini akan diperiksa apakah mengandung autokorelasi, dengan menggunakan plot fungsi autokorelasi (ACF) terhadap masing-masing karakteristik kualitas.



(a)



(b)



Gambar 4. 3 Plot ACF (a) FGF, (b) CP, (c) GT

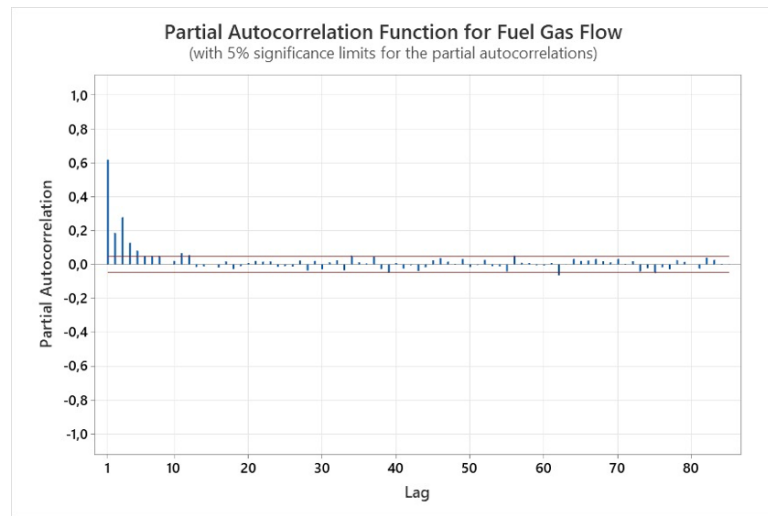
Gambar 4.4 menunjukkan hasil uji autokorelasi untuk ketiga variabel karakteristik kualitas, yaitu *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Gas Temperature* (GT). Berdasarkan grafik ACF, seluruh variabel menunjukkan adanya autokorelasi yang signifikan, terutama pada lag awal. Variabel FGF memperlihatkan autokorelasi positif yang konsisten hingga sekitar lag ke-15, sementara CP juga menunjukkan pola serupa meskipun dengan intensitas yang sedikit lebih rendah. GT merupakan variabel dengan autokorelasi paling kuat dan bertahan hingga lag yang lebih panjang, mencerminkan ketergantungan waktu yang tinggi. Hal ini dapat terjadi karena sistem *Heat Recovery Steam Generator* (HRSG) beroperasi secara kontinu, sehingga setiap variabel proses cenderung memiliki ketergantungan terhadap waktu sebelumnya. Oleh karena itu, penggunaan model *Random Forest Regressor* digunakan untuk memodelkan ketergantungan waktu tersebut dan menghasilkan residual yang lebih independen. Selanjutnya, residual dari model digunakan dalam diagram kontrol Max-MEWMA untuk mendeteksi perubahan proses secara statistik dengan lebih akurat.

4.5 Pemodelan Menggunakan *Random Forest Regressor*

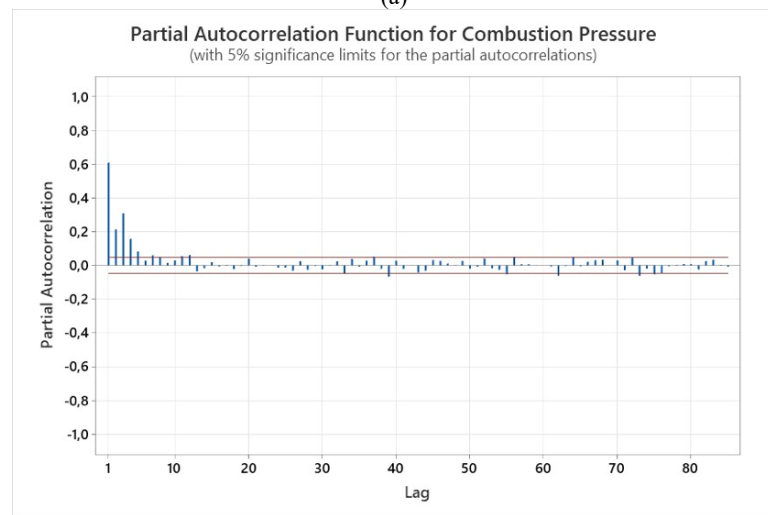
Pemodelan menggunakan algoritma *Random Forest Regressor* diterapkan pada data proses produksi listrik di PT PLN Nusantara Power UP Gresik untuk mengatasi permasalahan autokorelasi yang terdeteksi pada variabel karakteristik kualitas. Data input pada model dibentuk dengan mempertimbangkan nilai lag dari masing-masing variabel, sesuai dengan pola autokorelasi yang ditunjukkan pada analisis sebelumnya. Pemilihan *hyperparameter* model dilakukan melalui metode *Grid Search* untuk memperoleh kombinasi parameter yang optimal. Model yang telah dibangun dan dioptimasi pada Fase I kemudian digunakan untuk melakukan prediksi dan monitoring pada Fase II, dengan residual hasil prediksi digunakan dalam diagram kontrol berbasis residual seperti Max-MEWMA.

4.5.1 Pemodelan Menggunakan *Random Forest Regressor* Fase I

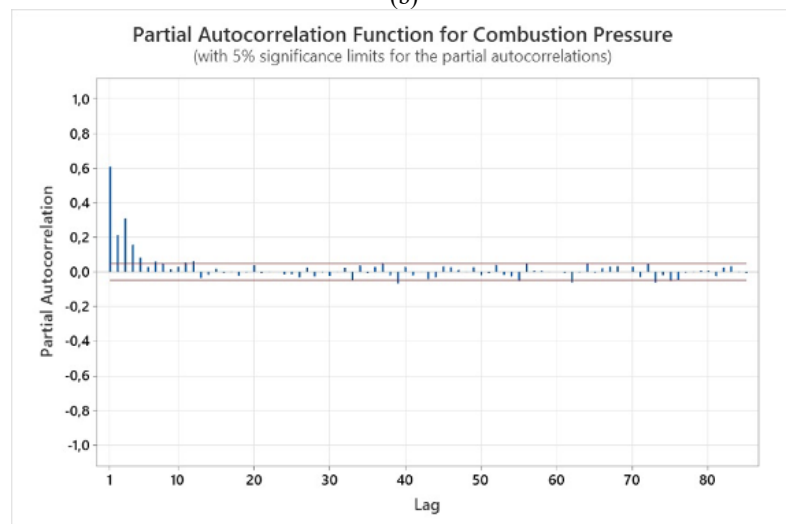
Langkah pertama pada pemodelan *Random Forest Regressor* adalah mencari lag yang signifikan pada plot PACF setiap variabel karakteristik kualitas dengan menggunakan data GT 23 Operation 202207010005 atau Blok II pada Pola operasi (1,1,1) 29 dimana 1 Turbin Gas, 1 HRSG, 1 Turbin Uap di PLTGU PT. PLN Nusantara Power UP Gresik pada tanggal 1 Juli – 6 Juli 2022.



(a)



(b)

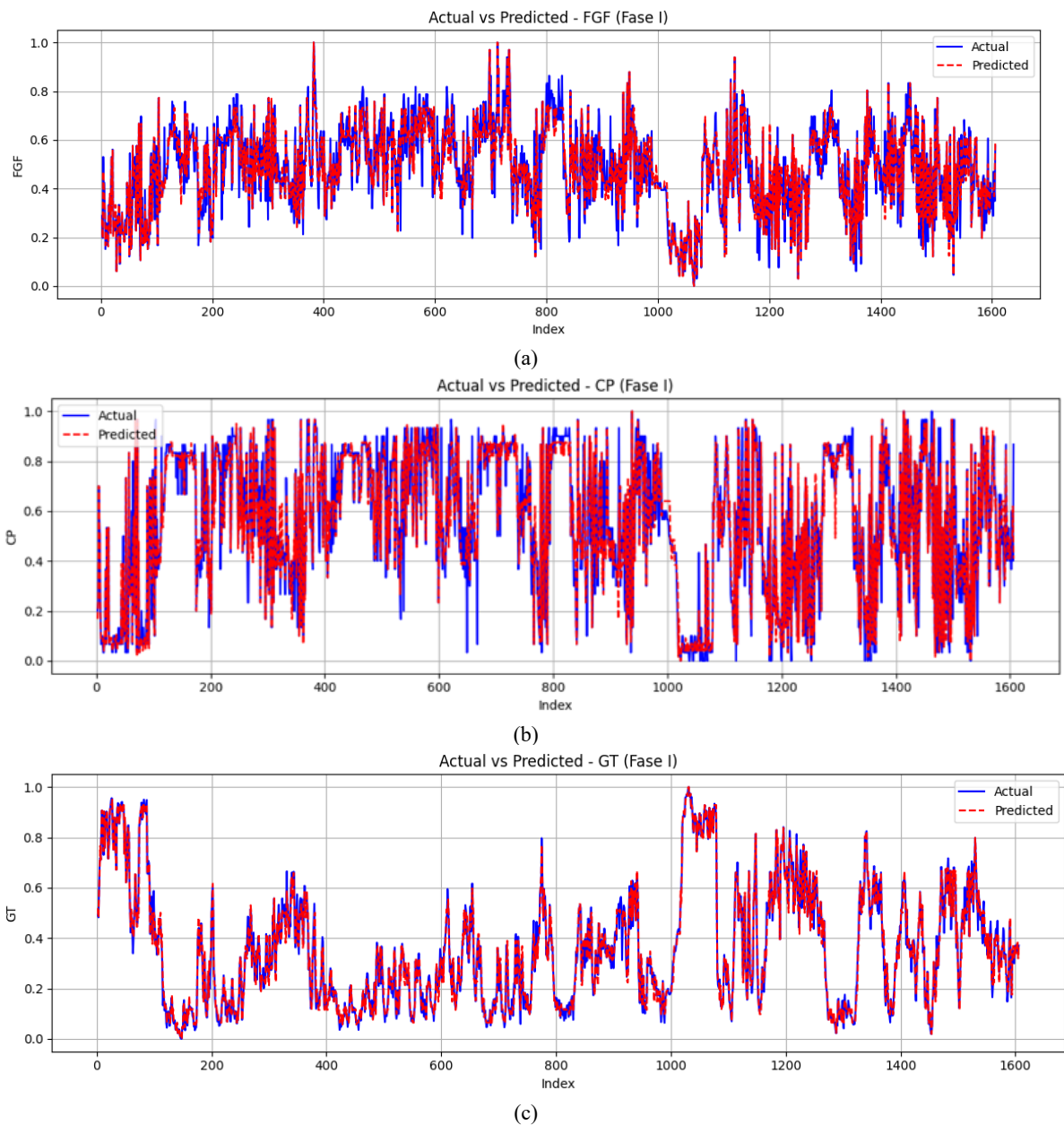


(c)

Gambar 4. 4 Plot PACF (a) Fuel Gas Flow, (b) Combustion Pressure, dan (c) Gas Temperature

Dapat dilihat pada Gambar 4.4, plot PACF fase I dengan taraf signifikansi 0,05 menunjukkan bahwa setiap variabel karakteristik kualitas memiliki lag-lag signifikan pada

beberapa titik. Karakteristik laju aliran bahan bakar (FGF) signifikan pada lag ke-1, 2, 3, 4, dan 5. Karakteristik tekanan pembakaran (CP) menunjukkan lag signifikan pada lag ke-1, 2, 3, 4, dan 5. Sementara itu, karakteristik temperatur gas buang (GT) memiliki lag signifikan pada lag ke-1, 3, dan 4. Berdasarkan hasil PACF ini, *input* model *Random Forest Regressor* pada $y_1(i-1), y_1(i-2), y_1(i-3), y_1(i-4), y_1(i-5), y_2(i-2), y_2(i-3), y_2(i-4), y_2(i-5), y_3(i-1), y_3(i-2), y_3(i-3),$ dan $y_3(i-4)$. Pemilihan kombinasi lag ini digunakan sebagai input utama pada *model Random Forest*, dengan *tuning hyperparameter* melalui metode *Grid Search*. Kombinasi *hyperparameter* terbaik digunakan dalam pemodelan fase I untuk mendapatkan hasil prediksi optimal terhadap karakteristik kualitas sistem *Heat Recovery Steam Generator* (HRSG). Setelah model dibentuk, dilakukan perbandingan antara data aktual dengan nilai prediksi *Random Forest Regressor* pada fase I. Berikut merupakan hasil plot deret waktu antara data aktual dan prediksi model untuk ketiga variabel karakteristik kualitas HRSG.

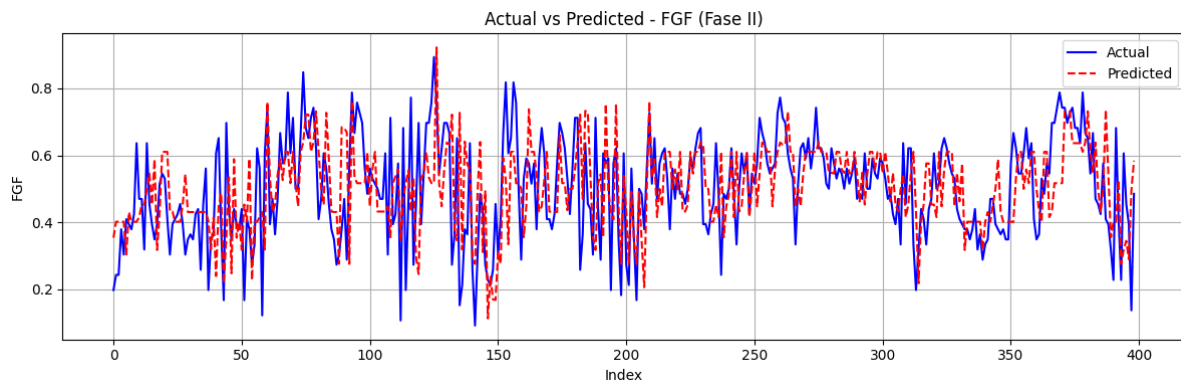


Gambar 4. 5 Aktual vs Prediksi Fase I (a) Fuel Gas Flow, (b) Combustion Pressure, (c) Gas Temperature.

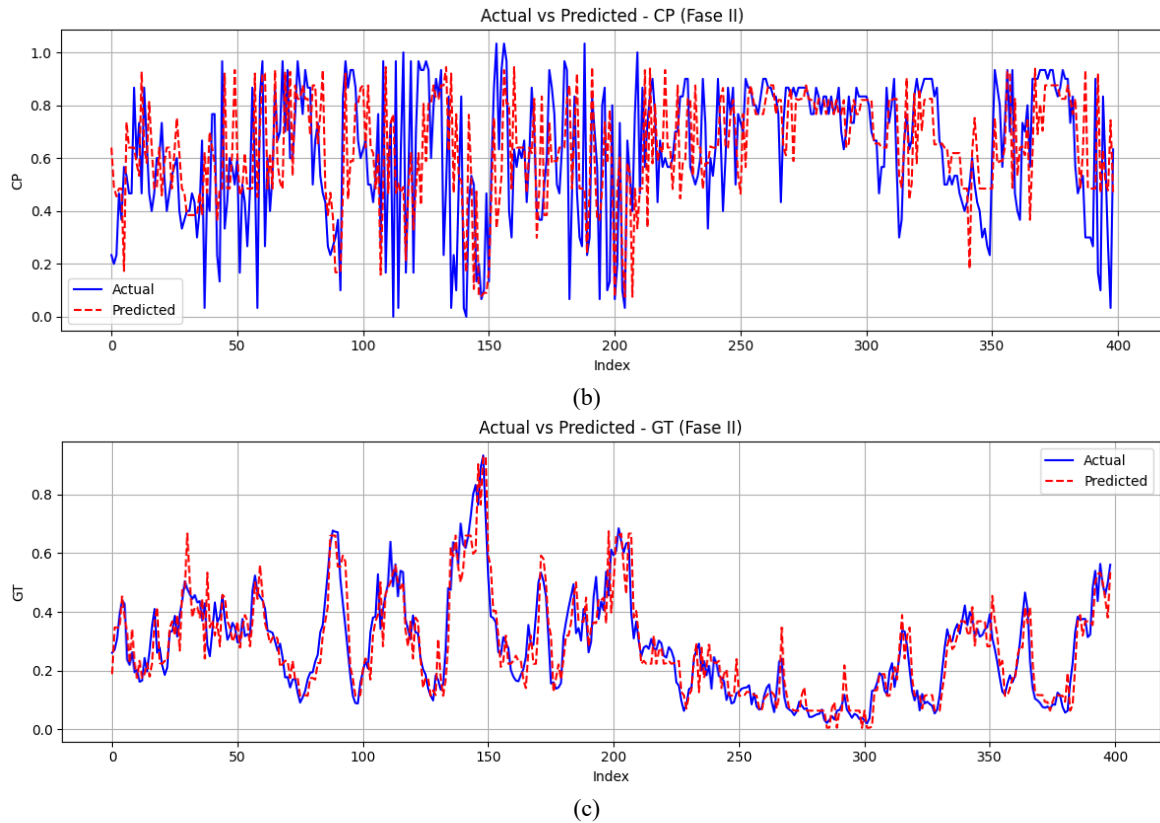
Berdasarkan Gambar 4.6, dapat dilihat bahwa hasil prediksi variabel *Fuel Gas Flow* pada fase I menggunakan model *Random Forest Regressor* belum sepenuhnya menyerupai pola dari data aktual. Meski demikian, dibandingkan dengan percobaan kombinasi *hyperparameter* lainnya, hasil ini merupakan yang paling mendekati data aktual. Untuk variabel tekanan pembakaran *Combustion Pressure*, prediksi model menunjukkan pola yang lebih menyerupai data aktual dibandingkan dengan variabel sebelumnya. Adapun prediksi untuk variabel temperatur gas buang *Gas Temperature* terlihat paling mendekati pola aktual dan memiliki kesesuaian yang lebih tinggi dibandingkan dengan dua variabel lainnya. Oleh karena itu, model *Random Forest* dengan kombinasi *hyperparameter* optimal yang diperoleh melalui *Grid Search* dipilih sebagai model terbaik dengan hasil prediksi pada fase I yang paling representatif terhadap data aktual sistem HRSG. Dengan kombinasi *hyper-parameter* optimal yang digunakan yaitu $n_estimators = 100$, $max_depth = 10$, $min_samples_split = 2$, dan $min_samples_leaf = 1$ yang menghasilkan nilai MSE minimum sebesar $MSE_{FGF} = 0,004595$, $MSE_{CP} = 0,008962$, dan $MSE_{GT} = 0,001178$ yang menunjukkan bahwa prediksi model terhadap variabel *Gas Temperature* paling mendekati nilai aktual, disusul oleh *Combustion Pressure* dan *Fuel Gas Flow*.

