

TUGAS BESAR MATEMATIKA TEKNIK LANJUT
ANALISIS TREN DAN PREDIKSI PRODUKSI
LISTRIK MENGGUNAKAN MODEL ARIMA

Ditujukan untuk memenuhi salah satu syarat kelulusan Mata Kuliah

Matematika Teknik Lanjut



Disusun oleh kelompok 04 :

Azmi Taqiuddin Syah	1103213078
Geovani Prihanda Mahkdagasta	1101213102
Fadhil Dzikri Aqila	1103213136

PROGRAM STUDI S2 TEKNIK ELEKTRO
FAKULTAS TEKNIK ELEKTRO
UNIVERSITAS TELKOM




2025

KATA PENGANTAR

Puji syukur kami panjatkan ke hadirat Tuhan Yang Maha Esa atas segala limpahan rahmat dan karunia-Nya sehingga laporan ini dapat diselesaikan dengan baik. Laporan ini disusun sebagai bagian dari upaya kami untuk menganalisis dan memprediksi produksi listrik menggunakan metode deret waktu dan model ARIMA. Dalam laporan ini, kami memaparkan langkah-langkah mulai dari pemahaman pola data, transformasi stasioneritas, pemodelan, hingga evaluasi performa prediksi.

Penelitian ini menggunakan model ARIMA (Autoregressive Integrated Moving Average), yang terbukti efektif dalam menganalisis data deret waktu, meskipun terdapat tantangan seperti ketidakstasioneran data awal. Melalui laporan ini, kami berharap dapat memberikan kontribusi dalam pengembangan metodologi analisis data deret waktu, yang dapat digunakan untuk berbagai kebutuhan, seperti perencanaan energi, prediksi konsumsi, dan pengelolaan sumber daya listrik.

Kami menyadari bahwa laporan ini masih memiliki keterbatasan, sehingga kami mengharapkan saran dan kritik yang membangun untuk penyempurnaan lebih lanjut. Akhir kata, kami mengucapkan terima kasih kepada semua pihak yang telah memberikan dukungan, baik secara langsung maupun tidak langsung, dalam penyelesaian laporan ini.

Bandung, 1 Januari 2025  Penulis, Azmi Taquiuddin Syah	Bandung, 1 Januari 2025  Penulis, Geovani Prihanda Mahkdagasta	Bandung, 1 Januari 2025  Penulis, Fadhil Dzikri Aqila
--	--	---

DAFTAR ISI

KATA PENGANTAR	1
DAFTAR ISI	1
BAB I PENDAHULUAN.....	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah.....	2
1.3 Tujuan Penelitian	3
BAB II TINJAUAN PUSTAKA.....	4
2.1 Deret Waktu (Time Series).....	4
2.2 Stasioneritas dalam Deret Waktu	4
2.3 Model ARIMA	5
2.4 Kriteria Pemilihan Model	5
2.4.1 Akaike Information Criterion (AIC)	5
2.4.2 Bayesian Information Criterion (BIC):.....	5
2.4.3 Log-Likelihood	6
2.5 Program ARIMA pada Produksi Listrik	6
2.6 Uji Statistik	6
BAB III METODOLOGI PENELITIAN	7
3.1 Desain Penelitian	7
3.2 Pengelolaan Data	7
3.2.1 Dataset.....	7
3.2.2 Pembagian Data Latih dan Data Test.....	8
3.3 Uji Stasioneritas Menggunakan Metode Differencing	9
3.3.1 ACF	10
3.3.2 PACF.....	11
3.4 Identifikasi Parameter Model ARIMA.....	12

3.4.2 Analisis Model ARIMA(1, 1, 1)	13
3.4.2 Analisis Model ARIMA(1, 1, 0)	13
3.4.3 Analisis Model ARIMA(1, 1, 2)	14
3.4.4 Perbandingan Model ARIMA	14
3.5 Pembangunan Model ARIMA	16
3.5.1 Transformasi Data	16
3.5.2 Analisis Model ARIMA(8, 1, 2) dengan Log-Transformasi	17
3.6 Evaluasi Model	19
3.6.1 Variasi Residual	19
3.6.2 Kernel Density Estimation (KDE)	20
3.6 Prediksi Model	22
BAB IV PENUTUP	23
4.1 Kesimpulan	23
4.2 Saran	23

BAB I

PENDAHULUAN

1.1 Latar Belakang

Produksi listrik merupakan salah satu indikator penting dalam mengukur perkembangan dan kebutuhan energi suatu negara. Seiring dengan meningkatnya permintaan energi, analisis data produksi listrik menjadi sangat penting untuk memastikan ketersediaan energi yang cukup dan pengelolaan sumber daya yang efisien. Data deret waktu produksi listrik, yang mencatat jumlah energi yang dihasilkan secara berkala, dapat memberikan wawasan mendalam mengenai pola, tren, dan fluktuasi yang terjadi selama periode tertentu.