4.5.2 Pemodelan Menggunakan Random Forest Regressor Fase II

Setelah dilakukan pemodelan menggunakan *Random Forest Regressor* pada data GT 23 Operation 202207010005 atau Blok II dengan konfigurasi pola operasi (1,1,1) yang terdiri dari 1 Turbin Gas, 1 HRSG, dan 1 Turbin Uap di PLTGU PT PLN Nusantara Power UP Gresik pada periode 1 Juli hingga 8 Juli 2022, model yang telah dilatih menggunakan kombinasi *hyperparameter* optimal kemudian diterapkan untuk melakukan prediksi pada data Fase II. Hasil prediksi tersebut dibandingkan dengan data aktual untuk mengevaluasi performa model dalam memproyeksikan kondisi sistem. Berikut ditampilkan *time series plot* antara nilai aktual dan nilai prediksi dari model *Random Forest* untuk ketiga variabel karakteristik kualitas sistem *Heat Recovery Steam Generator* (HRSG) pada fase II, guna menilai sejauh mana model mampu mengikuti pola proses aktual pada data pengujian.



(a)



Gambar 4. 6 Aktual vs Prediksi Fase II (a) Fuel Gas Flow, (b) Combustion Pressure, (c) Gas Temperature.

Berdasarkan Gambar 4.6, dapat dilihat bahwa pola data aktual pada variabel *Fuel Gas Flow* (FGF) pada fase II cenderung berbeda dengan pola hasil prediksi dari model Random Forest. Hal serupa juga terlihat pada variabel *Combustion Pressure* (CP), di mana meskipun pola prediksi relatif mendekati, tetap terdapat deviasi dari data aktual. Adapun pada variabel *Gas Temperature* (GT), hasil prediksi tampak sedikit lebih konsisten namun tetap menunjukkan perbedaan dengan pola aktual. Secara keseluruhan, hasil *time series plot* antara nilai prediksi dan nilai aktual dari ketiga karakteristik kualitas pada fase II menunjukkan adanya ketidaksesuaian pola yang dapat mengindikasikan pergeseran proses, perbedaan pola ini selanjutnya akan dianalisis lebih lanjut menggunakan diagram kontrol.

4.4 Pengujian Asumsi Diagram Kontrol

Karakteristik kualitas sistem *Heat Recovery Steam Generator* (HRSG) di PLTGU PT PLN Nusantara Power UP Gresik pada fase I telah dimodelkan menggunakan algoritma *Random Forest Regressor*. Sebelum dilakukan pengujian terhadap asumsi distribusi normal multivariat yang diperlukan dalam analisis diagram kontrol, terlebih dahulu perlu dipastikan bahwa model telah berhasil mengurangi autokorelasi pada data. Oleh karena itu, dilakukan pemeriksaan terhadap autokorelasi residual model melalui analisis *Matriks Cross-Correlation* (MCC). Berikut disajikan hasil plot MCC untuk residual pada fase I guna mengevaluasi apakah residual telah bersifat acak dan memenuhi asumsi independensi.

Tabel 4. 5 MCC Residual Fase I

Variabel /Lag	1	2	3	4	5	6	7	8	9	10	11	12	13
FGF	...	...	...	..+	...	...	...	...	...	...	...	...	+..
CP	...	...	...	...	...	...	.+.	...	...	..+	...	...	...
GT	...	...	...	...	..+	...	...	...	...	...	...	-..	...

Variabel /Lag	14	15	16	17	18	19	20
FGF	.-.	...	...	...	...	...	...
CP	...	...	...	...	...	...	...
GT	.-.	...	...	...	...	...	+..

Hasil plot MCC dari ketiga karakteristik kualitas pada Fase I menunjukkan bahwa pemodelan menggunakan *Random Forest Regressor* telah cukup berhasil dalam mengurangi autokorelasi. Hal ini ditunjukkan oleh dominasi simbol titik (“...”) pada sebagian besar lag, yang mengindikasikan tidak adanya korelasi signifikan antar waktu. Dengan demikian, dapat disimpulkan bahwa asumsi independensi telah relatif terpenuhi untuk kelanjutan analisis pengendalian proses menggunakan diagram kontrol Max-MEWMA.

Selain itu, pengujian asumsi residual independen juga bisa dilakukan dengan pengujian *portmanteau* untuk menguji asumsi *white noise*. Berikut ini akan disajikan hasil dari uji *portmanteau* pada Tabel 4.6.

Tabel 4. 6 Hasil Uji *Portmanteau* Residual Fase I

Lag	Statistik Uji	df	<i>P-value</i>
1	10,9	9	0,28
12	107,9	108	0,49
24	220,4	216	0,40
36	326,5	324	0,45

Berdasarkan hasil uji *portmanteau* yang disajikan pada Tabel 4.6, diketahui bahwa nilai *p-value* untuk berbagai lag (misalnya lag ke-1, 12, 24, dan 36) berada di atas tingkat signifikansi 5%. Hal ini menunjukkan bahwa tidak terdapat autokorelasi signifikan pada residual multivariat hingga lag ke-36. Dengan demikian, dapat disimpulkan bahwa residual telah bersifat *white noise* secara multivariat, sehingga asumsi independensi terpenuhi dan model layak untuk digunakan dalam pemantauan proses multivariat seperti Max-MEWMA.

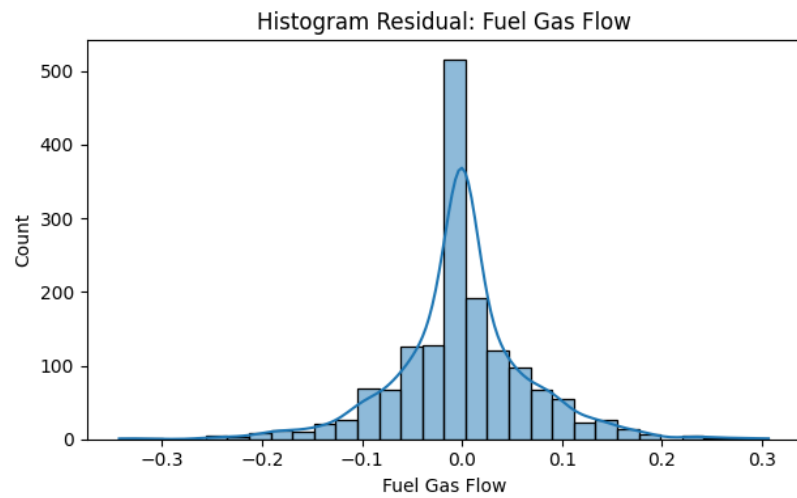
Setelah dilakukan pengecekan autokorelasi melalui MCC dan uji *portmanteau*, langkah selanjutnya adalah melakukan pengujian terhadap distribusi normal multivariat untuk ketiga karakteristik kualitas pada residual fase I. Uji normalitas dilakukan menggunakan metode *Shapiro-Wilk*.

Tabel 4. 7 Uji Normal Multivariat Residual Fase I

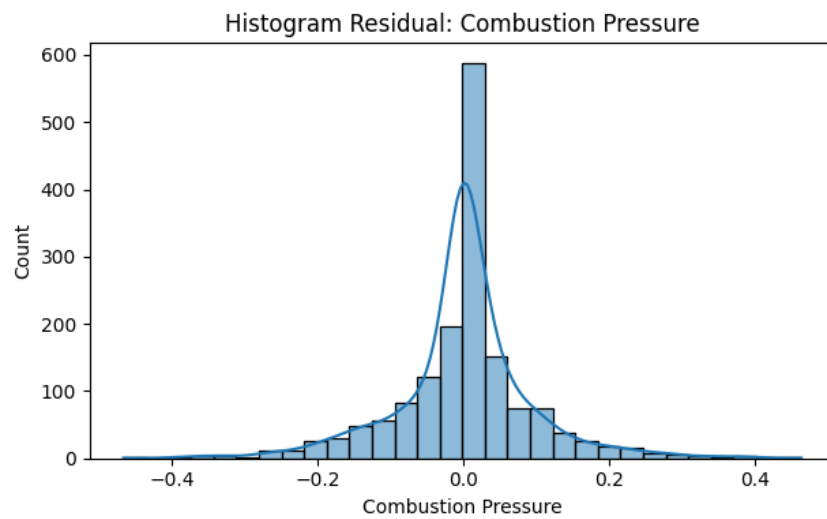
Statistik W	<i>p-value</i>
0,96572	4,67818e-19

Berdasarkan hasil uji normalitas multivariat terhadap data residual fase I menggunakan metode *Shapiro-Wilk*, diperoleh nilai statistik W sebesar 0,7155 dengan *p-value* sebesar 4,67818e-19. Karena nilai *p-value* tersebut jauh lebih kecil dari tingkat signifikansi α sebesar

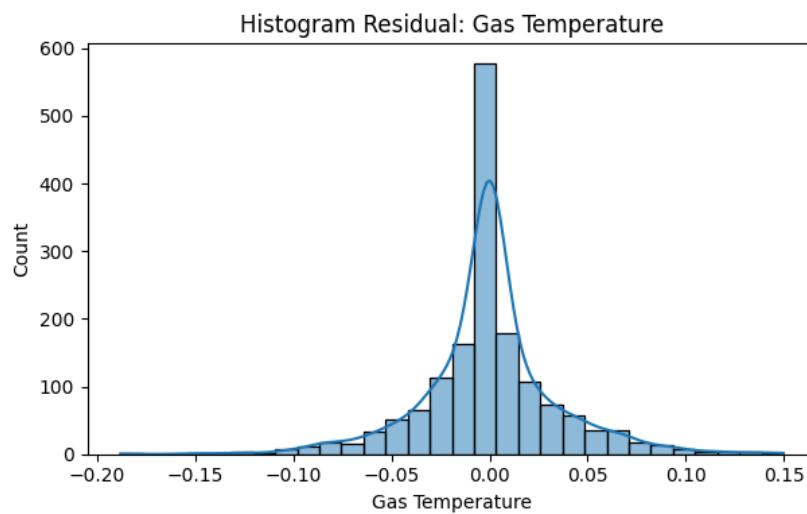
5%, maka dapat disimpulkan bahwa residual dari model pada fase I tidak memenuhi asumsi distribusi normal multivariat.



(a)



(b)



(c)

Gambar 4. 7 Histogram (a) Fuel Gas Flow, (b) Combustion Pressure, dan (c) Gas Temperature.

Hasil ini juga didukung oleh visualisasi pada Gambar 4.8, yang menunjukkan bentuk distribusi residual untuk ketiga variabel proses (Fuel Gas Flow, Combustion Pressure, dan Gas Temperature). Pada histogram residual, terlihat bahwa mayoritas nilai residual terkonsentrasi di sekitar nol, namun distribusinya tidak membentuk pola *bell curve* yang simetris seperti distribusi normal.

Kondisi residual yang tidak normal ini kemungkinan besar disebabkan oleh karakteristik data proses yang memang tidak bersifat linier dan cenderung kompleks secara hubungan antar variabel. Model *Random Forest* yang digunakan memang mampu menangkap pola non-linier dengan baik, namun residual yang dihasilkan dari model non-linier seperti ini umumnya tidak mengikuti distribusi normal secara sempurna. Ketidaklinieran data juga menyebabkan adanya variasi *error* (residual) yang lebih besar pada kondisi-kondisi tertentu, sehingga distribusi residual menjadi lebih berat di ekor (*heavy-tailed*) dan tidak simetris.

Pola-pola ini semakin menegaskan bahwa residual yang dihasilkan oleh model tidak memenuhi asumsi kenormalan multivariat, sebagaimana yang juga ditunjukkan oleh hasil uji *Shapiro-Wilk* secara statistik. Dengan demikian, baik secara visual maupun statistik, residual model pada fase I belum dapat dikatakan berdistribusi normal, sehingga penggunaan metode pengendalian berbasis residual seperti Max-MEWMA yang tidak mengharuskan asumsi kenormalan menjadi lebih sesuai untuk data ini.

4.5 Diagram Kontrol Max-MEWMA Berbasis Residual *Random Forest Regressor*

Setelah dilakukan serangkaian uji asumsi terhadap residual, seperti uji autokorelasi dan uji normalitas multivariat, diperoleh bahwa residual hasil pemodelan *Random Forest Regressor* tidak memenuhi asumsi kenormalan. Namun demikian, residual tersebut telah memenuhi syarat sebagai *white noise* karena tidak terdapat autokorelasi yang signifikan. Oleh karena itu, penerapan diagram kontrol Max-MEWMA berbasis residual tetap dapat dilanjutkan, mengingat metode ini tidak mengharuskan adanya distribusi normal pada residual. Pada bagian ini, akan dibahas penerapan diagram kontrol Max-MEWMA pada residual hasil pemodelan *Random Forest Regressor*.

4.5.1 Diagram Kontrol Max-MEWMA Pada Data Fase I

Pada tahap ini, setelah dilakukan validasi terhadap asumsi dasar diagram kontrol, yaitu autokorelasi dan distribusi normal multivariat pada residual fase I, langkah selanjutnya adalah penerapan diagram kontrol Max-MEWMA berbasis residual dari model *Random Forest Regressor*. Diagram kontrol ini digunakan untuk memantau proses secara simultan berdasarkan residual yang telah terbukti tidak mengikuti distribusi normal multivariat berdasarkan hasil pengujian statistik. Oleh karena itu, batas kendali atas (UCL) dalam penelitian ini ditentukan menggunakan metode *bootstrap* tanpa bergantung pada asumsi distribusi tertentu.

Penentuan UCL dilakukan dengan metode *bootstrap* sebanyak 1000 iterasi dengan nilai alpha (α) sebesar 0,0027 yang sesuai dengan target ARL sebesar 370. Simulasi *bootstrap* ini dilakukan untuk beberapa nilai parameter lambda, yaitu 0,05; 0,1; 0,2; 0,3; 0,4; 0,5; 0,6; 0,7; 0,8; dan 0,9, untuk membandingkan sensitivitas grafik kontrol terhadap jumlah titik *out-of-control* (OOC) serta nilai absolut dari selisih antara UCL dengan nilai maksimum statistik M_i ($|UCL - \text{Max}M_i|$). Hasil lengkap dari analisis ini disajikan dalam Tabel 4.8.

Tabel 4. 8 Hasil Kombinasi λ UCL Simulasi Bootstrap

λ	Jumlah <i>Out of Control</i>	Max M_i	UCL	$ UCL - \text{Max } M_i $
0,05	53	3,843548	3,518849	0,324699
0,1	0	3,161539	3,080642	0,080897
0,2	5	3,976342	3,354574	0,621767
0,3	5	4,320383	3,540544	0,779839
0,4	5	4,941217	3,728173	1,213044
0,5	7	5,250721	3,874572	1,376149
0,6	6	5,618488	4,011695	1,606793
0,7	8	5,887612	4,121038	1,766574
0,8	13	6,056256	4,240442	1,815814
0,9	12	6,198774	4,388237	1,810536

Pemilihan parameter lambda optimal didasarkan pada jumlah titik *out-of-control* (OOC) terbanyak dan nilai absolut dari selisih antara UCL dengan nilai maksimum statistik M_i terkecil ($|UCL - \text{Max } M_i|$). Berdasarkan hasil analisis yang disajikan pada Tabel 4.6, parameter lambda optimal untuk analisis diagram kontrol Max-MEWMA fase I adalah $\lambda = 0,2$, dengan nilai UCL sebesar 3,354. Oleh karena itu, dalam studi ini parameter $\lambda = 0,2$ dipilih sebagai parameter eksponensial dalam penerapan diagram kontrol Max-MEWMA pada data fase I untuk meningkatkan sensitivitas terhadap deteksi variasi kecil dalam proses produksi.

Untuk membuat diagram kontrol Max-MEWMA diperlukan perhitungan nilai vektor rata-rata (U_i) dan variabilitas proses (V_i). Sebagai contoh perhitungan statistik diagram kontrol Max-MEWMA dengan $\lambda = 0,2$ adalah sebagai berikut.