Dataset yang digunakan dalam analisis ini berisi data produksi listrik dengan kode variabel IPG2211A2N. Data ini terdiri dari 397 entri, mencakup rentang waktu dari tahun 1985 hingga beberapa tahun ke depan dalam bentuk tahunan. Data ini memberikan gambaran yang mendalam tentang produksi listrik selama lebih dari tiga dekade, yang dapat digunakan untuk menganalisis tren historis dan membuat proyeksi jangka panjang. Namun, data deret waktu sering kali memiliki tantangan seperti ketidakstasioneran, faktor musiman, dan tren jangka panjang yang perlu ditangani sebelum proses analisis lebih lanjut. Ketidakstasioneran dalam data deret waktu dapat menyebabkan kesulitan dalam memprediksi nilai masa depan, karena pola dalam data terus berubah seiring waktu. Selain itu, faktor musiman seperti perubahan pola produksi berdasarkan musim atau kebijakan tertentu juga dapat mempengaruhi kestabilan data. Oleh karena itu, pendekatan yang tepat diperlukan untuk menangani masalah tersebut.

Dengan menggunakan model Autoregressive Integrated Moving Average (ARIMA), pendekatan ini mampu menangkap pola yang kompleks dan memberikan prediksi yang akurat untuk produksi listrik di masa depan. Model ARIMA merupakan salah satu metode statistik yang sangat efektif dalam memodelkan dan meramalkan data deret waktu, dengan mempertimbangkan elemen autoregresif, perbedaan, dan rata-rata bergerak dalam data. ARIMA memungkinkan untuk mereduksi dampak ketidakstasioneran dan memproyeksikan produksi listrik secara lebih tepat, bahkan dengan adanya fluktuasi musiman dan tren jangka panjang.

Selain itu, pemahaman lebih lanjut terhadap tren produksi listrik sangat penting untuk merencanakan kebijakan energi yang efisien, mengantisipasi kebutuhan pasokan, dan memastikan keberlanjutan energi di masa depan. Oleh karena itu, analisis ini bertujuan untuk memberikan gambaran yang lebih jelas dan detail tentang pola produksi listrik dengan memanfaatkan data deret waktu serta metode prediksi yang tepat.

1.2 Rumusan Masalah

Dalam penelitian ini, terdapat beberapa masalah utama yang ingin diselesaikan:

1. Bagaimana mengidentifikasi pola dan tren pada data produksi listrik dari tahun 1985 hingga 2018?
2. Apakah data produksi listrik bersifat stasioner, dan bagaimana cara mentransformasikannya jika tidak stasioner?
3. Bagaimana membangun model ARIMA yang optimal untuk memprediksi produksi listrik di masa mendatang?

1.3 Tujuan Penelitian

Penelitian ini bertujuan untuk:

1. Menganalisis pola dan karakteristik data produksi listrik dalam rentang waktu tertentu.
2. Mengidentifikasi dan mentransformasi data menjadi stasioner untuk memenuhi persyaratan analisis deret waktu.
3. Membuat model ARIMA yang optimal untuk prediksi produksi listrik dan mengevaluasi kinerjanya.

BAB II TINJAUAN PUSTAKA

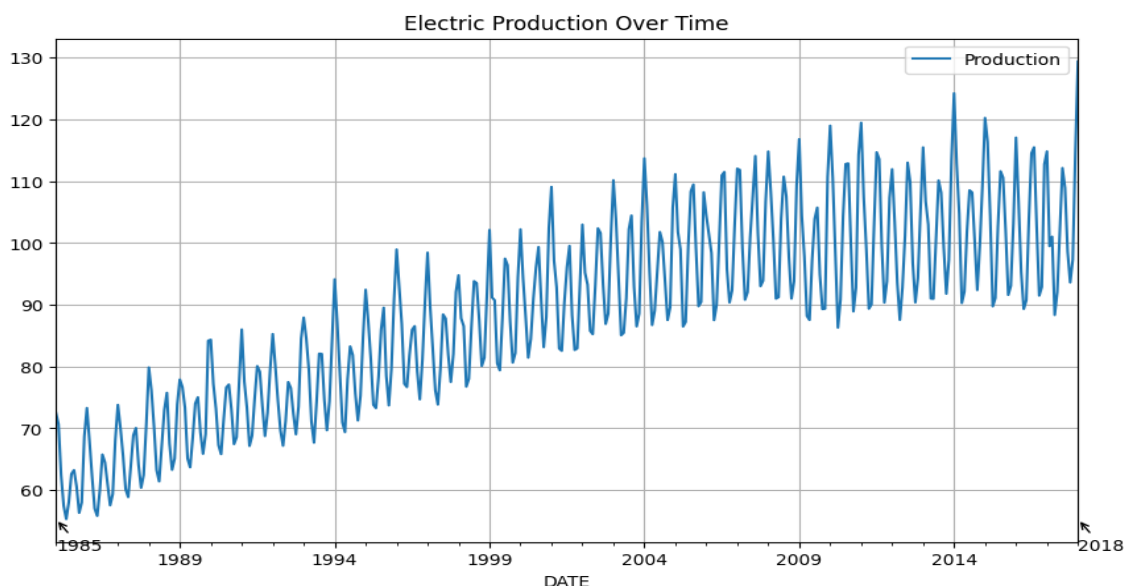
2.1 Deret Waktu (Time Series)

Deret waktu merupakan sekumpulan data yang diambil secara berkala pada interval waktu tertentu, seperti harian, bulanan, atau tahunan. Analisis deret waktu bertujuan untuk memahami pola data historis dan membuat prediksi untuk periode waktu mendatang. Dalam analisis ini, terdapat beberapa komponen utama deret waktu, yaitu:

1. **Tren:** Pola jangka panjang yang mencerminkan arah kenaikan atau penurunan data.
2. **Musiman (Seasonality):** Pola berulang dalam data yang terjadi secara periodik.
3. **Residual:** Fluktuasi acak yang tidak dapat dijelaskan oleh tren atau musiman.

2.2 Stasioneritas dalam Deret Waktu

Stasioneritas merupakan sifat penting dalam deret waktu, di mana rata-rata, variansi, dan autokorelasi data tidak berubah seiring waktu. Data yang tidak stasioner dapat menyebabkan kesalahan interpretasi dan prediksi yang tidak akurat. Oleh karena itu, transformasi seperti differencing sering digunakan untuk membuat data menjadi stasioner. Metode Differencing adalah teknik yang digunakan untuk menghilangkan tren dalam data deret waktu guna membuatnya stasioner. Differencing dilakukan dengan mengurangi nilai data saat ini dengan nilai data pada lag sebelumnya. Dengan cara ini, komponen tren dapat dihilangkan, sehingga data menjadi lebih stabil secara statistik.



Pada plot grafik di atas, terlihat bahwa dataset yang digunakan tidak stasioner, yang ditunjukkan oleh adanya tren kenaikan grafik yang signifikan seiring waktu. Hal ini mengindikasikan bahwa data memiliki tren atau pola jangka panjang, sehingga perlu dilakukan transformasi, seperti differencing, untuk menjadikan data stasioner sebelum membangun model ARIMA.

2.3 Model ARIMA

Model *Autoregressive Integrated Moving Average* (ARIMA) adalah metode yang banyak digunakan untuk analisis dan peramalan deret waktu. Model ini terdiri dari tiga komponen:

1. Autoregressive (AR): Ketergantungan nilai saat ini pada nilai sebelumnya.
2. Integrated (I): Differencing data untuk mencapai stasioneritas.
3. Moving Average (MA): Ketergantungan nilai saat ini pada kesalahan sebelumnya.

Model ARIMA sering ditulis dalam bentuk $ARIMA(p,d,q)$, di mana:

- **p**: Jumlah lag untuk komponen AR.
- **d**: Tingkat differencing untuk membuat data stasioner.
- **q**: Jumlah lag untuk komponen MA.

2.4 Kriteria Pemilihan Model

Dalam memilih model ARIMA yang optimal, beberapa kriteria evaluasi digunakan:

2.4.1 Akaike Information Criterion (AIC)

ACF menunjukkan hubungan antara data pada lag tertentu dengan nilai data sebelumnya. Pada model ARIMA, ACF digunakan untuk mengidentifikasi parameter q untuk Moving Average (MA). Jika ACF menunjukkan cutoff tiba-tiba setelah lag tertentu, maka model MA adalah yang dominan. Biasanya, nilai signifikan pada ACF hanya muncul pada beberapa lag pertama, yang menunjukkan bahwa parameter q (jumlah lag yang digunakan dalam model MA) harus ditentukan berdasarkan jumlah lag yang signifikan.

2.4.2 Bayesian Information Criterion (BIC):

PACF menunjukkan hubungan antara data pada lag tertentu setelah menghilangkan pengaruh lag-lag sebelumnya. PACF digunakan untuk mengidentifikasi parameter p untuk Autoregressive (AR). PACF menunjukkan cutoff setelah lag tertentu, maka model AR adalah yang dominan. Biasanya, nilai signifikan pada PACF hanya muncul pada beberapa lag pertama,

yang menunjukkan bahwa parameter p (jumlah lag yang digunakan dalam model AR) harus ditentukan berdasarkan jumlah lag yang signifikan.

2.4.3 Log-Likelihood

Log-Likelihood (LL) mengukur sejauh mana data yang diamati mendekati model yang digunakan. Ini adalah logaritma dari fungsi likelihood, yang berfungsi untuk mengukur seberapa besar kemungkinan data yang diamati muncul berdasarkan parameter model yang ada. Dalam konteks model ARIMA, Log-Likelihood mengukur kesesuaian antara model ARIMA yang dibangun dengan data yang ada.

2.5 Program ARIMA pada Produksi Listrik

Model ARIMA telah digunakan dalam berbagai penelitian untuk menganalisis dan memprediksi produksi listrik. Metode ini dapat menangkap pola tren dan musiman yang terdapat dalam data deret waktu produksi listrik. Prediksi yang dihasilkan dapat digunakan untuk mendukung perencanaan kapasitas listrik dan pengambilan keputusan strategis.