Langkah awal dimulai dengan menentukan selisih antara nilai residual pengamatan ke-1 dengan rata-rata residual seluruh data pada masing-masing variabel, yaitu *Fuel Gas Flow*, *Combustion Pressure*, dan *Gas Temperature*. Selisih tersebut kemudian dikalikan dengan parameter peluruhan λ untuk menghasilkan vektor Z_1 , yang merepresentasikan penyimpangan terhalus dari rata-rata pada saat awal proses. Berdasarkan data residual fase I, diperoleh

$$x_1 = \begin{bmatrix} -0,1894 \\ -0,0217 \\ 0,0272 \end{bmatrix}, \mu_0 = \begin{bmatrix} -6,28 \times 10^{17} \\ -1,82 \times 10^{-16} \\ -1,59 \times 10^{-17} \end{bmatrix}$$

$$Z_1 = \lambda(x_1 - \mu_0) = 0,2 \cdot \begin{bmatrix} -0,1894 \\ -0,0217 \\ 0,0272 \end{bmatrix} = \begin{bmatrix} -0,0379 \\ -0,0043 \\ 0,0054 \end{bmatrix}$$

Setelah mendapatkan Z_1 , dilakukan pengukuran kontribusi multivariatnya terhadap perubahan proses dengan menghitung nilai statistik T_1 . Nilai ini diperoleh dari perkalian $Z_1^T \Sigma_0^{-1} Z_1$, yang dikalibrasi dengan faktor koreksi terhadap efek peluruhan λ pada posisi waktu ke-1. Hasil perhitungan menunjukkan bahwa.

$$T_1 = Z_1^T \Sigma_0^{-1} Z_1 \cdot \left(\frac{2-\lambda}{\lambda(1-(1-\lambda)^2(1))} \right)$$

$$T_1 = 9,119$$

Nilai T_1 kemudian ditransformasikan ke dalam bentuk statistik normal standar melalui fungsi distribusi kumulatif *chi-square* dengan derajat bebas 3. Diperoleh probabilitas kumulatif.

$$H_1 = P(\chi_3^2 \leq T_1) = 0,9722$$

Kemudian nilai U_i diperoleh dari fungsi kuantil distribusi normal standar.

$$U_1 = \Phi^{-1}(0,9722) = 1,915$$

Secara paralel, untuk memantau perubahan variabilitas proses, dihitung jarak *Mahalanobis* antara x_1 dan μ_0 , yang kemudian dikonversi ke probabilitas distribusi kumulatif *chi-square*.

$$\phi_1 = P(\chi_3^2 \leq D^2) = 0,9722$$

Nilai ϕ_1 ini kemudian digunakan untuk membentuk statistik eksponensial Y_i .

$$Y_1 = \lambda \cdot \phi_1 = 0,2 \cdot 0,9722 = 0,1944$$

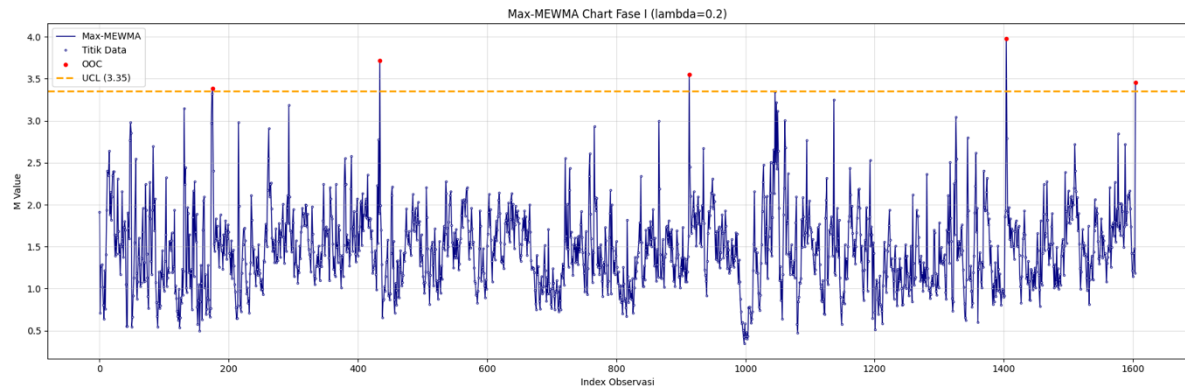
Nilai ini dikalibrasi menggunakan faktor *scaling* sebagai berikut:

$$V_1 = \sqrt{\frac{2-\lambda}{\lambda(-1(1-\lambda)^2)}} \cdot Y_1 = 0,9722$$

Akhirnya, statistik pengendali Max-MEWMA untuk pengamatan ke-1 ditentukan dengan mengambil nilai maksimum dari nilai absolut U_i dan V_i .

$$M_1 = \max(|U_i|, |V_i|) = \max(1,915; 0,9722) = 0,915$$

Berdasarkan hasil perhitungan di atas, diperoleh bahwa nilai statistik Max-MEWMA pada observasi ke-1 adalah sebesar 0,915. Nilai ini berada di bawah batas kendali atas (UCL) yang telah ditetapkan sebelumnya, sehingga dapat disimpulkan bahwa pengamatan pertama berada dalam kondisi *in-control*, atau tidak menunjukkan gejala penyimpangan proses yang signifikan secara statistik. Pada setiap titik waktu, nilai statistik MEWMA dibandingkan dengan batas kendali atas (UCL). Jika nilai M_i melebihi batas tersebut, maka hal tersebut menunjukkan adanya indikasi pergeseran dalam proses. Hasil diagram kontrol Max-MEWMA untuk data fase I yang menggunakan nilai UCL disajikan sebagai berikut.



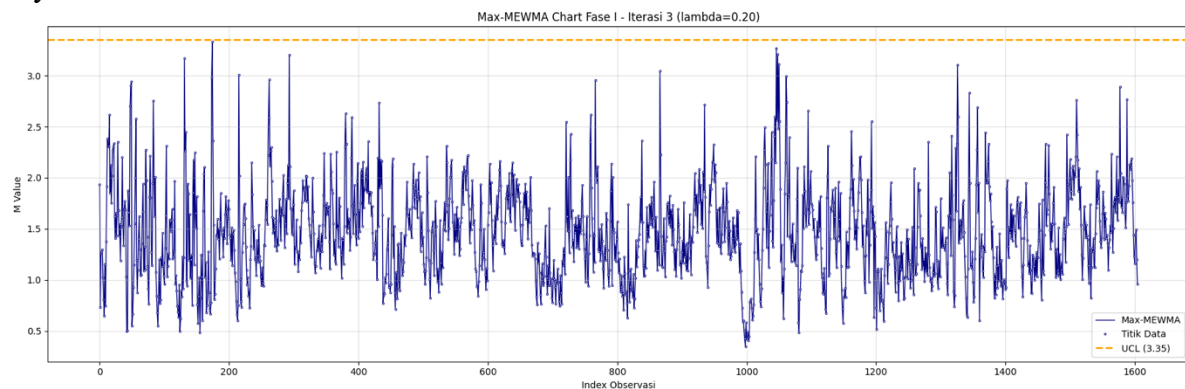
Gambar 4. 8 Diagram kendali Max - MEWMA $\lambda=0,2$ fase I

Berdasarkan Gambar 4.9, terlihat bahwa pada diagram kontrol Max-MEWMA untuk data residual fase I, sumbu horizontal menunjukkan urutan observasi, sedangkan sumbu vertikal merepresentasikan nilai statistik Max-MEWMA (M_i) untuk setiap titik pengamatan. Garis putus-putus berwarna oranye menunjukkan batas kendali atas (UCL), yang ditetapkan sebesar 3,354 berdasarkan simulasi *bootstrap* dengan parameter peluruhan $\lambda = 0,2$.

Hasil pengamatan terhadap grafik menunjukkan bahwa sebagian besar nilai M_i berada di bawah batas kendali, yang menandakan bahwa proses berada dalam kondisi yang relatif stabil. Namun demikian, terdapat 5 titik observasi yang berada di atas UCL dan dikategorikan sebagai *out-of-control* (OOC), mengindikasikan adanya potensi penyimpangan terhadap kestabilan proses.

Untuk memastikan analisis selanjutnya hanya mempertimbangkan titik-titik yang berada dalam kondisi terkendali, dilakukan proses pembersihan terhadap data fase I. Pembersihan dilakukan secara iteratif, yaitu dengan mengganti nilai residual menjadi nol pada setiap observasi dengan $M_i > \text{UCL}$, kemudian menghitung ulang nilai M_i hingga tidak ada lagi titik yang melebihi batas kendali. Proses ini diulang sampai seluruh observasi berada dalam batas pengendalian statistik.

Gambar berikut menunjukkan grafik Max-MEWMA setelah proses pembersihan iteratif dilakukan. Setiap titik pada grafik merepresentasikan nilai M_i dari data residual yang telah dinyatakan berada dalam kondisi kendali statistik.



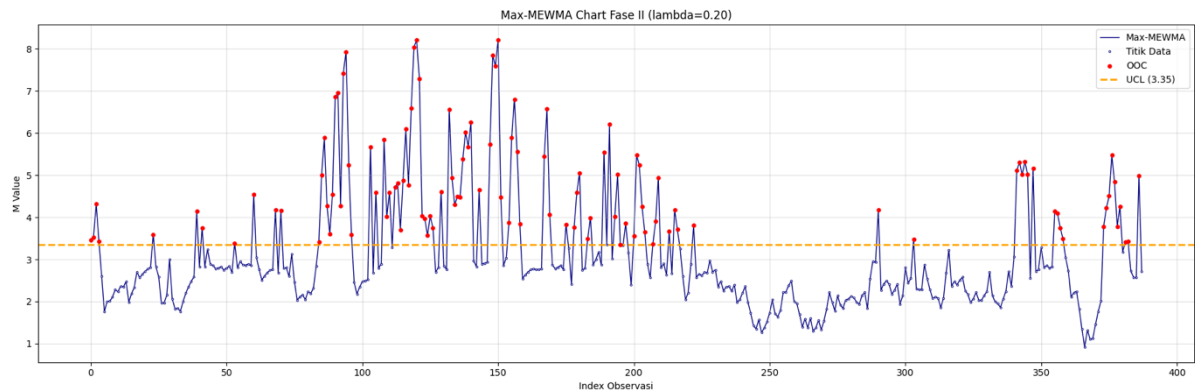
Gambar 4. 9 Diagram kendali Max - MEWMA $\lambda=0,2$ fase I *in control*

Berdasarkan Gambar 4.10, terlihat bahwa setelah dilakukan pembersihan data secara iteratif hingga iterasi ke-3, seluruh nilai pengamatan telah berada di bawah batas kendali atas (UCL). Hal ini menunjukkan bahwa tidak ada lagi observasi yang melebihi batas kendali,

sehingga dapat disimpulkan bahwa setelah proses pembersihan, variabilitas sistem *Heat Recovery Steam Generator* (HRSG) pada fase I telah terkendali secara statistik.

4.5.2 Diagram Kontrol Max-MEWMA Pada Data Fase II

Pengendalian proses pada fase II dilakukan dengan menerapkan batas kendali atas (UCL) sebesar 3,354 dan parameter mean kovarians, yang diperoleh dari analisis fase I parameter peluruhan $\lambda = 0,2$. Data fase II dianalisis menggunakan batas kendali tersebut untuk menilai apakah proses berjalan dalam kondisi yang stabil.



Gambar 4. 10 Diagram kendali Max - MEWMA $\lambda=0,2$ fase II

Berdasarkan Gambar 4.11, terlihat bahwa diagram kontrol Max-MEWMA pada fase II menunjukkan proses yang masih mengalami fluktuasi signifikan secara statistik, ditandai dengan banyaknya titik observasi yang berada di atas batas kendali atas (UCL = 3,354). Hasil analisis menunjukkan terdapat 114 titik *out-of-control* (OOC) pada data fase II. Hal ini mengindikasikan bahwa proses pada sistem *Heat Recovery Steam Generator* (HRSG) di PLTGU PT. PLN Nusantara Power UP Gresik pada fase II belum sepenuhnya berada dalam keadaan yang terkendali secara statistik. Keberadaan sejumlah besar titik OOC ini menjadi perhatian penting dan memerlukan evaluasi lebih lanjut terhadap potensi gangguan maupun penyimpangan dalam proses yang sedang berjalan. Selanjutnya, dilakukan proses identifikasi lebih lanjut terhadap setiap titik OOC yang terdeteksi pada fase II. Identifikasi ini bertujuan untuk menelusuri penyebab terjadinya penyimpangan dan memberikan rekomendasi tindakan perbaikan yang sesuai guna meningkatkan kestabilan proses di masa mendatang.

4.6 Identifikasi Out Of Control

Identifikasi titik *out-of-control* (OOC) dilakukan untuk mengetahui karakteristik kualitas sistem *Heat Recovery Steam Generator* (HRSG) di PLTGU PT PLN Nusantara Power UP yang dapat memengaruhi efisiensi kinerja sistem. Berdasarkan hasil analisis pada fase II menggunakan diagram kontrol Max-MEWMA, diperoleh bahwa sistem HRSG tidak berada dalam kondisi terkendali secara statistik, yang ditandai dengan munculnya beberapa sinyal OOC. Untuk menelusuri penyebab terjadinya kondisi tidak terkendali, dilakukan identifikasi titik *out-of-control* (OOC) berdasarkan dua aspek utama.

4.5.1 Identifikasi Berdasarkan Statististik U_i dan V_i

Pendekatan analisis dilakukan dengan membandingkan nilai statistik U_i dan V_i pada setiap titik yang teridentifikasi sebagai *out-of-control* (OOC) pada diagram kontrol Max-MEWMA fase II. Dalam konteks pengendalian kualitas multivariat, U_i mengindikasikan

pergeseran rata-rata (*mean shift*), sedangkan V_i mencerminkan perubahan variansi (*variance shift*).

Berdasarkan hasil deteksi OOC dengan parameter $\lambda = 0,2$ dan UCL sebesar 3,354, ditemukan 114 titik OOC pada fase II. Hasil analisis menunjukkan bahwa seluruh titik OOC disebabkan oleh nilai U_i yang melebihi batas kendali, sementara tidak terdapat titik OOC yang disebabkan oleh V_i . Temuan ini mengindikasikan bahwa sinyal penyimpangan yang teridentifikasi pada proses lebih banyak berkaitan dengan pergeseran rata-rata (*mean shift*), bukan perubahan ragam atau variansi proses (*variance shift*).

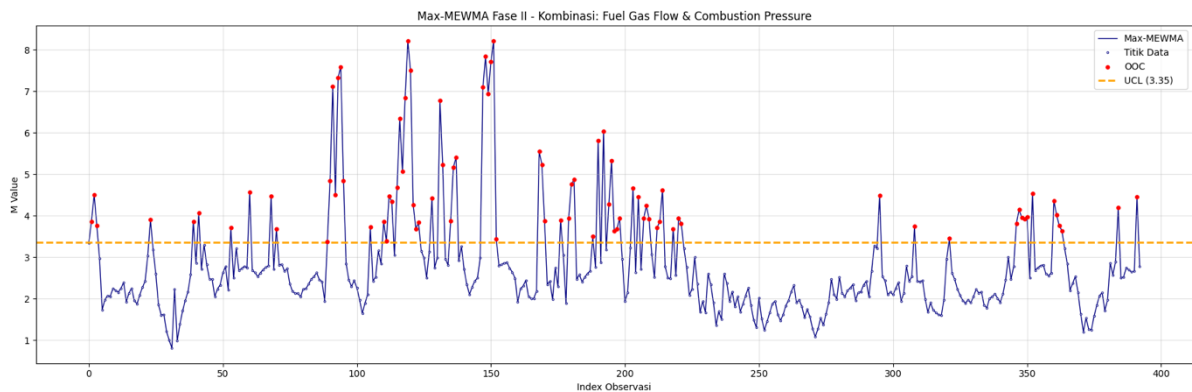
Dengan demikian, dapat disimpulkan bahwa pada fase II, sistem *Heat Recovery Steam Generator* (HRSG) di PLTGU PT. PLN Nusantara Power UP Gresik lebih sensitif terhadap perubahan rata-rata data residual dibandingkan perubahan variasi. Oleh karena itu, strategi pengendalian proses pada fase ini sebaiknya difokuskan pada upaya untuk menjaga stabilitas rata-rata proses, seperti monitoring ketat terhadap parameter proses yang dapat menyebabkan terjadinya mean shift, evaluasi dan pemeliharaan peralatan secara berkala, serta peningkatan sistem deteksi dini untuk mengantisipasi pergeseran rata-rata yang tidak diharapkan.

4.5.2 Identifikasi Berdasarkan Kombinasi Karakteristik Kualitas

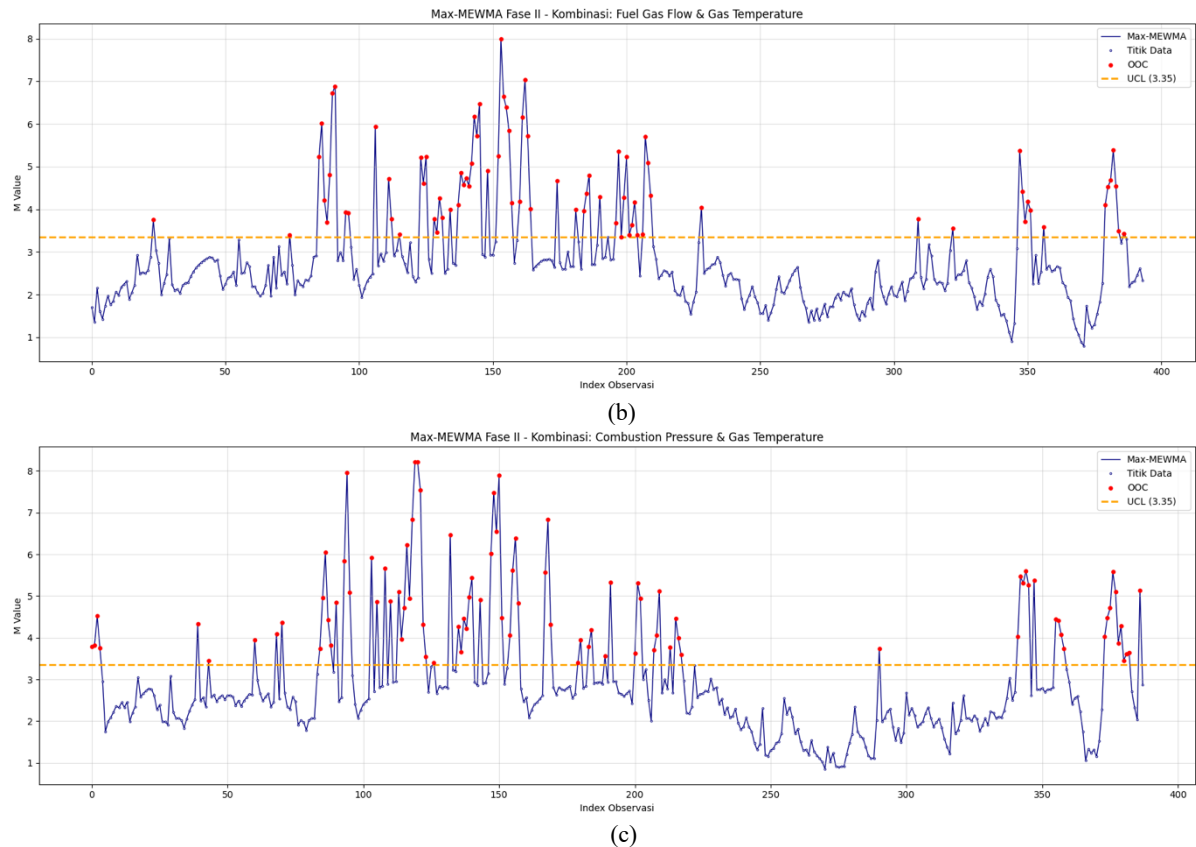
Selain mengamati sinyal *out-of-control* (OOC) secara statistik melalui komponen U_i dan V_i , analisis lanjutan dilakukan untuk menelusuri kontribusi relatif dari masing-masing variabel residual terhadap kemunculan sinyal OOC. Salah satu pendekatan yang digunakan dalam penelitian ini adalah dengan membandingkan performa pengendalian proses dari beberapa kombinasi pasangan variabel residual.

Tujuan dari analisis ini adalah untuk mengidentifikasi variabel mana yang secara konsisten terlibat dalam sinyal OOC, sehingga dapat diindikasikan sebagai faktor dominan penyebab ketidakterkendalian proses. Dengan mengetahui kombinasi pasangan variabel yang paling sering menghasilkan sinyal OOC, maka dapat ditentukan variabel yang paling berpengaruh terhadap kestabilan sistem produksi.

Dalam penelitian ini, dilakukan pengujian terhadap tiga kombinasi pasangan variabel residual, yaitu pasangan *Fuel Gas Flow* (FGF) dan *Combustion Pressure* (CP), pasangan *Fuel Gas Flow* (FGF) dan *Gas Temperature* (GT), serta pasangan *Combustion Pressure* (CP) dan *Gas Temperature* (GT). Masing-masing kombinasi diuji menggunakan diagram kontrol Max-MEWMA dengan nilai parameter bobot $\lambda = 0,2$ dan UCL sebesar 3,354.



(a)



Gambar 4.11 Kombinasi (a) Fuel Gas Flow dan Combustion Pressure, (b) Fuel Gas Flow dan Gas Temperatur, dan (c) Combustion Pressure dan Gas Temperatur

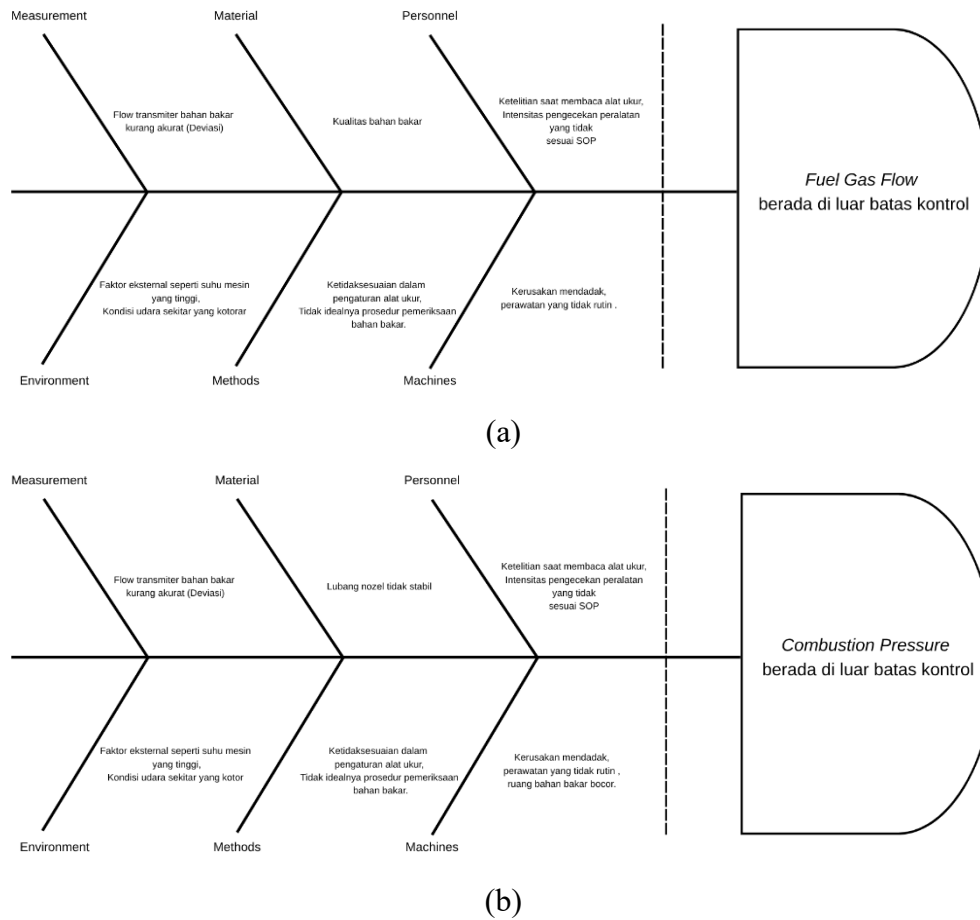
Berdasarkan Gambar 4.12, hasil analisis diagram kontrol Max-MEWMA pada kombinasi karakteristik kualitas menunjukkan bahwa pasangan FGF dan CP menghasilkan jumlah titik OOC tertinggi, yaitu sebanyak 92 titik OOC. Kombinasi CP dan GT menghasilkan 84 titik OOC, sedangkan kombinasi FGF dan GT menghasilkan 79 titik OOC. Temuan ini mengindikasikan bahwa pasangan FGF dan CP memiliki potensi ketidakstabilan proses yang paling dominan pada kestabilan proses produksi pada sistem *Heat Recovery Steam Generator* (HRSG), sehingga perlu mendapat perhatian lebih lanjut dalam upaya pengendalian dan perbaikan proses.

Kondisi ini menunjukkan bahwa ketidakterkendalian proses lebih sering dipicu oleh fluktuasi pada aliran bahan bakar (FGF) dan tekanan pembakaran (CP). Terutama ketika kedua variabel tersebut berubah secara bersamaan, dapat terjadi gangguan signifikan terhadap kestabilan proses. Pergeseran kecil dalam aliran bahan bakar, jika tidak diiringi dengan penyesuaian tekanan pembakaran, dapat menyebabkan ketidakseimbangan termal dan menurunkan efisiensi konversi panas pada sistem HRSG.

Oleh karena itu, penyebab utama ketidakterkendalian proses pada sistem ini dapat diarahkan pada ketidakteraturan atau variabilitas dari FGF dan CP, baik secara individu maupun dalam kombinasi. Informasi ini menjadi penting dan strategis sebagai dasar dalam proses evaluasi sistem produksi, penyusunan rencana pengendalian kualitas, serta dalam perumusan strategi perbaikan yang lebih terarah, terutama dalam menjaga kestabilan parameter utama seperti aliran bahan bakar dan tekanan pembakaran.

Untuk mendukung proses identifikasi akar penyebab, dapat digunakan pendekatan diagram sebab-akibat (Ishikawa), yang mengelompokkan faktor-faktor pengaruh berdasarkan kategori seperti manusia, mesin, metode, material, lingkungan, dan pengukuran. Diagram ini

memberikan gambaran sistematis terhadap kemungkinan penyebab dominan dari sinyal OOC dan ditunjukkan pada Gambar 4.11 berikut



Gambar 4. 12 Diagram Ishikawa (a) *Fuel Gas Flow* dan (b) *Combustion Pressure*

Berdasarkan Gambar 4.11 (a), yang menampilkan diagram Ishikawa untuk karakteristik kualitas laju aliran bahan bakar, dapat diidentifikasi sejumlah faktor utama yang berkontribusi terhadap kondisi proses yang tidak terkendali pada sistem HRSG di PLTGU PT. PLN Nusantara Power UP Gresik. Enam kategori utama penyebab ketidakterkendalian meliputi faktor manusia, mesin, metode, bahan baku, lingkungan, dan pengukuran. Dari sisi manusia, ketidakteelitian dalam pengukuran, pemeriksaan *flow transmitter*, serta ketidaksesuaian prosedur kerja dengan SOP menjadi sumber potensi deviasi. Faktor mesin mencakup suhu mesin yang terlalu tinggi dan kurangnya pemeliharaan rutin, yang dapat berdampak pada ketidakstabilan gas buang. Pada aspek bahan baku, perbedaan kualitas minyak hidrolik dan kondisi *air intake filter* turut memengaruhi performa kontrol valve dan aliran bahan bakar. Sementara itu, pada sisi metode, pengaturan alat ukur yang tidak optimal dan prosedur pemeriksaan yang belum memadai turut memperbesar kemungkinan terjadinya sinyal OOC. Di sisi pengukuran, akurasi *flow transmitter* bahan bakar yang rendah menyebabkan deviasi yang dapat membawa proses keluar dari batas kendali.

Selanjutnya, Gambar 4.11(b) menunjukkan diagram Ishikawa untuk karakteristik tekanan pembakaran, yang juga menjadi salah satu indikator tidak efisiennya sistem HRSG. Enam faktor penyebab yang serupa kembali teridentifikasi. Dari sisi tenaga kerja, ketidakteelitian

dalam pengecekan flow transmitter dan penerapan prosedur kerja yang tidak disiplin kembali menjadi sorotan. Masalah pada mesin, seperti suhu tinggi dan kurangnya perawatan, berdampak pada stabilitas hasil pembakaran. Dari aspek bahan baku, material *nozzle* yang tidak konsisten menyebabkan pembakaran yang tidak merata. Metode yang kurang tepat dalam pengaturan alat ukur dan inspeksi bahan bakar menambah potensi deviasi. Terakhir, pengukuran yang tidak akurat pada *flow transmitter* tekanan pembakaran turut memicu sinyal OOC yang menjauhkan proses dari kondisi stabil.

4.7 Analisis Kapabilitas Proses

Analisis kapabilitas proses digunakan untuk mengevaluasi sejauh mana suatu proses mampu memenuhi spesifikasi yang telah ditetapkan, serta mencerminkan kualitas kinerja proses secara keseluruhan. Suatu proses dikatakan kapabel apabila telah berada dalam kondisi terkendali secara statistik dan seluruh produk yang dihasilkan berada dalam batas spesifikasi yang ditentukan oleh perusahaan. Kapabilitas ini biasanya dinyatakan melalui nilai indeks kapabilitas, di mana nilai lebih dari satu menunjukkan bahwa proses memiliki kemampuan yang baik dalam menghasilkan produk sesuai standar.

Dalam konteks ini, indeks kapabilitas yang digunakan bergantung pada jenis data yang dianalisis. Untuk data multivariat, digunakan indeks MP_p dan MP_{pk} sebagai ukuran kapabilitas proses, sedangkan untuk data univariat digunakan indeks P_p dan P_{pk} . Analisis kapabilitas proses secara univariat dilakukan secara terpisah untuk setiap karakteristik kualitas dari sistem *Heat Recovery Steam Generator* (HRSG) pada fase II, sebagaimana akan dijelaskan pada bagian berikutnya.

Tabel 4.7 Perhitungan Kapabilitas Proses secara Univariat

Karakteristik Kualitas	P_p	P_{pk}
<i>Fuel Gas Flow</i> (FGF)	2,0882	1,7784
<i>Combustion Pressure</i> (CP)	0,8420	0,6833
<i>Gas Temperature</i> (GT)	0,5977	0,5719

Tabel 4.7 menunjukkan bahwa nilai C_p dan C_{pk} untuk masing-masing karakteristik kualitas sistem *Heat Recovery Steam Generator* (HRSG) memiliki nilai . Kemudian melakukan analisis kapabilitas proses secara multivariat dimana untuk perhitungannya sebagai berikut.

$$MP_p = \frac{1}{3} \times (2,0882 + 0,8420 + 0,5977) = \frac{1}{3} \times 3,528 = 1,176$$

$$MP_{pk} = \frac{1}{3} \times (1,7784 + 0,6833 + 0,5719) = \frac{1}{3} \times 3,0336 = 1,0112$$

Hasil perhitungan di atas menunjukkan bahwa nilai MP_p dan MP_{pk} keduanya berada di atas angka satu, yang secara umum mengindikasikan bahwa proses memiliki kapabilitas yang baik. Namun demikian, interpretasi terhadap kapabilitas proses tidak dapat hanya didasarkan pada nilai indeks semata. Suatu proses hanya dapat dinyatakan benar-benar kapabel apabila proses tersebut telah berada dalam kondisi terkendali secara statistik. Berdasarkan hasil analisis menggunakan diagram kendali Max-MEWMA berbasis residual *Random Forest Regressor* pada fase II, diketahui bahwa proses masih belum sepenuhnya terkendali. Oleh karena itu,

meskipun nilai indeks menunjukkan kapabilitas, secara keseluruhan proses belum dapat dinyatakan kapabel karena belum memenuhi syarat kestabilan proses secara statistik.

BAB V

KESIMPULAN DAN SARAN

5.1 Kesimpulan

Berdasarkan hasil analisis yang telah dilakukan, didapatkan kesimpulan sebagai berikut:

1. Pengendalian proses dilakukan dengan menerapkan diagram kontrol Max-MEWMA berbasis residual dari model *Random Forest Regressor*. Input dari model *Random Forest Regressor* ditentukan berdasarkan kombinasi variabel *Fuel Gas Flow* (FGF), *Combustion Pressure* (CP), dan *Gas Temperature* (GT) sebagai prediktor dalam membentuk model residual multivariat. Penentuan input model dilakukan berdasarkan lag-lag signifikan yang diperoleh dari hasil analisis *Partial Autocorrelation Function* (PACF) terhadap masing-masing karakteristik kualitas proses. Hasil analisis menunjukkan bahwa karakteristik *Fuel Gas Flow* (FGF) memiliki lag signifikan pada lag ke-1, 2, 3, 4, dan 5. Sementara itu, lag signifikan pada karakteristik *Combustion Pressure* (CP) teridentifikasi pada lag ke-1, 2, 3, 4 dan 5. Adapun karakteristik *Gas Temperature* (GT) memiliki lag signifikan pada lag ke-1, 3, dan 4. Model *Random Forest* yang dibentuk dengan konfigurasi tersebut menghasilkan residual yang telah mampu mengurangi autokorelasi dalam data pengamatan. Residual hasil model tersebut kemudian digunakan dalam pengendalian proses menggunakan diagram kontrol Max-MEWMA dengan $\lambda = 0,2$ yang diterapkan dalam dua fase. Pada fase I, dilakukan iterasi pembersihan titik *out-of-control* hingga seluruh data berada dalam kendali statistik dan menghasilkan batas kendali atas (UCL) sebesar 3,354. Nilai UCL ini kemudian digunakan sebagai acuan dalam pengendalian proses pada fase II untuk memantau kestabilan sistem secara menyeluruh.
2. Berdasarkan nilai statistik U_i dan V_i dan kombinasi karakteristik kualitas *out of control* disebabkan oleh pergeseran rata-rata (*mean shift*). Kombinasi karakteristik *Fuel Gas Flow* dan *Combustion Pressure* secara konsisten menghasilkan jumlah sinyal OOC tertinggi menjadikan karakteristik kualitas yang paling berpengaruh terhadap ketidakterkendalian proses.
3. Berdasarkan hasil perhitungan kapabilitas proses, diketahui bahwa secara multivariat kinerja *Heat Recovery Steam Generator* (HRSG) di PLTGU PT PLN Nusantara Power UP Gresik belum dapat dikatakan kapabel baik dari segi presisi maupun akurasi. Meskipun nilai indeks MP_p dan MP_{pk} berada di atas satu, maka sistem belum dapat dinyatakan kapabel secara menyeluruh. Artinya, meskipun secara presisi dan akurasi tergolong baik, kestabilan sistem masih perlu ditingkatkan.

5.2 Saran

Adapun saran yang diberikan untuk kebaikan dan kemajuan penelitian lebih lanjut adalah sebagai berikut:

1. Jika penelitian ini dilanjut bisa menggunakan pemodelan yang berbeda seperti model prediktif berbasis *deep learning* seperti LSTM.
2. Penelitian ini bisa dilanjutkan dengan menggunakan turbin gas dan menambah karakteristik kualitas.
3. Jika peneltian ini dilanjut, sebaiknya lebih difokuskan pada sistem di turbin gas dengan menggunakan variabel pada proses pembakaran.

(halaman ini sengaja dikosongkan)

DAFTAR PUSTAKA

- Rohana. (2008). Analisis Perilaku Pelanggan PLN dalam Pemakaian Daya Listrik. *Sarjana Teknik Industri Universitas Sumatera Utara*.
- Tanjung, Y. P., Sentinuwo, S., & Jacobus, A. (2016, 11 17). Penentuan Daya Listrik Rumah Tangga Menggunakan Metode Decision Tree. *Jurnal Teknik Informatika Universitas Sam Ratulangi*, Vol 9(Vol. 9 No. 1 (2016): Jurnal Teknik Informatika). doi:<https://doi.org/10.35793/jti.v9i1.14141>
- Widiawati, W. Y., & Atok, R. M. (2018). Analisis Klasifikasi Pelanggan Listrik Rumah Tangga Bersubsidi Kota Surabaya Menggunakan Support Vector Machine dan Naive Bayes Classifier. *Jurnal Sains dan Seni ITS*(JURNAL SAINS DAN SENI ITS Vol. 7, No. 2 (2018)).
- Yulia, & Azwanti, N. (2018, August). Penerapan Algoritma C4.5 Untuk Memprediksi Besarnya Penggunaan Listrik Rumah Tangga di Kota Batam. *Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)*, Vol 2. No.2. doi:10.29207/resti.v2i2.503
- Saleh, A. (2015, Mei). Implementasi Metode Klasifikasi Naive Bayes Dalam Memprediksi Besarnya Penggunaan Listrik Rumah Tangga. *AMIKOM OJS Journal*, Vol. 2 No. 3.
- Sumarna, & Triharso, I. B. (2020). *Dasar Teknik Listrik*. Gramedia Pustaka.
- Witten, I., Frank, E., & Hall, M. (2016). *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann.
- Tan, P.-N., Steinbach, M., & Kumar, V. (2006). *Introduction to Data Mining*. Boston, Massachusetts, AS: Pearson/Addison Wesley.
- Mitchell, T. (1997). *Machine Learning*. McGraw-Hill.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer.
- Amin, R. K., Indwiarti, D., & M.Si Yuliant Sibaroni, S. M. (2015). Implementasi Klasifikasi Decision Tree Dengan Algoritma C4.5 Dalam Pengambilan Keputusan Permohonan Oleh Debitur . *Studi Kasus: Bank Pasar Daerah Istimewa Yogyakarta*, 1768.
- International Energy Agency. (2023). *World Energy Outlook*. International Energy Agency. Retrieved from <https://www.iea.org/reports/world-energy-outlook-2023>
- Intergovernmental Panel on Climate Change. (2021). *Climate Change 2021: The Physical Science Basis*. IPCC. Retrieved from https://www.ipcc.ch/report/ar6/wg1/downloads/report/IPCC_AR6_WGI_Full_Report.pdf