2.6 Uji Statistik

Untuk memastikan validitas model dan stasioneritas data, beberapa uji statistik digunakan:

1. **Augmented Dickey-Fuller (ADF) Test:** Menguji stasioneritas data dengan hipotesis nol bahwa data memiliki akar unit (tidak stasioner).
2. **Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF):** Membantu dalam menentukan nilai p dan q pada model ARIMA.

BAB III

METODOLOGI PENELITIAN

3.1 Desain Penelitian

Penelitian ini bertujuan untuk memodelkan data produksi listrik menggunakan analisis deret waktu dan memprediksi produksinya di masa depan. Langkah-langkah penelitian terdiri dari:

1. Eksplorasi dataset.
2. Transformasi data menjadi stasioner.
3. Pemilihan parameter p , d , q menggunakan ACF dan PACF.
4. Pembangunan dan evaluasi model ARIMA.
5. Prediksi data untuk beberapa tahun ke depan.

3.2 Pengelolaan Data

Pengelolaan data merupakan salah satu aspek penting dalam setiap sistem informasi dan penelitian. Proses ini mencakup pengumpulan, penyimpanan, pengolahan, dan analisis data yang diperoleh untuk menghasilkan informasi yang dapat digunakan untuk pengambilan keputusan. Dalam pengelolaan data yang efektif, penting untuk memastikan bahwa data yang digunakan adalah data yang valid, akurat, dan dapat dipercaya.

3.2.1 Dataset

Dataset merupakan kumpulan data yang digunakan sebagai bahan untuk analisis dalam suatu penelitian atau sistem. Data yang terkumpul harus memenuhi kriteria validitas, akurasi, dan kepercayaan agar dapat digunakan untuk tujuan yang diinginkan. Dalam tahap pengelolaan data, pemilihan dataset yang tepat menjadi langkah pertama yang krusial, karena dataset yang tidak sesuai atau berkualitas rendah dapat mengarah pada kesimpulan yang tidak tepat.

1	DATE	IPG2211A2N		
2	1/1/1985	72.5052		
3	2/1/1985	70.672		
4	3/1/1985	62.4502		
5	4/1/1985	57.4714		
6	5/1/1985	55.3151		
7	6/1/1985	58.0904		
8	7/1/1985	62.6202		

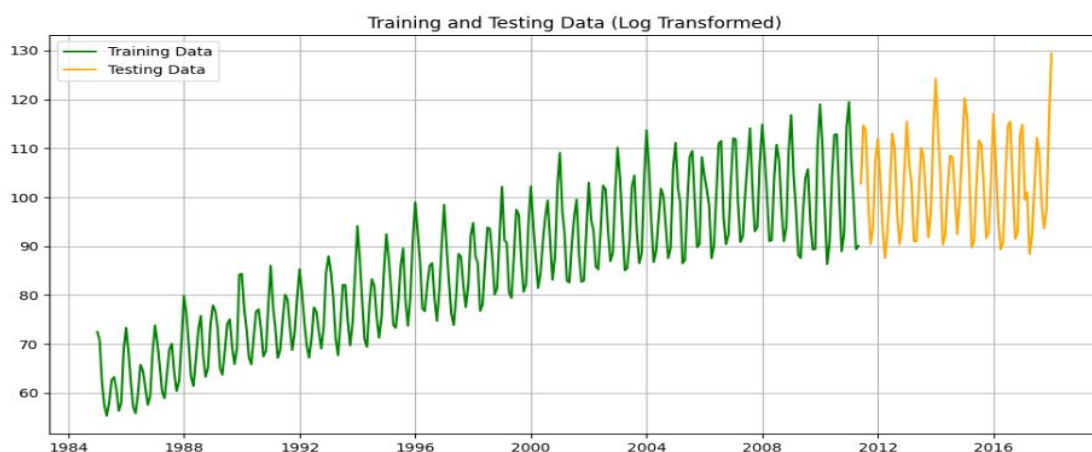
Dataset yang digunakan terdiri dari 397 entri produksi listrik bulanan dari Januari 1985. Variabel kunci:

- DATE: Waktu pengamatan (tahunan).
- IPG2211A2N: Produksi listrik pada Tahun tersebut.