- Kementerian Energi dan Sumber Daya Mineral. (2022). *Statistik Ketenagalistrikan 2022*. Kementerian Energi dan Sumber Daya Mineral. Retrieved from <https://www.esdm.go.id/>
- Montgomery, D. C. (2020). *Introduction to Statistical Quality Control*. John Wiley; Sons.
- Vapnik, V. (1995). *The Nature of Statistical Learning Theory*. Springer.
- Huwang, J., Yeh, A. B., & Wu, Z. (2007). Multivariate control charts for detecting small process shifts. *Journal of Quality Technology*, 39(1), 52-64.
- Khusna, H., Mashuri, M., Suhartono, Prastyo, D. D., & Ahsan, M. (2018). Multioutput least square SVR-based multivariate EWMA control chart: The performance evaluation and application. *Cogent Engineering*, 5(1). doi:<https://doi.org/10.1080/23311916.2018.1531456>
- Syahrozahdl, N. L. (2023). Pengendalian Kualitas Sistem HRSG dengan MEWMV dan MEWMA Berbasis Residual MLS-SVR di PT PLN Nusantara Power Unit Pembangkitan Gresik. *Tugas Akhir Institut Teknologi Sepuluh Nopember*.
- Iskandar, E. A. (2020). Pengendalian Kualitas Sistem Heat Recovery Steam Generator (HRSG) PLTGU di PT. X dengan Diagram Kontrol Maximum.
- Ilmar, A., & Sandra, A. (2012). Analisis Unjuk Kerja Heat Recovery Steam Generator (HRSG) pada PLTGU Muara Tawar Blok 5.
- Aulia, T. R. (2021). Pengendalian Kualitas Proses Produksi Air di PDAM Tirtanadi Provinsi Sumatera Utara Menggunakan Diagram Kontrol MEWMA Berbasis Residual Model Multioutput Least Square Support Vector Regression (MLS-SVR).
- Yunika, A. (2023). Pengendalian Kualitas Proses Produksi Gula Kristal Putih di Pabrik Gula Djatiroto PT Perkebunan Nusantara XI Menggunakan Diagram Kontrol Max-MCUSUM Berbasis Residual Model MLS-SVR.
- Setyoko, B. (2006). Analisa Efisiensi Performa HRSG pada PLTGU.
- Syahrozahdl, N. L. (2023). Pengendalian Kualitas Sistem Heat Recovery Steam Generator (HRSG) di PLTGU PT PLN Nusantara Power Unit Pembangkitan Gresik Menggunakan Diagram Kontrol MEWMV dan MEWMA Berbasis Residual Model Multioutput Least Square Support Vector Regression (MLS-SVR).
- Kumparan. (2023). *Pengertian Energi Listrik: Sumber, Manfaat, dan Tantangan Penggunaannya*. Retrieved from <https://kumparan.com/pengertian-dan-istilah/pengertian-energi-listrik-sumber-manfaat-dan-tantangan-penggunaannya-21XmpF2NSNF>
- Sainstekno. (2024). *Produksi Listrik: Proses, Jenis, dan Teknologi Terkini*. Retrieved from Sainstekno: <https://sainstekno.net/2024/12/30/produksi-listrik-proses-jenis-dan-teknologi-terkini/>
- Kimia, P. (2023). *Energi Listrik: Pengertian, Rumus, dan Contoh Soal*. Retrieved from Pakar Kimia: <https://www.pakarkimia.com/energi-listrik/>

- Polines. (2024). Efisiensi HRSG dalam Sistem PLTGU. *Eksergi*. Retrieved from <https://jurnal.polines.ac.id>
- Walpole, R. E. (1995). *Pengantar Statistika. Alih bahasa : Ir. Bambang Sumantri*. Jakarta: PT. Gramedia Pustaka Utama.
- Johnson, R., & Wichern, D. (2007). *Applied Multivariate Analysis*.
- Mardia, K. V. (1970). *Multivariate Analysis*.
- Wei, W. W. (2006). *Time Series Analysis: Univariate and Multivariate Methods*. Pearson Addison Wesley.
- Xu, S., An, X., Qiao, X., Zhu, L., & Li, L. (2013). Multi-Output Least-Squares Support Vector Regression Machines. *Pattern Recognition Letters*.
- Khusna, H., Mashuri, M., Suhartono, Prastyo, D. D., & Ahsan, M. (2019). Multioutput least square SVR based multivariate EWMA. *Journal of Physics : Conference Series* .
- Heskes, T. M. (2000). *Empirical Bayes for Learning to Learn*. San Fransisco.
- Hsu, C. W., Chang, C., & Lin, C. (2016). *A Practical Guide to Support Vector Classification*. . National Taiwan University.
- Xu, S., Ma, F., & Tao, L. (2007). *Learn from the information contained in the false splice sites as well as in the true splice sites using SVM*.
- Raissi, S. (2009). Multivariate Process Capability Indices on The Presence of Priority for Quality Characteristics. *Journal of Industrial Engineering International, Vol. 5 No. 9*, 27-36.
- Ridgon, S. E., & Champ, C. W. (1987). *Multivariate Geometric Moving Average Charts*.
- Lowry, J. M. (1989). *A Multivariate Exponentially Weighted*.
- Morrison, D. F. (2005). *Multivariate Statistical Methods* (4th ed.). Jakarta: The Whartoon School University of Pennsylvania.
- Walpole, R. E., Myers, R. H., Myers, S. L., & Ye, K. (2012). *Probability & statistics for engineers & scientists* (9th ed.). Boston: Pearson Education, Inc.
- Duyo, R. A. (2020). Analisis penyebab gangguan jaringan pada distribusi listrik menggunakan metode fault tree analysis di PT. PLN (Persero) Rayon Daya Makassar. *Jurnal Teknik Elektro UNISMUH*.
- Fachrizal, W. (2020). Pengendalian kualitas garam konsumsi beryodium di PT Garam (Persero) menggunakan Diagram Maximum Multivariate Exponential Weighted Moving Average (MAX-MEWMA). *Tugas Akhir Institut Teknologi Sepuluh Nopember*.
- Chen, G., Cheng, S. W., & Xie, H. (2005). A New Multivariate Control Chart for Monitoring Both Location and Dispersion. *Communications in Statistics-Simulation and Computation*, 203-217.

- Shapiro, S., & Wilk, M. (1965). An Analysis of Variance Test for Normality (Complete Samples). *Biometrika*, 591-611.
- Breiman, L. (2001). Machine Learning. *Random Forest*.
- Wang, Z., Qin, S., & Ma, C. (2022). A Review on Data-Driven Process Monitoring Methods. *Processes*, 335.
- Weix, T., Resch, M., & Guh, R. S. (2022). Monitoring covariance in multivariate time series. *Comparing VAR with multivariate random forest. Journal of Quality Technology*, 227–242. doi: <https://doi.org/10.1080/00224065.2021.1884606>
- Prabowo, H., Suhartono, & Prastyo, D. D. (2020, March). The Performance of Ramsey Test, White Test and Terasvirta Test in Detecting Nonlinierity. *Inferensi*, 1-2.
- Prabowo, H., Suhartono, & Prastyo, D. D. (2020). The Performance of Ramsey Test. *White Test and Terasvirta Test in Detecting Nonlinierity*.
- Montgomery, D. C. (2013). *DESIGN AND ANALYSIS OF EXPERIMENTS* (8th ed.). Hoboken: John Wiley & Sons, Inc.
- Wei, W. W. (2006). *Time Series Analysis Univariate and Multivariate Methods*. New York: Pearson Education Inc.
- Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. New York: Springer.
- Johnson, R. A., & Wichern, D. W. (2007). *Applied Multivariate Statistical Analysis* (6th ed.). United States of America: Pearson Education, Inc.
- Ho, T. (1998). The random subspace method for constructing decision forests. *Pattern Analysis and Machine Intelligence. IEEE Transactions on*, 20(8):832– 844.
- Linusson, H., Rudenwall, R., & Olausson, A. (2012). Random forest och glesa datarepresentationer.
- Segal, & R., M. (1992). Tree-structured methods for longitudinal data. *Journal of the American Statistical Association*, 87(418):407–418.
- Glocker, B., Pauly, O., Konukoglu, E., & Criminisi, A. (2012). Joint classification-regression forests for spatially structured multi-object segmentation. *In Computer Vision–ECCV 2012*, 870–881.
- Linusson, H., Rudenwall, R., & Olausson, A. (2012). Random forest och glesa datarepresentationer.
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., . . . Duchesnay, É. (2011). Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12, 2825–2830.
- Parejo, A., Bracco, S., Personal, E., Larios, D. F., Delfino, F., & Leon, C. (2021). Short-Term Power Forecasting Framework for Microgrids Using Combined Baseline and Regression Models.

LAMPIRAN

Lampiran 1 Data GT 23 Operation 202207010005 atau Blok II pada Pola operasi (1,1,1) PLTGU PT. PLN Nusantara Power UP Gresik pada tanggal 1 Juli – 7 Juli 2022.

<i>Timestamp</i>	<i>Fuel Gas Flow</i>	<i>Combustion Pressure</i>	<i>Gas Temperature</i>
2022-07-01 00:00:00	24,7	9,2	544,5
2022-07-01 00:05:00	27,1	10,8	547,4
2022-07-01 00:10:00	26	9,4	542,1
2022-07-01 00:15:00	25,8	10	532,2
2022-07-01 00:20:00	28,1	11,7	527,2
2022-07-01 00:25:00	28,1	10,7	524,2
2022-07-01 00:30:00	25,9	9,3	520,2
2022-07-01 00:35:00	26	9,7	515,4
2022-07-01 00:40:00	25,4	9,8	514,7
2022-07-01 00:45:00	27,6	11,2	521,9
2022-07-01 00:50:00	26,9	10,8	529,3
2022-07-01 00:55:00	26,9	10,6	530,8
2022-07-01 01:00:00	26,1	9,7	530,2
2022-07-01 01:05:00	25,1	9,3	538,2
2022-07-01 01:10:00	25,8	9,4	543,5
2022-07-01 01:15:00	25,8	9,4	541
2022-07-01 01:20:00	25,4	9,3	534,9
2022-07-01 01:25:00	25,2	9,2	531,5
2022-07-01 01:30:00	25,3	9,2	543,2
2022-07-01 01:35:00	26,3	9,4	533,4
2022-07-01 01:40:00	25,9	9,4	539,7
2022-07-01 01:45:00	26,6	10,2	541,5
2022-07-01 01:50:00	25,6	9,3	543,2
2022-07-01 01:55:00	26,8	10,7	542,3
2022-07-01 02:00:00	27,6	10,5	534,8
2022-07-01 02:05:00	26,5	9,5	530,5
2022-07-01 02:10:00	27,8	10,7	533,6
2022-07-01 02:15:00	25,5	9,3	537,4
2022-07-01 02:20:00	25,7	9,3	543,3
2022-07-01 02:25:00	25,6	9,3	544,8
2022-07-01 02:30:00	25,6	9,3	546,2
2022-07-01 02:35:00	25,9	9,3	546,8
2022-07-01 02:40:00	25,9	9,3	546,4
2022-07-01 02:45:00	24,5	9,2	537,1
2022-07-01 02:50:00	25	9,3	532,9
2022-07-01 02:55:00	26	9,4	533,8
2022-07-01 03:00:00	25,5	9,3	541,6
2022-07-01 03:05:00	26,3	9,4	536,1
2022-07-01 03:10:00	26,3	9,3	531,5
2022-07-01 03:15:00	24,7	9,2	527,5
2022-07-01 03:20:00	25,6	9,4	541,2
⋮	⋮	⋮	⋮
2022-07-07 23:55:00	27,3	11	520

Lampiran 2 Syntax R Uji Normalitas

```
library(mvnormtest)
library(readr)
data <- read_csv("resid_fase1_stacked.csv")

# Pastikan hanya kolom numerik
data_numeric <- data[sapply(data, is.numeric)]

# Pastikan jumlah observasi > jumlah variabel
if (nrow(data_numeric) > ncol(data_numeric)) {
  # Jalankan uji normalitas multivariat Shapiro-Wilk
  result <- mshapiro.test(t(as.matrix(data_numeric)))

  # Tampilkan hasil
  print(result)
  cat("Nilai p-value:", result$p.value, "\n")
} else {
  cat("Error: Jumlah observasi harus lebih banyak dari jumlah variabel untuk
mshapiro.test.\n")
}
```

Lampiran 3 Output Uji Normalitas

```
Shapiro-Wilk normality test

data: Z
W = 0.96572
Nilai p-value: 4.67818e-19
```

Lampiran 4 Syntax R Uji Dependensi

```
library(readxl)
library(psych)

# Baca data fase 1 dan fase 2
fase1 <- read_excel("/Users/husnanalihusnain/Downloads/Skripsi/df_fase1.xlsx")
fase2 <- read_excel("/Users/husnanalihusnain/Downloads/Skripsi/df_fase2.xlsx")

# Buat data jadi numeric matrix (jika masih ada karakter)
data_fase1 <- as.matrix(fase1[sapply(fase1, is.numeric)])
data_fase2 <- as.matrix(fase2[sapply(fase2, is.numeric)])

# Uji Bartlett untuk fase 1
cat("UJI BARTLETT - FASE 1:\n")
bartlett_fase1 <- cortest.bartlett(cor(data_fase1), n = nrow(data_fase1))
print(bartlett_fase1)

# Uji Bartlett untuk fase 2
cat("\nUJI BARTLETT - FASE 2:\n")
bartlett_fase2 <- cortest.bartlett(cor(data_fase2), n = nrow(data_fase2))
print(bartlett_fase2)
```

Lampiran 5 Output Uji Dependensi

```
> # Uji Bartlett untuk fase 1
> cat("UJI BARTLETT - FASE 1:\n")
UJI BARTLETT - FASE 1:
> bartlett_fase1 <- cortest.bartlett(cor(data_fase1), n = nrow(data_fase1))
> print(bartlett_fase1)
$chisq
[1] 4289.82

$P.value
[1] 0

$df
[1] 3

> # Uji Bartlett untuk fase 2
> cat("\nUJI BARTLETT - FASE 2:\n")

UJI BARTLETT - FASE 2:
> bartlett_fase2 <- cortest.bartlett(cor(data_fase2), n = nrow(data_fase2))
> print(bartlett_fase2)
$chisq
[1] 882.6329

$P.value
[1] 5.176288e-191
```

```
$df  
[1] 3
```

Lampiran 6 Syntax Deteksi *Outlier*

```
# Deteksi outlier dengan z-score  
z_scores = df[['fgf', 'cp', 'gt']].apply(zscore)  
outlier_mask = (np.abs(z_scores) > 3).any(axis=1)  
df[outlier_mask] = np.nan  
df = df.interpolate(method='time')  
print(z_scores)
```


Lampiran 7 Syntax Random Forest Regressor

```
import pandas as pd
import numpy as np
from sklearn.preprocessing import MinMaxScaler
from sklearn.ensemble import RandomForestRegressor
from sklearn.multioutput import MultiOutputRegressor
from sklearn.model_selection import TimeSeriesSplit, GridSearchCV
from statsmodels.tsa.stattools import pacf

# 1. Load data
df_fase1 = pd.read_excel("df_fase1.xlsx")
df_fase2 = pd.read_excel("df_fase2.xlsx")

# 2. Simpan timestamp
timestamp_fase1 = df_fase1['Timestamp']
timestamp_fase2 = df_fase2['Timestamp']

# 3. Hapus kolom waktu
df_fase1_num = df_fase1.drop(columns=['Timestamp'])
df_fase2_num = df_fase2.drop(columns=['Timestamp'])

# 4. Min-Max Scaling
scaler = MinMaxScaler()
df_fase1_scaled = pd.DataFrame(scaler.fit_transform(df_fase1_num),
                                columns=df_fase1_num.columns)
df_fase2_scaled = pd.DataFrame(scaler.transform(df_fase2_num),
                                columns=df_fase2_num.columns)

# 5. Cari lag PACF signifikan
lag_max = 36
lag_sig_dict = {}
for col in df_fase1_scaled.columns:
    pacf_vals = pacf(df_fase1_scaled[col], nlags=lag_max, method='yw')
    n = len(df_fase1_scaled[col])
    conf = 1.96 / np.sqrt(n)
    signif_lags = [lag for lag in range(1, lag_max+1) if abs(pacf_vals[lag]) > conf]
    lag_sig_dict[col] = signif_lags[:5]

# 6. Bangun fitur lag
def build_lagged_features(df, lag_sig_dict):
    df_lag = df.copy()
    for col, lags in lag_sig_dict.items():
        for lag in lags:
            df_lag[f'{col}_lag{lag}'] = df_lag[col].shift(lag)
    return df_lag

df_fase1_lagged = build_lagged_features(df_fase1_scaled.copy(), lag_sig_dict)
df_fase2_lagged = build_lagged_features(df_fase2_scaled.copy(), lag_sig_dict)
```

```

# 7. Tambah timestamp dan drop NA
df_fase1_lagged['Timestamp'] = timestamp_fase1
df_fase2_lagged['Timestamp'] = timestamp_fase2
df_fase1_clean = df_fase1_lagged.dropna().reset_index(drop=True)
df_fase2_clean = df_fase2_lagged.dropna().reset_index(drop=True)

# 8. Split X-Y
target_cols = ['fgf', 'cp', 'gt']
X_fase1 = df_fase1_clean.drop(columns=target_cols + ['Timestamp'])
Y_fase1 = df_fase1_clean[target_cols]
X_fase2 = df_fase2_clean.drop(columns=target_cols + ['Timestamp'])
Y_fase2 = df_fase2_clean[target_cols]
timestamps1 = df_fase1_clean['Timestamp']
timestamps2 = df_fase2_clean['Timestamp']

# 9. Model awal
param_grid = {
    'estimator__n_estimators': [50, 100, 150],
    'estimator__max_depth': [3, 5, 10, None],
    'estimator__min_samples_split': [2, 4, 6],
    'estimator__min_samples_leaf': [1, 3, 5]
}
base_rf = RandomForestRegressor(random_state=42, bootstrap=False)
multi_rf = MultiOutputRegressor(base_rf)
grid = GridSearchCV(multi_rf, param_grid, cv=TimeSeriesSplit(n_splits=3),
                    scoring='neg_mean_squared_error', n_jobs=-1)
grid.fit(X_fase1, Y_fase1)
model_initial = grid.best_estimator_

# 10. Residual awal
Y_pred_initial = model_initial.predict(X_fase1)
residuals_initial = Y_fase1.values - Y_pred_initial

# 11. Tambah stacking lag residual (lag 1 dan 2)
X_stacked = X_fase1.copy()
for i, col in enumerate(target_cols):
    for lag in [1, 2]:
        resid_lagged = pd.Series(residuals_initial[:, i], index=X_fase1.index).shift(lag)
        X_stacked[f'{col}_resid_lag{lag}'] = resid_lagged
X_stacked = X_stacked.dropna()
Y_stacked = Y_fase1.loc[X_stacked.index]
timestamps_stacked = timestamps1.loc[X_stacked.index]

# 12. Model stacked
grid.fit(X_stacked, Y_stacked)
model_stacked = grid.best_estimator_

# 13. Residual fase 1
Y_pred_stacked_f1 = model_stacked.predict(X_stacked)

```

```

resid_f1 = Y_stacked.values - Y_pred_stacked_f1
resid_df_f1 = pd.DataFrame(resid_f1, columns=[f'{col}_resid' for col in target_cols])
resid_df_f1['Timestamp'] = timestamps_stacked.values

# 14. Prediksi fase 2
X_stacked_f2 = X_fase2.copy()
for i, col in enumerate(target_cols):
    for lag in [1, 2]:
        X_stacked_f2[f'{col}_resid_lag{lag}'] = 0 # Tidak tahu residual sebelumnya →
        asumsikan 0

Y_pred_stacked_f2 = model_stacked.predict(X_stacked_f2)
resid_f2 = Y_fase2.values - Y_pred_stacked_f2
resid_df_f2 = pd.DataFrame(resid_f2, columns=[f'{col}_resid' for col in target_cols])
resid_df_f2['Timestamp'] = timestamps2.values

# 15. Save
resid_df_f1.to_csv("resid_fase1_stacked.csv", index=False)
resid_df_f2.to_csv("resid_fase2_stacked.csv", index=False)

```

Lampiran 8 Syntax Uji MSE

```

from sklearn.metrics import mean_squared_error
for i, col in enumerate(target_cols):
    mse_col = mean_squared_error(Y_stacked.iloc[:, i], Y_pred_stacked_f1[:, i])
    print(f'MSE untuk {col}: {mse_col:.6f}')

```

Lampiran 9 Output Uji MSE

```

MSE untuk fgf: 0.004595
MSE untuk cp: 0.008962
MSE untuk gt: 0.001178

```

Lampiran 10 Data Residual Fase I

<i>Timestamp</i>	<i>Fuel Gas Flow</i>	<i>Combustion Pressure</i>	<i>Gas Temperature</i>
2022-07-01 00:15:00	-0.18939393939393917	-0.02166666666666686	0.027204195345787108
2022-07-01 00:20:00	-2,78E-01	-0.0111111111111110461	0.02359882005899805
2022-07-01 00:25:00	0.10454545454545544	-1,55E+01	0.0007374631268450837
2022-07-01 00:30:00	-0.04242424242424403	0.18541666666666617	0.02872594465514655
2022-07-01 00:35:00	1,05E+01	0.016666666666665886	0.050849838460456875
2022-07-01 00:40:00	0.015151515151514972	0.03333333333333349	7,77E-01
2022-07-01 00:45:00	-4,16E-01	-0.014285714285713999	-1,44E+01
2022-07-01 00:50:00	-0.00378787878787884	0.019047619047619216	0.0017305801376586993
2022-07-01 00:55:00	-0.00378787878787884	0.019047619047619216	-0.04648560675414393
2022-07-01 01:00:00	-0.06439393939393984	-0.014285714285713999	0.015117994100295906
2022-07-01 01:05:00	0.0021645021645025075	-0.025000000000000286	-4,44E-01
2022-07-01 01:10:00	2,22E-01	-0.025000000000000286	-5,55E-01
2022-07-01 01:15:00	0.07142857142857101	0.025362318840579642	-4,44E-01
2022-07-01 01:20:00	-0.0015151515151512474	0.025362318840579642	1,33E+01
2022-07-01 01:25:00	0.015151515151515083	-6,66E-01	-0.003933136676498639
2022-07-01 01:30:00	5,55E-01	-0.007971014492753573	-5,55E-01
2022-07-01 01:35:00	-0.0075757575757571805	1,22E+01	-0.027311565456209497
2022-07-01 01:40:00	-7,77E-01	-1,11E+01	-1,11E+01
2022-07-01 01:45:00	0.015151515151515083	0.05869565217391286	-7,77E-01
2022-07-01 01:50:00	-1,33E+01	1,22E+01	-1,33E+01
2022-07-01 01:55:00	5,27E-01	1,80E+00	2,11E+01
2022-07-01 02:00:00	00.00	0.008333333333333373	-0.0012192723697171948
2022-07-01 02:05:00	-0.03463203463203496	-0.014285714285713999	0.009561590885973192
2022-07-01 02:10:00	0.004559884559884819	1,80E+00	0.030210558437597346
2022-07-01 02:15:00	0.010822510822510012	-0.08333333333333337	0.007005899705015639
2022-07-01 02:20:00	-0.0023232323232326713	-0.007971014492753573	1,55E+01
2022-07-01 02:25:00	00.00	-0.04130434782608723	1,11E+00
2022-07-01 02:30:00	-3,05E+00	-0.008333333333333248	-2,22E-01
2022-07-01 02:35:00	-3,89E-01	0.02499999999999967	1,33E+01
2022-07-01 02:40:00	-0.0621212121212118	-0.007971014492753573	-0.001474926253687836
2022-07-01 02:45:00	0.07196969696969685	-0.04000000000000037	-0.025811209439526417
2022-07-01 02:50:00	3,89E-01	-0.014285714285713999	-4,44E-01
2022-07-01 02:55:00	1,11E+00	-0.04761904761904766	1,11E+01
2022-07-01 03:00:00	5,55E-01	2,64E-01	-1,33E+01
2022-07-01 03:05:00	0.0075757575757576245	-0.007971014492753573	1,33E+01
2022-07-01 03:10:00	-0.004329004329004682	0.019047619047619216	0.019886074661786157
2022-07-01 03:15:00	-0.0631313131313132	0.019047619047619216	0.012511443393349087
2022-07-01 03:20:00	5,55E-01	0.05869565217391286	3,33E+00
2022-07-01 03:25:00	5,27E-01	-0.007971014492753573	0.001966568338249597
2022-07-01 03:30:00	5,27E-01	-0.007971014492753573	0.009561590885973192
2022-07-01 03:35:00	2,22E-01	-0.007971014492753573	0.011036517139662028
2022-07-01 03:40:00	-0.019480519480520264	0.008333333333333373	0.021361000915473216
2022-07-01 03:45:00	-0.005050505050505305	0.11666666666666636	-0.007743362831856726
2022-07-05 13:50:00	00.00	0.025362318840579642	0.009561590885973192
2022-07-05 13:55:00	0.0075757575757576245	-0.07333333333333358	-0.003712745397212114
2022-07-05 14:00:00	0.0075757575757575135	0.04999999999999949	-0.006662597904586232
⋮	⋮	⋮	⋮
2022-07-06 14:15:00	-0.01515151515151636	0.46369047619047576	0.023725242309313388

Lampiran 11 Data Residual Fase II

<i>Timestamp</i>	<i>Fuel Gas Flow</i>	<i>Combustion Pressure</i>	<i>Gas Temperature</i>
2022-07-06 14:35:00	-0.09272727272727316	-0.36698412698412797	0.07267363904532054
2022-07-06 14:40:00	-0.11604683195592308	-0.22913650793650814	-0.0462143559488688
2022-07-06 14:45:00	-0.11604683195592308	-0.23784761904761892	-0.0427728613569322
2022-07-06 14:50:00	0.020316804407713607	-0.08730158730158777	0.0476843657817122
2022-07-06 14:55:00	0.05090909090909157	-0.036309523809524236	-0.012481563421829323
2022-07-06 15:00:00	0.10626262626262578	0.18666666666666676	0.08538348082595743
2022-07-06 15:05:00	-0.07204116638078956	0.19333333333333313	-0.09292035398230003
2022-07-06 15:10:00	-0.05522971652003961	-0.17499999999999905	0.011612447546636567
2022-07-06 15:15:00	-0.009775171065494637	-0.13365079365079457	-0.05506391347099204
2022-07-06 15:20:00	0.20234604105571818	0.2663492063492049	-0.013461298766047874
2022-07-06 15:25:00	0.030303030303031386	-0.10962962962963041	0.03993226264612734
2022-07-06 15:30:00	0.021505376344086558	0.13301587301587203	-0.09242871189773871
2022-07-06 15:35:00	-0.13000977517106527	-0.1297916666666672	0.004534032557631712
2022-07-06 15:40:00	0.0742115027829322	0.2663492063492049	0.054974524001073166
2022-07-06 15:45:00	0.015151515151516692	0.0056666666666654875	-0.010511446258673757
2022-07-06 15:50:00	-0.05425219941348958	-0.28	0.06879817299457675
2022-07-06 15:55:00	-0.07188910380399749	-0.200317460317461	0.07772414409582651
2022-07-06 16:00:00	0.0354683195592283	0.0636904761904763	0.10384402654867242
2022-07-06 16:05:00	0.049170954831332436	0.10909090909090852	-0.15531361589815357
2022-07-06 16:10:00	-0.015151515151516914	-0.24761904761904796	0.07060949769412869
2022-07-06 16:15:00	-0.03030303030303161	0.38000000000000009	-0.05796907124340689
2022-07-06 16:20:00	-0.1515151515151527	-0.31933333333333236	-0.022310856288171116
2022-07-06 16:25:00	-0.2591218305504013	-0.200317460317461	0.011276049100771679
2022-07-06 16:30:00	-0.1682127396413109	-0.13365079365079457	-0.00848082595870292
2022-07-06 16:35:00	-0.02492668621700933	-0.02063492063492134	0.028622787610619094
2022-07-06 16:40:00	-0.009775171065494637	-0.07619047619047736	0.08024520648967504
2022-07-06 16:45:00	0.020527859237536528	-0.04285714285714415	0.009448746312683776
2022-07-06 16:50:00	-0.05425219941348958	0.046666666666667855	0.04584913611462371
2022-07-06 16:55:00	-0.055440771349862084	-0.06964285714285745	0.03604351032448333
2022-07-06 17:00:00	0.0031769305962862093	-0.013333333333333863	0.0994653392330393
2022-07-06 17:05:00	0.018328445747800903	0.02000000000000024	-0.06588003933136799
2022-07-06 17:10:00	0.0031769305962862093	0.02000000000000024	-0.04052840836218963
2022-07-06 17:15:00	0.07893450635386146	0.08666666666666667	0.038993362831859224
2022-07-06 17:20:00	-0.026586620926243698	0.05333333333333455	0.0408515758422599
2022-07-06 17:25:00	-0.20840480274442535	-0.08000000000000003	0.01577782952957546
2022-07-06 17:30:00	0.10923753665689262	-0.26666666666666655	0.03161873156342215
2022-07-06 17:35:00	0.0946255002858774	0.28666666666666664	-0.03475294985250704
2022-07-06 17:40:00	-0.24242424242424154	-0.31999999999999993	0.0827222924568064
2022-07-06 17:45:00	-0.12878787878787817	-0.19999999999999907	-0.11734882005899655
2022-07-06 17:50:00	-0.04666666666666641	-0.3296666666666668	-0.09589970501474954
2022-07-06 17:55:00	0.3633333333333342	0.12500000000000078	0.017207472959686276
2022-07-06 18:00:00	0.21212121212121304	-0.03333333333333188	0.08567214496418052
2022-07-06 18:05:00	-0.23694638694638792	-0.10666666666666669	0.004424778761062287
2022-07-06 18:10:00	-0.28263403263403336	-0.2696428571428572	0.07729535398230003
2022-07-06 18:15:00	0.24766899766899736	0.36190476190476073	0.030058997050146974
⋮	⋮	⋮	⋮
2022-07-07 23:55:00	-0.05681818181818066	0.2533333333333332	0.016715830875121518

Lampiran 12 Uji MCC

```
data TAin;
  infile "/home/u64038375/residual_fase1_datalines.txt" dlm=' ' firstobs=1;
  input FGF CP GT;
run;

proc varmax data=TAin lagmax=36 printall;
  model FGF CP GT / p=1 minic=(p=12) print=(corry pcorr);
  output lead=12 out=hasil;
run;
```

Lampiran 13 Syntax Diagram Kontrol Max-MEWMA Fase I

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import chi2, norm

def max_mewma_stat(residual, lambda_, mu_0, Sigma_0):
    n, p = residual.shape
    Sigma_0_inv = np.linalg.inv(Sigma_0)
    Z = np.zeros((n, p))
    U = np.zeros(n)
    Y = np.zeros(n)
    V = np.zeros(n)
    M = np.zeros(n)
    for i in range(n):
        x_i = residual[i]
        if i == 0:
            Z[i] = lambda_ * (x_i - mu_0)
        else:
            Z[i] = (1 - lambda_) * Z[i-1] + lambda_ * (x_i - mu_0)
            Ti = (Z[i].T @ Sigma_0_inv @ Z[i]) * ((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1))))))
            Hi = chi2.cdf(Ti, df=p)
            U[i] = norm.ppf(Hi)
            phi_i = chi2.cdf((x_i - mu_0).T @ Sigma_0_inv @ (x_i - mu_0), df=p)
            if i == 0:
                Y[i] = lambda_ * phi_i
            else:
                Y[i] = (1 - lambda_) * Y[i-1] + lambda_ * phi_i
            V[i] = np.sqrt((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1)))))) * Y[i]
            M[i] = max(abs(U[i]), abs(V[i]))
    return M

def bootstrap_ucl_fase1(residual, lambda_, alpha=0.0027, N=500):
    n, p = residual.shape
    percentile_list = []
```

```

for l in range(N):
    idx_boot = np.random.choice(np.arange(n), size=n, replace=True)
    boot_data = residual[idx_boot, :]
    mu_0 = np.mean(boot_data, axis=0)
    Sigma_0 = np.cov(boot_data, rowvar=False)
    M = max_mewma_stat(boot_data, lambda_, mu_0, Sigma_0)
    perc = np.percentile(M, 100 * (1 - alpha))
    percentile_list.append(perc)
UCL_bootstrap = np.mean(percentile_list)
return UCL_bootstrap

# --- Main script untuk multi-lambda ---
lambda_list = np.arange(0.05, 0.91, 0.05)
alpha = 0.0027
N = 1000
residual_fase1 = resid_df_fase1_best.values # Ubah sesuai variabelmu

summary = []

for lambda_ in lambda_list:
    mu_0 = np.mean(residual_fase1, axis=0)
    Sigma_0 = np.cov(residual_fase1, rowvar=False)
    # Hitung Max-MEWMA
    M = max_mewma_stat(residual_fase1, lambda_, mu_0, Sigma_0)
    # Bootstrap UCL
    UCL_bootstrap = bootstrap_ucl_fase1(residual_fase1, lambda_, alpha=alpha, N=N)
    # Deteksi OOC
    idx_OOC = np.where(M > UCL_bootstrap)[0]
    ooc_count = len(idx_OOC)
    max_Mi = np.max(M)
    abs_diff = abs(UCL_bootstrap - max_Mi)
    # Simpan summary
    summary.append({
        "lambda": round(lambda_, 2),
        "UCL_bootstrap": UCL_bootstrap,
        "OOC_count": ooc_count,
        "Max_Mi": max_Mi,
        "Abs_UCL_minus_MaxMi": abs_diff
    })
    # Plot
    plt.figure(figsize=(12,5))
    plt.plot(M, label='Max-MEWMA', color='navy')
    plt.scatter(idx_OOC, M[idx_OOC], color='red', label='OOC', zorder=3)
    plt.axhline(UCL_bootstrap, color='orange', linestyle='--', label=f'UCL bootstrap
({UCL_bootstrap:.2f})')
    plt.title(f'Max-MEWMA Chart Fase I (lambda={lambda_:.2f})')
    plt.xlabel('Index Observasi')
    plt.ylabel('M Value')
    plt.legend()

```