3.2.2 Pembagian Data Latih dan Data Test

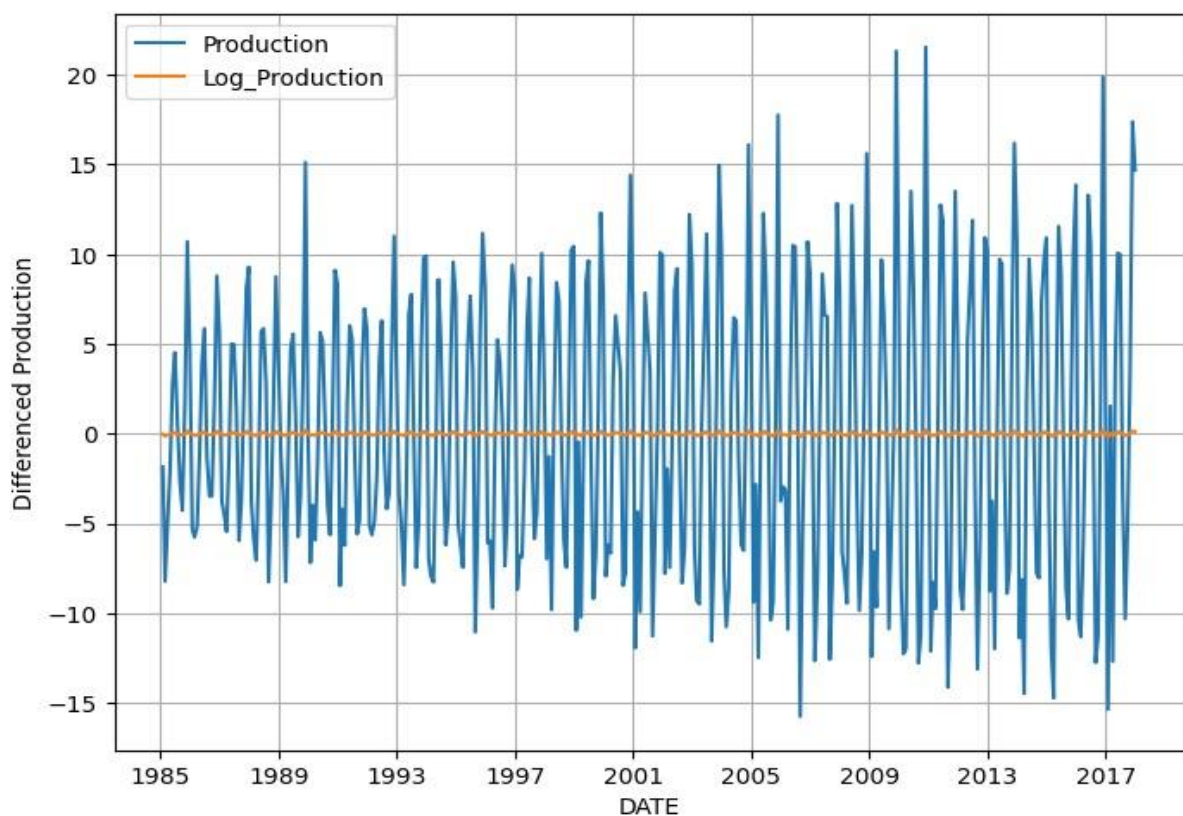
Pembagian data latih (training data) dan data uji (testing data) merupakan langkah penting dalam proses pengembangan model machine learning. Tujuan dari pembagian ini adalah untuk memastikan bahwa model yang dibangun dapat mempelajari pola dari data latih dan kemudian diuji kemampuannya pada data yang belum pernah dilihat sebelumnya, yakni data uji

Dalam tugas ini kami melakukan pembagian sebesar 80% data untuk training dan 20% untuk testing, meskipun rasio ini dapat disesuaikan tergantung pada ukuran dataset dan kebutuhan analisis. Pembagian yang seimbang antara data latih dan data uji sangat penting untuk menghindari bias dalam evaluasi model dan untuk memastikan bahwa model dapat menggeneralisasi dengan baik pada data baru.



3.3 Uji Stasioneritas Menggunakan Metode Differencing

Uji stasioneritas sangat penting dalam analisis deret waktu karena sebagian besar model deret waktu, seperti ARIMA (AutoRegressive Integrated Moving Average), mengasumsikan bahwa data yang digunakan bersifat stasioner. Data stasioner adalah data yang statistiknya (seperti rata-rata dan varians) tidak berubah seiring waktu, dan pola-pola dalam data cenderung konsisten. Dari dataset yang dianalisis, terlihat pola pada grafik yang tidak stasioner, sehingga dilakukan proses differencing untuk membuat pola dalam dataset menjadi stasioner.