```
plt.grid(True, alpha=0.4)
plt.tight_layout()
plt.show()
```

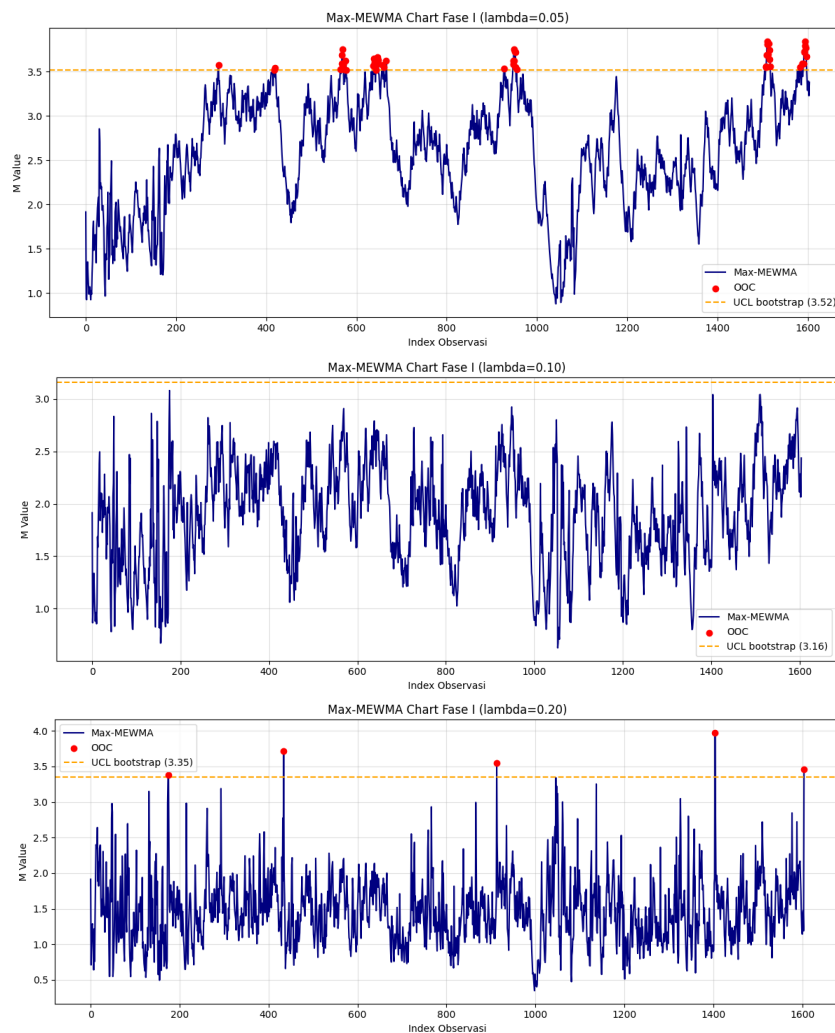
```
# Buat dataframe summary
```

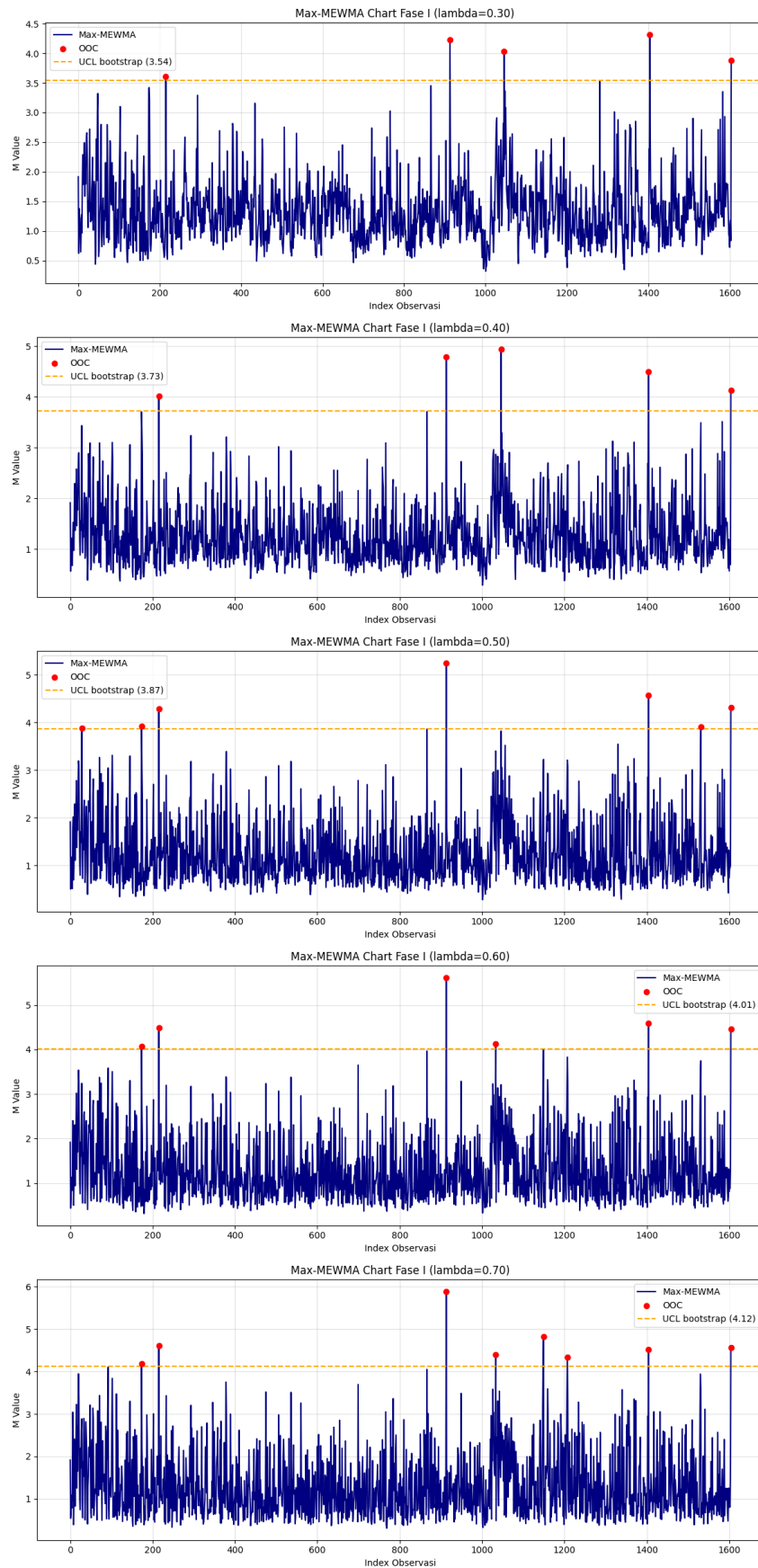
```
df_summary = pd.DataFrame(summary)
```

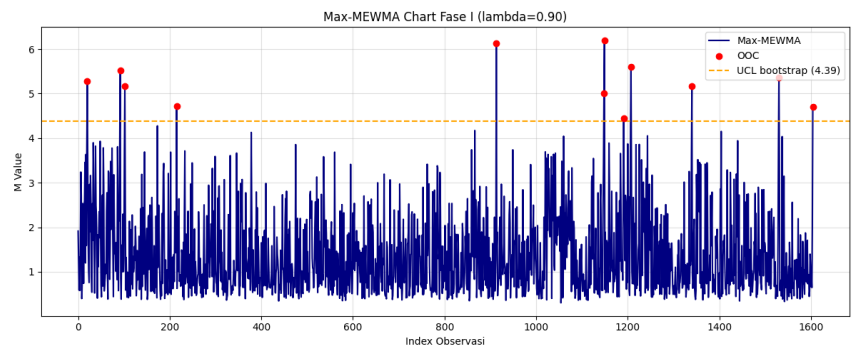
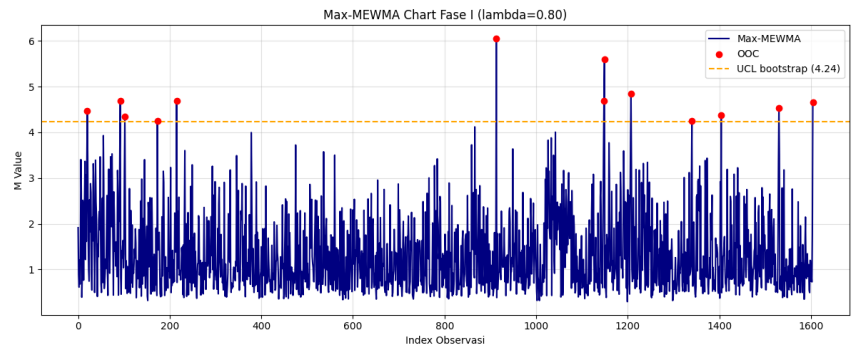
```
df_summary = df_summary.sort_values(['OOC_count', 'Abs_UCL_minus_MaxMi'],
ascending=[False, True])
```

```
print(df_summary)
```

Lampiran 14 Output Diagram Kontrol Max-MEWMA Fase I







Lampiran 15 Output Diagram Kontrol Max-MEWMA $\lambda = 0, 2$

Index	U_i	V_i	M_i	OOC
1	19.148.675.697.894.300	0.9722453027132107	19.148.675.697.894.300	Tidak
2	0.7102211861076202	0.6672102105381337	0.7102211861076202	Tidak
3	-12.756.811.150.027.500	0.8774323044003369	12.756.811.150.027.500	Tidak
4	0.6077425613799531	12.706.008.383.056.300	12.706.008.383.056.300	Tidak
5	10.296.086.290.590.500	12.892.691.576.550.100	12.892.691.576.550.100	Tidak
6	0.5084826005230338	1.017.610.262.964.130	1.017.610.262.964.130	Tidak
7	-0.1480413771741366	0.8041919171949936	0.8041919171949936	Tidak
8	-0.41305131078572194	0.6407198478834609	0.6407198478834609	Tidak
9	-1.218.187.645.581.730	0.7520418141803903	1.218.187.645.581.730	Tidak
10	-0.7583003925770417	0.737277274240352	0.7583003925770417	Tidak
11	-14.060.333.052.327.900	0.5928198547271519	14.060.333.052.327.900	Tidak
12	-19.331.344.555.676.300	0.47726968373350454	19.331.344.555.676.300	Tidak
13	-2.399.861.485.202.360	0.5199726398799611	2.399.861.485.202.360	Tidak
14	-2.365.500.009.291.070	0.41994686504010414	2.365.500.009.291.070	Tidak
15	-23.484.802.538.529.300	0.33876312352927984	23.484.802.538.529.300	Tidak
16	-26.421.375.253.316.400	0.27107065706526023	26.421.375.253.316.400	Tidak
17	-18.846.400.957.621.300	0.2866735222361498	18.846.400.957.621.300	Tidak
18	-21.539.120.844.707.200	0.22931787809068177	21.539.120.844.707.200	Tidak
19	-1.819.960.621.262.430	0.21769801026523095	1.819.960.621.262.430	Tidak
20	-2.094.324.826.824.060	0.17415189639337056	2.094.324.826.824.060	Tidak
21	-23.475.925.795.481.100	0.13931818337011323	23.475.925.795.481.100	Tidak
22	-23.929.852.013.661.900	0.11160893810159934	23.929.852.013.661.900	Tidak
23	-2.393.567.672.273.420	0.11662097660554055	2.393.567.672.273.420	Tidak
24	-16.417.363.899.489.100	0.1818178806527636	16.417.363.899.489.100	Tidak
25	-13.912.348.135.345.700	0.2759048669286572	13.912.348.135.345.700	Tidak
26	-16.346.018.802.956.600	0.22081944829193714	16.346.018.802.956.600	Tidak
27	-14.149.793.957.585.400	0.19238897605828953	14.149.793.957.585.400	Tidak
28	-16.094.508.756.186.900	0.15404873913486467	16.094.508.756.186.900	Tidak
29	-2.305.565.413.946.170	0.12687739281988497	2.305.565.413.946.170	Tidak
30	-16.798.010.457.470.000	0.2097498297678738	16.798.010.457.470.000	Tidak
31	-14.108.842.056.021.700	0.482972964511215	14.108.842.056.021.700	Tidak
32	-1.526.982.987.286.840	0.3870684719817846	1.526.982.987.286.840	Tidak
33	-11.706.969.202.057.500	0.333246512472039	11.706.969.202.057.500	Tidak
34	-15.031.287.541.894.900	0.26659719069939164	15.031.287.541.894.900	Tidak
35	-15.931.734.675.524.200	0.21414685476767126	15.931.734.675.524.200	Tidak
36	-2.158.685.408.353.090	0.20937161922714262	2.158.685.408.353.090	Tidak
37	-1.530.254.865.660.130	0.3378074081541635	1.530.254.865.660.130	Tidak
38	-13.760.270.832.134.000	0.3124742861274615	13.760.270.832.134.000	Tidak
39	-17.340.403.749.100.200	0.2501598251292984	17.340.403.749.100.200	Tidak
40	-1.800.595.183.806.510	0.20383891696360446	1.800.595.183.806.510	Tidak
41	-17.034.633.279.235.300	0.16857234252003905	17.034.633.279.235.300	Tidak
42	-10.514.389.834.607.300	0.18596917584516282	10.514.389.834.607.300	Tidak
43	-0.5488170139720218	0.38971883621563885	0.5488170139720218	Tidak
44	-0.5506473233187084	0.32273693152448807	0.5506473233187084	Tidak
⋮	⋮	⋮	⋮	⋮
1605	3.456.762.636.395.180	15.505.301.227.874.100	3.456.762.636.395.180	Ya

Lampiran 16 Syntax Pembersihan OOC Diagram Kontrol Fase I $\lambda = 0, 2$

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import chi2, norm

def max_mewma_stat_all(residual, lambda_, mu_0, Sigma_0):
    n, p = residual.shape
    Sigma_0_inv = np.linalg.inv(Sigma_0)
    Z = np.zeros((n, p))
    U = np.zeros(n)
    Y = np.zeros(n)
    V = np.zeros(n)
    M = np.zeros(n)
    for i in range(n):
        x_i = residual[i]
        if i == 0:
            Z[i] = lambda_ * (x_i - mu_0)
        else:
            Z[i] = (1 - lambda_) * Z[i-1] + lambda_ * (x_i - mu_0)
            Ti = (Z[i].T @ Sigma_0_inv @ Z[i]) * ((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1))))))
            Hi = chi2.cdf(Ti, df=p)
            U[i] = norm.ppf(Hi)
            phi_i = chi2.cdf((x_i - mu_0).T @ Sigma_0_inv @ (x_i - mu_0), df=p)
            if i == 0:
                Y[i] = lambda_ * phi_i
            else:
                Y[i] = (1 - lambda_) * Y[i-1] + lambda_ * phi_i
            V[i] = np.sqrt((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1)))))) * Y[i]
            M[i] = max(abs(U[i]), abs(V[i]))
    return U, V, M

# Parameter
lambda_ = 0.20
UCL = 3.354574
residual = resid_df_fase1_best.values.copy()
max_iter = 30

mask_valid = np.ones(residual.shape[0], dtype=bool) # Menyimpan baris yang BELUM pernah OOC

for it in range(max_iter):
    # Hitung mean dan cov HANYA dari baris yang belum pernah OOC
    mu_0 = np.mean(residual[mask_valid], axis=0)
    Sigma_0 = np.cov(residual[mask_valid], rowvar=False)
    U, V, M = max_mewma_stat_all(residual, lambda_, mu_0, Sigma_0)
```

```

idx_OOC = np.where((M > UCL) & mask_valid)[0]

plt.figure(figsize=(18,6))
plt.plot(M, color='navy', linewidth=1, label='Max-MEWMA', zorder=1)
plt.plot(M, linestyle='', marker='o', markersize=2, markerfacecolor='white',
markedgecolor='navy', markedgewidth=0.7, label='Titik Data', zorder=2)
if len(idx_OOC) > 0:
    plt.scatter(idx_OOC, M[idx_OOC], color='red', s=14, label='OOC', zorder=3)
plt.axhline(UCL, color='orange', linestyle='--', linewidth=2, label=f'UCL ( {UCL:.2f} )')
plt.title(f'Max-MEWMA Chart Fase I - Iterasi {it+1} (lambda={lambda_:.2f})')
plt.xlabel('Index Observasi')
plt.ylabel('M Value')
plt.legend()
plt.grid(True, alpha=0.4)
plt.tight_layout()
plt.show()

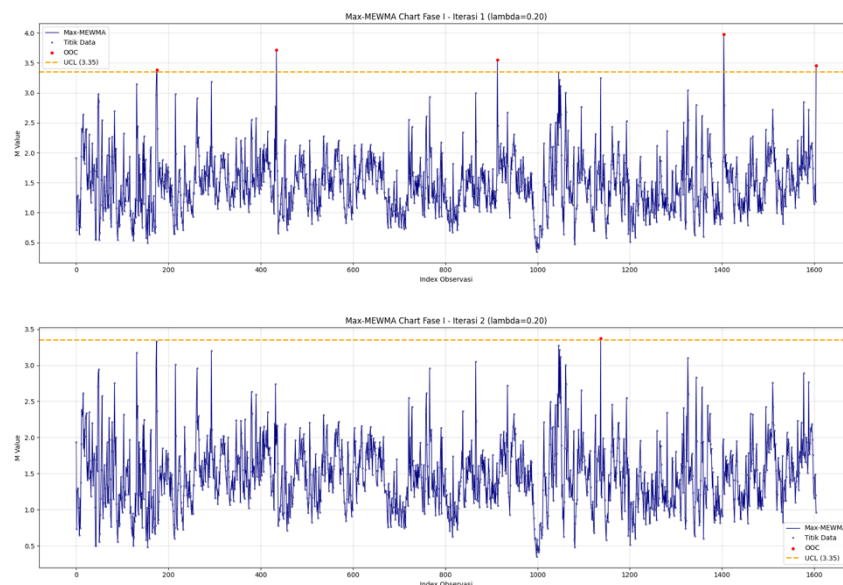
print(f'Iterasi {it+1} | Jumlah OOC: {len(idx_OOC)}')
if len(idx_OOC) == 0:
    print("Tidak ada OOC tersisa. Proses cleaning selesai.")
    break

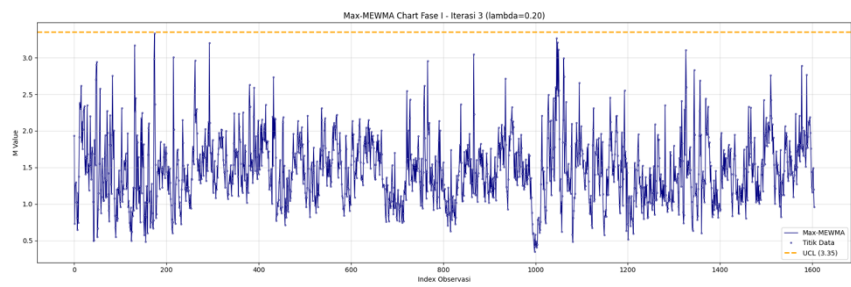
# Set OOC jadi 0, dan tandai sudah OOC (agar di-skip di iterasi berikut)
residual[idx_OOC, :] = 0
mask_valid[idx_OOC] = False

# Setelah cleaning selesai, simpan hasil
df_cleaned = pd.DataFrame(residual, columns=resid_df_fase1_best.columns)
df_cleaned.to_csv("Residual_Fase1_Cleaned0.csv", index=False)
print("Cleaning iteratif selesai. Residual yang OOC sudah diisi 0, dan file disimpan:
Residual_Fase1_Cleaned0.csv")

```

Lampiran 17 Output Pembersihan OOC Diagram Kontrol Fase I $\lambda = 0,2$





Lampiran 18 Syntax Diagram Kontrol Max-MEWMA Fase II

```

import numpy as np
import pandas as pd
from scipy.stats import chi2, norm
import matplotlib.pyplot as plt

# --- PARAMETER ---
lambda_ = 0.2
UCL = 3.354574

# Baca data residual fase 1
res_f1_clean = pd.read_csv("Residual_Fase1_Cleaned0.csv").values
mu_0 = np.mean(res_f1_clean, axis=0)
Sigma_0 = np.cov(res_f1_clean, rowvar=False)

# Baca data residual fase 2 & bersihkan NA
resid_df_fase2 = pd.read_csv("resid_fase2_stacked.csv")
resid_df_fase2 = resid_df_fase2[['fgf_resid', 'cp_resid',
'gt_resid']].dropna().reset_index(drop=True)
residual_fase2 = resid_df_fase2.values
n2, p2 = residual_fase2.shape

# Inisialisasi array
Sigma_0_inv = np.linalg.inv(Sigma_0)
Z = np.zeros((n2, p2))
U = np.zeros(n2)
Y = np.zeros(n2)
V = np.zeros(n2)
M = np.zeros(n2)

for i in range(n2):
    x_i = residual_fase2[i]
    if i == 0:
        Z[i] = lambda_ * (x_i - mu_0)
    else:
        Z[i] = (1 - lambda_) * Z[i-1] + lambda_ * (x_i - mu_0)
        Ti = (Z[i].T @ Sigma_0_inv @ Z[i]) * ((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1))))))
        Hi = chi2.cdf(Ti, df=p2)
        U[i] = norm.ppf(Hi)
        phi_i = chi2.cdf((x_i - mu_0).T @ Sigma_0_inv @ (x_i - mu_0), df=p2)
        if i == 0:
            Y[i] = lambda_ * phi_i
        else:
            Y[i] = (1 - lambda_) * Y[i-1] + lambda_ * phi_i
        V[i] = np.sqrt((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1)))))) * Y[i]
        M[i] = max(abs(U[i]), abs(V[i]))

```

```

valid_idx = (~np.isnan(M)) & (~np.isinf(M))
M = M[valid_idx]
U = U[valid_idx]
V = V[valid_idx]

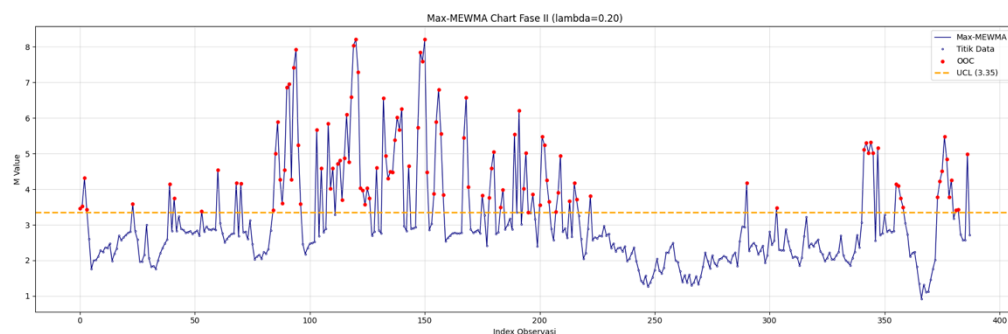
idx_OOC_f2 = np.where(M > UCL)[0]

plt.figure(figsize=(18,6))
plt.plot(np.arange(len(M)), M, color='navy', linewidth=1, label='Max-MEWMA', zorder=1)
plt.plot(np.arange(len(M)), M, linestyle='o', marker='o', markersize=2,
         markerfacecolor='white', markeredgewidth=0.7, label='Titik Data', zorder=2)
if len(idx_OOC_f2) > 0:
    plt.scatter(idx_OOC_f2, M[idx_OOC_f2], color='red', s=14, label='OOC', zorder=3)
plt.axhline(UCL, color='orange', linestyle='--', linewidth=2, label=f'UCL ( {UCL:.2f} )')
plt.title(f'Max-MEWMA Chart Fase II (lambda={lambda_:.2f})')
plt.xlabel('Index Observasi')
plt.ylabel('M Value')
plt.legend()
plt.grid(True, alpha=0.4)
plt.tight_layout()
plt.show()

# Simpan hasil Fase II
df_fase2_result = pd.DataFrame({
    "Index": np.arange(1, len(M)+1),
    "Ui": U,
    "Vi": V,
    "Mi": M,
    "OOC": ["Ya" if m > UCL else "Tidak" for m in M]
})
df_fase2_result.to_csv("MaxMEWMA_Fase2_Result_clean.csv", index=False)
print("Hasil deteksi OOC Fase II sudah disimpan: MaxMEWMA_Fase2_Result_clean.csv")
print(f"Jumlah OOC pada Fase II: {len(idx_OOC_f2)}")

```

Lampiran 19 Output Diagram Kontrol Max-MEWMA Fase II $\lambda = 0,2$



Lampiran 20 Ouput Diagram Kontrol Max-MEWMA Fase II

Index	U_i	V_i	M_i	OOC
1	34.589.607.333.198.800	0.9997288680231836	34.589.607.333.198.800	Ya
2	35.324.530.713.269.000	1.384.548.608.884.480	35.324.530.713.269.000	Ya
3	431.328.072.980.205	16.716.459.923.895.500	431.328.072.980.205	Ya
4	3.427.906.984.007.930	16.764.184.704.417.300	3.427.906.984.007.930	Ya
5	2.607.297.823.373.770	14.583.202.072.456.500	2.607.297.823.373.770	Tidak
6	10.753.338.929.574.500	17.593.334.100.751.300	17.593.334.100.751.300	Tidak
7	-14.471.771.035.144.100	20.017.610.810.916.700	20.017.610.810.916.700	Tidak
8	-0.4724300964989112	2.008.054.603.587.250	2.008.054.603.587.250	Tidak
9	0.7974605776234962	2.108.943.888.710.820	2.108.943.888.710.820	Tidak
10	0.07009490079920445	22.816.771.095.012.800	22.816.771.095.012.800	Tidak
11	0.024012905784285936	2.231.764.754.804.600	2.231.764.754.804.600	Tidak
12	0.7443363340784378	2.366.982.923.114.960	2.366.982.923.114.960	Tidak
13	-0.1525668544986735	2.349.644.747.434.700	2.349.644.747.434.700	Tidak
14	-0.21271481878560877	24.740.100.532.013.100	24.740.100.532.013.100	Tidak
15	-0.5586959340443453	19.868.571.874.503.300	19.868.571.874.503.300	Tidak
16	-0.32143463601163197	2.185.784.690.755.340	2.185.784.690.755.340	Tidak
17	17.106.073.038.369.800	233.273.085.556.346	233.273.085.556.346	Tidak
18	27.060.827.427.797.000	24.562.581.792.488.500	27.060.827.427.797.000	Tidak
19	-20.181.732.967.130.700	25.649.083.411.375.500	25.649.083.411.375.500	Tidak
20	1.030.469.798.028.280	26.459.615.939.854.500	26.459.615.939.854.500	Tidak
21	-0.4663971002800626	27.167.251.479.067.000	27.167.251.479.067.000	Tidak
22	0.2901005214317807	2.771.345.120.991.300	2.771.345.120.991.300	Tidak
23	2.685.258.886.496.120	2.816.131.241.551.340	2.816.131.241.551.340	Tidak
24	35.949.789.133.564.800	2.802.194.706.606.560	35.949.789.133.564.800	Ya
25	28.208.118.672.193.900	2.335.084.397.469.390	28.208.118.672.193.900	Tidak
26	2.591.991.043.449.480	2.401.540.448.504.120	2.591.991.043.449.480	Tidak
27	18.283.153.090.305.300	19.742.809.593.394.400	19.742.809.593.394.400	Tidak
28	19.086.343.838.349.300	19.719.788.847.830.000	19.719.788.847.830.000	Tidak
29	21.524.618.411.717.900	18.208.044.411.121.300	21.524.618.411.717.900	Tidak
30	30.073.523.073.229.800	20.353.424.090.981.400	30.073.523.073.229.800	Tidak
31	0.7878171262458474	20.581.231.324.752.000	20.581.231.324.752.000	Tidak
32	-0.5400451818105274	18.296.661.721.239.700	18.296.661.721.239.700	Tidak
33	-0.5041846879914809	18.435.363.793.499.600	18.435.363.793.499.600	Tidak
34	0.3983970091527843	176.879.484.934.644	176.879.484.934.644	Tidak
35	15.474.666.370.488.600	20.027.207.329.973.100	20.027.207.329.973.100	Tidak
36	13.249.430.186.318.400	22.018.046.438.557.200	22.018.046.438.557.200	Tidak
37	-0.8877137060282463	23.516.310.569.519.900	23.516.310.569.519.900	Tidak
38	19.047.696.217.918.700	24.812.758.763.473.700	24.812.758.763.473.700	Tidak
39	22.518.133.978.735.900	258.487.287.437.594	258.487.287.437.594	Tidak
40	4.145.345.083.417.660	26.678.684.996.652.300	4.145.345.083.417.660	Ya
41	28.246.823.673.773.100	2.734.293.950.018.890	28.246.823.673.773.100	Tidak
42	37.532.552.419.716.500	27.872.761.662.073.900	37.532.552.419.716.500	Ya
43	1.670.643.730.259.470	2.826.072.670.816.910	2.826.072.670.816.910	Tidak
44	3.243.127.708.424.320	28.608.406.490.161.600	3.243.127.708.424.320	Tidak
45	0.7456116481735429	28.886.387.707.464.900	28.886.387.707.464.900	Tidak
46	-0.5357770296453194	2.854.325.860.116.940	2.854.325.860.116.940	Tidak
47	0.8486220909490066	27.806.846.397.357.500	27.806.846.397.357.500	Tidak
⋮	⋮	⋮	⋮	⋮
388	2.505.377.645.969.130	27.165.606.997.825.500	27.165.606.997.825.500	Tidak

Lampiran 21 Syntax Kombinasi Karakteristik Kualitas

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
from scipy.stats import chi2, norm
from itertools import combinations

# --- PARAMETER ---
lambda_ = 0.2
UCL = 3.354574

# Data: Fase 2 (pastikan DataFrame dan tidak ada NA)
df_f2 = resid_df_fase2_best.copy()
df_f2 = df_f2[['fgf_resid', 'cp_resid', 'gt_resid']].dropna().reset_index(drop=True)
var_list = df_f2.columns.tolist()

label_map = {
    'fgf_resid': 'Fuel Gas Flow',
    'cp_resid': 'Combustion Pressure',
    'gt_resid': 'Gas Temperature'
}

# Mean & covariance baseline dari Fase 1
df_f1_clean = pd.read_csv("Residual_Fase1_Cleaned0.csv")
df_f1_clean = df_f1_clean[['fgf_resid', 'cp_resid', 'gt_resid']].dropna().reset_index(drop=True)

summary = []

for pair in combinations(var_list, 2): # Semua kombinasi 2 variabel
    v1, v2 = pair
    # Subset data, pastikan bersih dan index berurutan
    X_f1 = df_f1_clean[[v1, v2]].dropna().reset_index(drop=True).values
    X_f2 = df_f2[[v1, v2]].dropna().reset_index(drop=True).values

    n2, p2 = X_f2.shape
    mu_0 = np.mean(X_f1, axis=0)
    Sigma_0 = np.cov(X_f1, rowvar=False)
    Sigma_0_inv = np.linalg.inv(Sigma_0)

    Z = np.zeros((n2, p2))
    U = np.zeros(n2)
    Y = np.zeros(n2)
    V = np.zeros(n2)
    M = np.zeros(n2)

    for i in range(n2):
        x_i = X_f2[i]
        if i == 0:
```

```

        Z[i] = lambda_ * (x_i - mu_0)
    else:
        Z[i] = (1 - lambda_) * Z[i-1] + lambda_ * (x_i - mu_0)
        Ti = (Z[i].T @ Sigma_0_inv @ Z[i]) * ((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1))))))
        Hi = chi2.cdf(Ti, df=p2)
        U[i] = norm.ppf(Hi)
        phi_i = chi2.cdf((x_i - mu_0).T @ Sigma_0_inv @ (x_i - mu_0), df=p2)
        if i == 0:
            Y[i] = lambda_ * phi_i
        else:
            Y[i] = (1 - lambda_) * Y[i-1] + lambda_ * phi_i
        V[i] = np.sqrt((2 - lambda_) / (lambda_ * (1 - (1 - lambda_)**(2 * (i+1)))))) * Y[i]
        M[i] = max(abs(U[i]), abs(V[i]))

valid_idx = (~np.isnan(M)) & (~np.isinf(M))
M_plot = M[valid_idx]
U_plot = U[valid_idx]
V_plot = V[valid_idx]

idx_OOC = np.where(M_plot > UCL)[0]
ooc_count = len(idx_OOC)
max_m = np.nanmax(M_plot) if len(M_plot) > 0 else np.nan

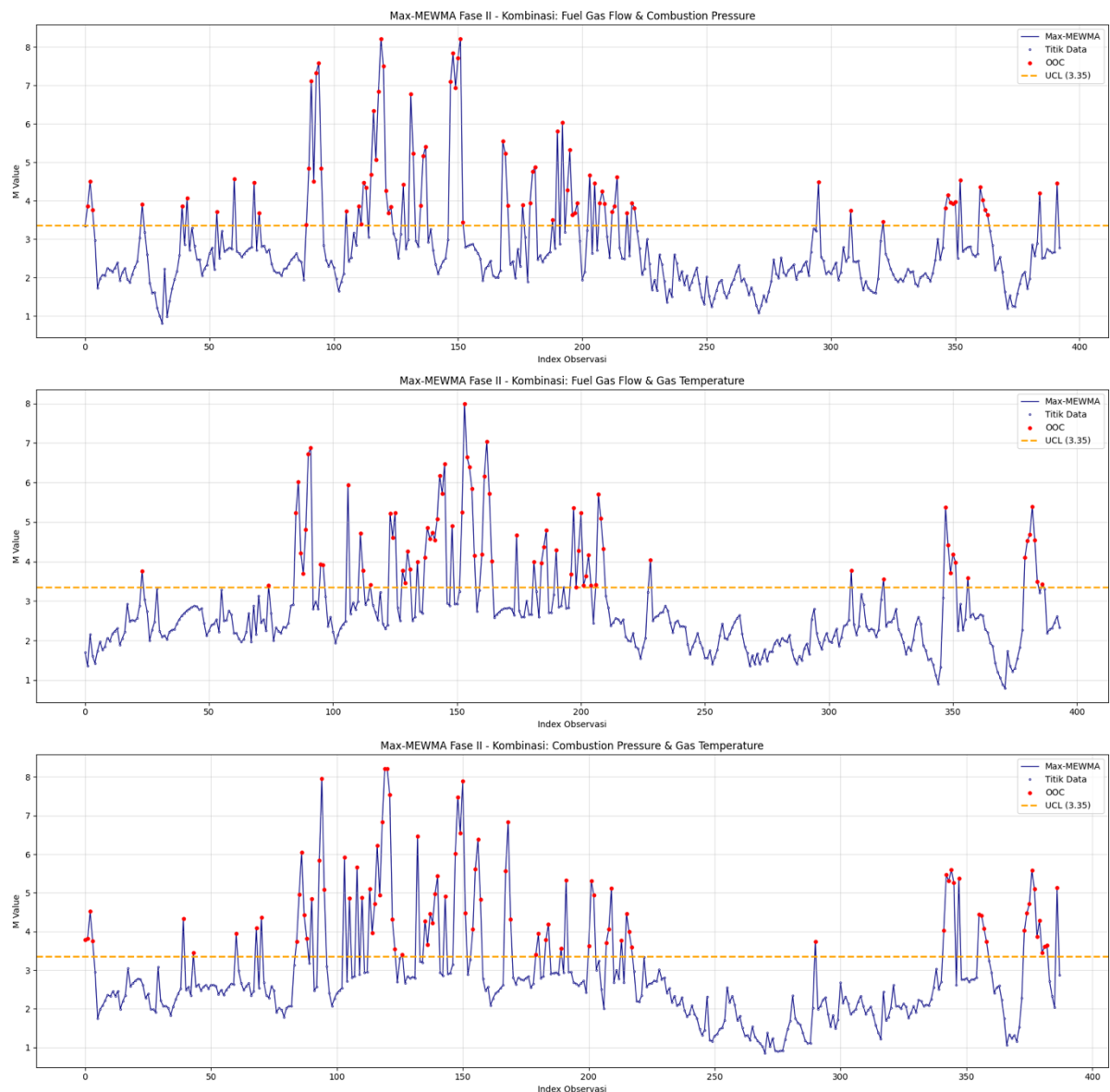
summary.append({
    "var1": v1,
    "var2": v2,
    "OOC_count": ooc_count,
    "Max_Mi": max_m,
    "Abs_UCL_minus_MaxMi": abs(UCL - max_m) if np.isfinite(max_m) else np.nan
})

# Plot dengan index urut agar tidak terputus
plt.figure(figsize=(18,6))
plt.plot(np.arange(len(M_plot)), M_plot, color='navy', linewidth=1, label='Max-MEWMA', zorder=1)
plt.plot(np.arange(len(M_plot)), M_plot, linestyle="", marker='o', markersize=2,
         markerfacecolor='white', markeredgewidth=0.7,
         label='Titik Data', zorder=2)
if len(idx_OOC) > 0:
    plt.scatter(idx_OOC, M_plot[idx_OOC], color='red', s=14, label='OOC', zorder=3)
plt.axhline(UCL, color='orange', linestyle='--', linewidth=2, label=f'UCL ({UCL:.2f})')
plt.title(f'Max-MEWMA Fase II - Kombinasi: {label_map[v1]} & {label_map[v2]}')
plt.xlabel('Index Observasi')
plt.ylabel('M Value')
plt.legend()
plt.grid(True, alpha=0.4)
plt.tight_layout()
plt.show()

```

```
# Rekap hasil kombinasi
df_pairwise = pd.DataFrame(summary)
df_pairwise = df_pairwise.sort_values(['OOC_count', 'Abs_UCL_minus_MaxMi'],
ascending=[False, True])
print(df_pairwise)
df_pairwise.to_excel("Summary_Pairwise_MaxMEWMA_Fase2.xlsx", index=False)
```

Lampiran 22 Output Kombinasi Karakteristik Kualitas



BIODATA PENULIS



Penulis dilahirkan Nama lengkap penulis adalah Husnan Ali Husnain, lahir pada tanggal 5 Mei 2003 di Pontianak. Penulis merupakan anak ketiga dari empat bersaudara, putra dari pasangan Sukmawati dan Anung Triwibowo. Riwayat pendidikan penulis dimulai dari SD Margahayu Raya Blok I/II, dilanjutkan ke SMP 45 Bandung, dan kemudian ke SMA 27 Bandung. Setelah menyelesaikan pendidikan menengah, penulis melanjutkan studi di Institut Teknologi Sepuluh Nopember, pada Fakultas Sains dan Analitika Data, Program Studi Statistika. Selama masa kuliah, penulis aktif dalam berbagai kegiatan organisasi. Penulis pernah menjabat sebagai Staf Departemen Minat dan Bakat di organisasi HIMASTA ITS Pijar Raya serta dipercaya menjadi Kepala Departemen Minat dan Bakat di HIMASTA ITS Amritaya. Penulis juga berperan sebagai Wakil Kepala Divisi Akomodasi dan Transportasi pada Pekan Raya Statistika 2023. Selain itu, penulis juga mengikuti berbagai pelatihan dan kegiatan pengembangan diri, seperti LKMM-Pra TD, Basic Media Schooling, dan Leadership Organization Training 10 (LOT 10.0).

Penulis dapat dihubungi melalui email: husnanalih@gmail.com.