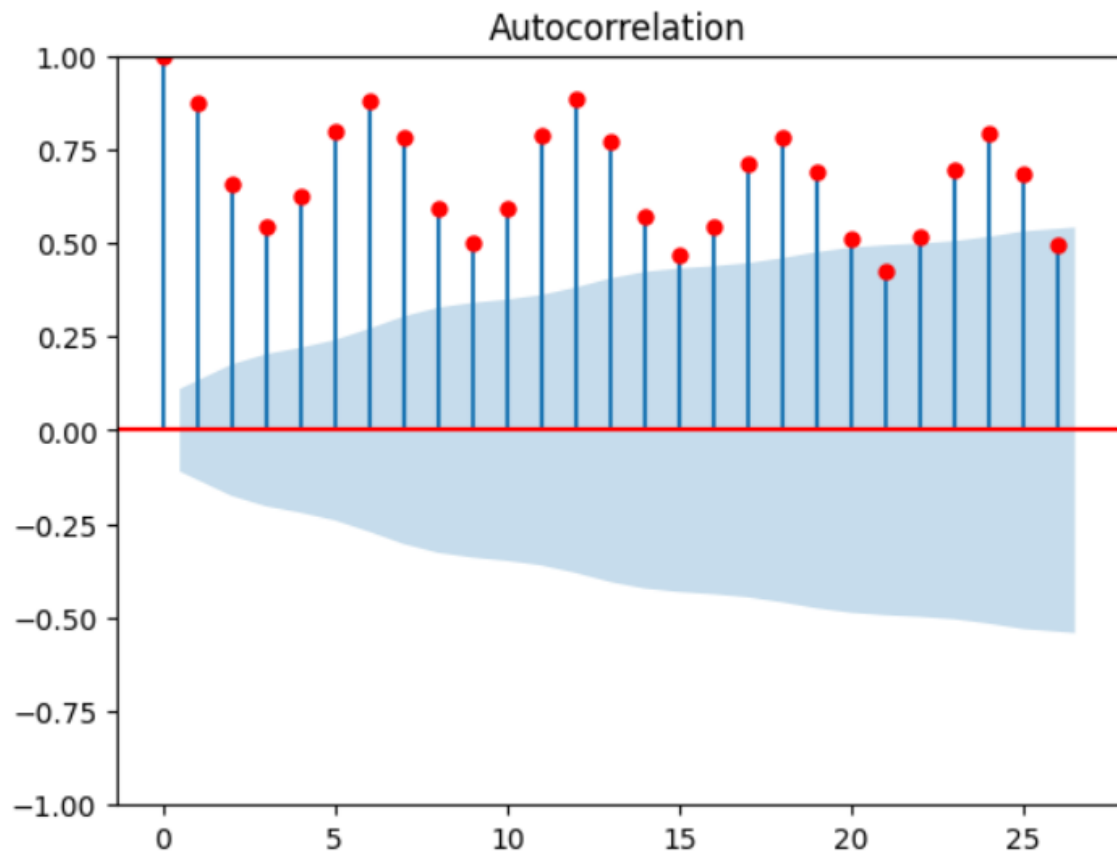


Terlihat pada grafik bahwa dataset yang digunakan telah disesuaikan menggunakan metode differencing. Pola kenaikan pada tren menjadi lebih stabil, dan lonjakan-lonjakan kenaikan yang sebelumnya terlihat pada grafik sudah tidak muncul lagi. Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF) adalah alat yang sering digunakan dalam analisis deret waktu untuk mengevaluasi ketergantungan antar lag dalam data.

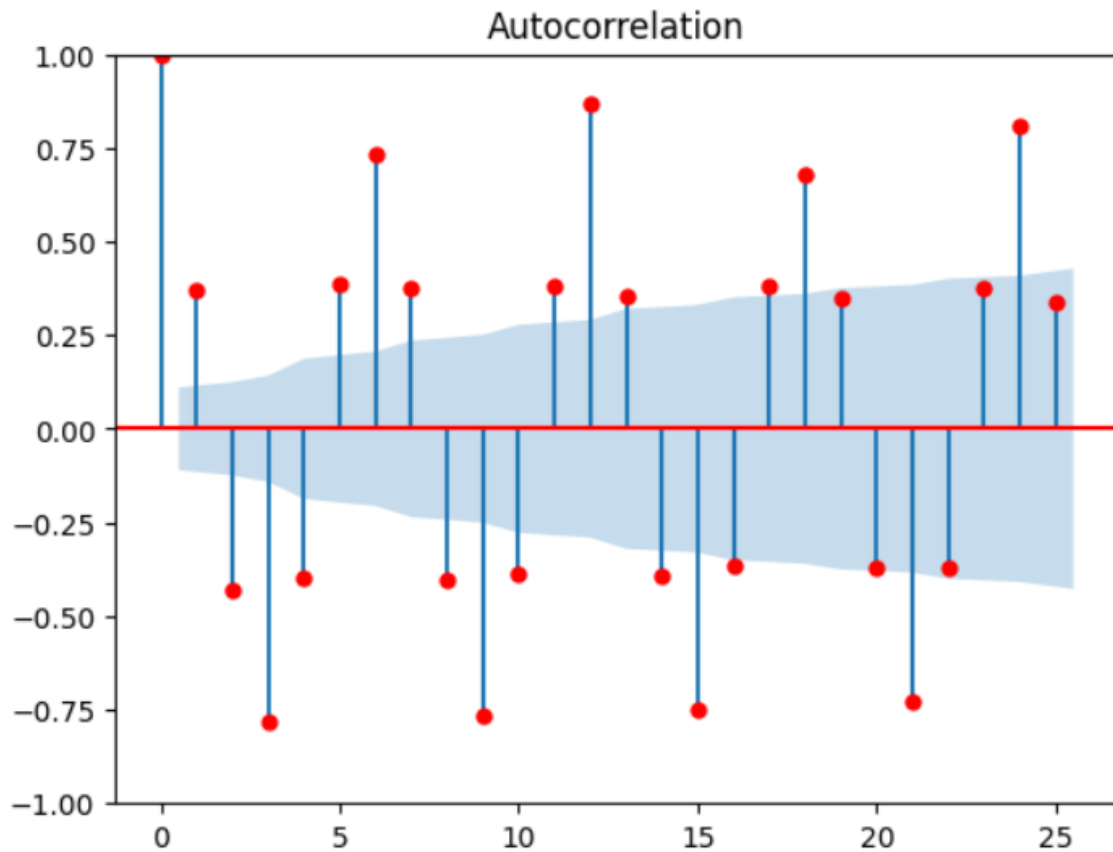
Namun, pengecekan untuk mengetahui apakah data stasioner atau tidak juga dapat dilakukan menggunakan Autocorrelation Function (ACF) dan Partial Autocorrelation Function (PACF), yang merupakan alat yang sering digunakan dalam analisis deret waktu untuk mengevaluasi ketergantungan antar lag dalam data.

3.3.1 ACF

Mengukur hubungan antara data dengan lag sebelumnya, sebelum dilakukan proses differencing, menunjukkan bahwa banyak nilai autokorelasi yang signifikan pada berbagai lag. Hal ini mengindikasikan bahwa data deret waktu ini kemungkinan besar tidak stasioner.

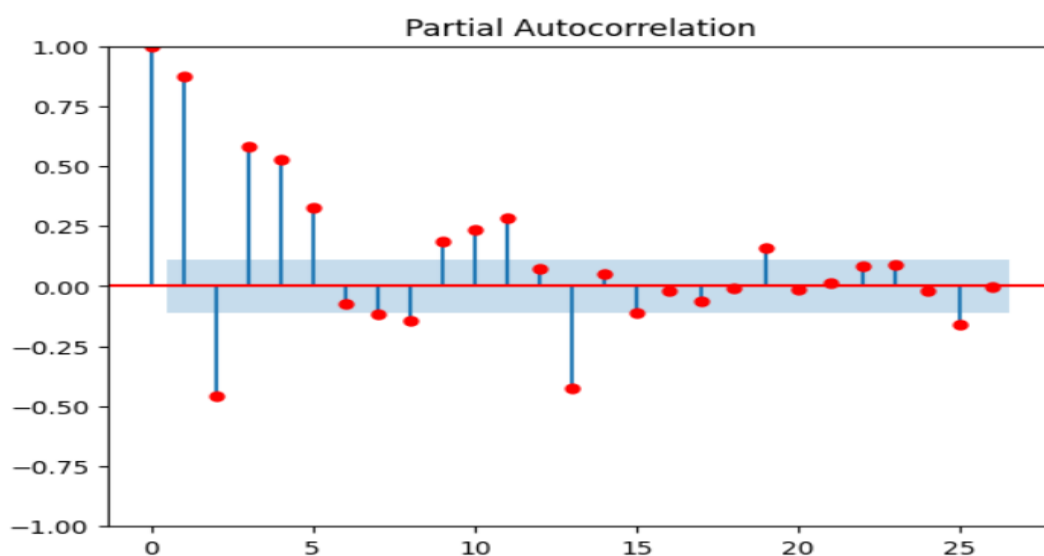


Namun, setelah dilakukan differencing, pola pada grafik ACF menunjukkan penurunan yang lebih cepat ke nol, dengan korelasi antar lag yang tidak signifikan pada lag lebih tinggi. Ini mengindikasikan bahwa data telah menjadi stasioner setelah proses differencing.

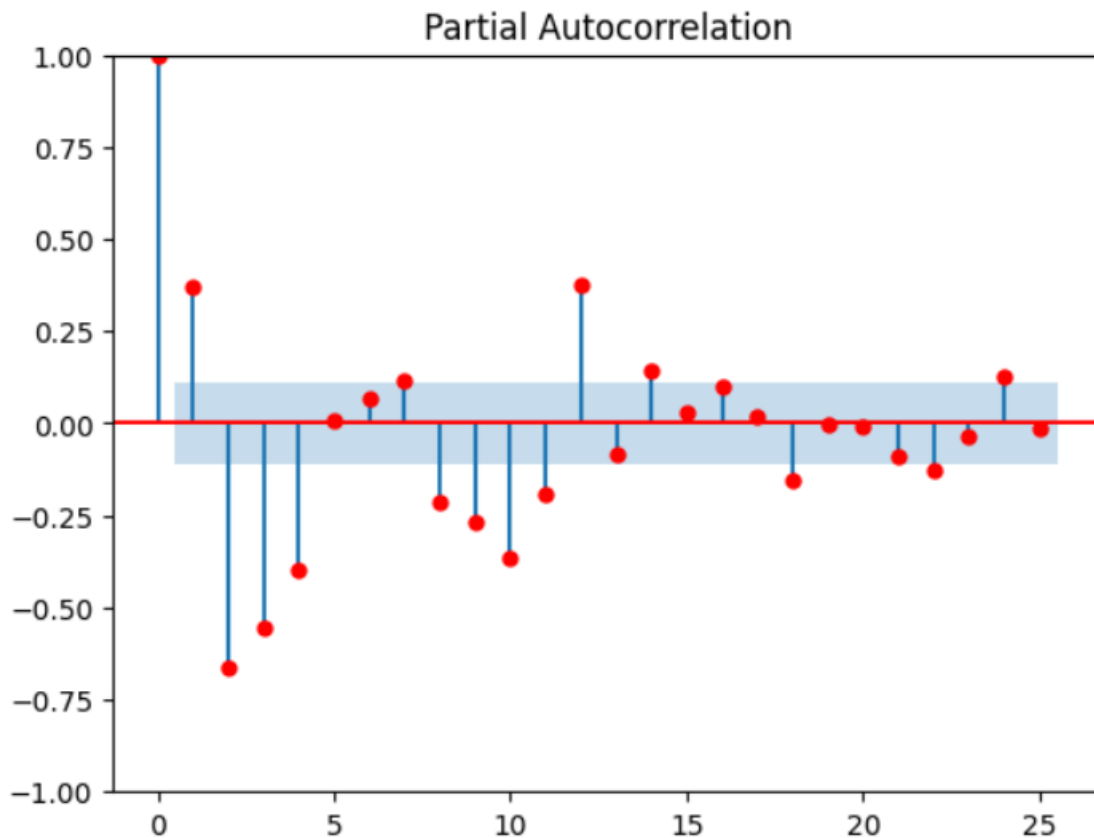


3.3.2 PACF

Mengukur hubungan antara data dengan lag sebelumnya, dengan menghilangkan pengaruh lag di antaranya. Sebelum dilakukan proses differencing, menunjukkan bahwa terdapat lonjakan signifikan pada lag 1 dan beberapa lag lainnya, tetapi sebagian besar nilai berada dalam interval kepercayaan, yang menunjukkan bahwa deret waktu ini mungkin stasioner.



Namun, setelah dilakukan differencing, pola pada grafik PACF menunjukkan bahwa korelasi parsial pada lag lebih tinggi tidak lagi signifikan, dengan hanya lag pertama yang menunjukkan nilai korelasi parsial yang signifikan. Hal ini mengindikasikan bahwa data telah menjadi stasioner setelah proses differencing.



3.4 Identifikasi Parameter Model ARIMA

Setelah data menjadi stasioner, langkah berikutnya adalah menentukan parameter model *ARIMA*: p , d , dan q . Parameter ini diperoleh melalui analisis grafik ACF (Autocorrelation Function) dan PACF (Partial Autocorrelation Function).

- **Parameter p :** Ditentukan dari jumlah lag signifikan pada grafik PACF.
- **Parameter q :** Ditentukan dari jumlah lag signifikan pada grafik ACF.
- **Parameter d :** Sudah ditentukan berdasarkan tingkat differencing yang dilakukan untuk mencapai stasioneritas.

3.4.2 Analisis Model ARIMA(1, 1, 1)

Informasi Umum Model

Model	ARIMA (1, 1, 1)
Data Latih	317 observasi
Log-Likelihood	-1022.680
AIC	2051.361
BIC	2062.628
HQIC	2055.862
Covariance Type	opg
ar.L1	0.0472
ma.L1	0.5736
sigma2	37.8392
Ljung-Box (L1)	0.05
Jarque-Bera (JB)	0.23
Heteroskedastisitas (H)	2.92

3.4.2 Analisis Model ARIMA(1, 1, 0)

Model	ARIMA (1, 1, 0)
Data Latih	317 observasi
Log-Likelihood	-1050.800
AIC	2105.600
BIC	2113.112
HQIC	2108.601
Covariance Type	opg
ar.L1	0.3706
ma.L1	0
sigma2	45.2535
Ljung-Box (L1)	19.32
Jarque-Bera (JB)	0.70
Heteroskedastisitas (H)	2.87

3.4.3 Analisis Model ARIMA(1, 1, 2)

Informasi Umum Model

Model	ARIMA (1, 1, 2)
Data Latih	317 observasi
Log-Likelihood	-973.506
AIC	1955.011
BIC	1970.034
HQIC	1961.013
Covariance Type	opg
ar.L1	0.3459
ma.L1	-0.2413
ma.L2	-0.6586
sigma2	27.6182
Ljung-Box (L1)	3.39
Jarque-Bera (JB)	1.50
Heteroskedastisitas (H)	2.59

3.4.4 Perbandingan Model ARIMA

Untuk menentukan model terbaik, perbandingan dilakukan berdasarkan beberapa kriteria evaluasi, seperti Log-Likelihood, AIC, BIC, dan diagnostik residu. Hasil perbandingan dapat dilihat pada tabel berikut:

Kriteria	ARIMA(1,1,0)	ARIMA(1,1,1)	ARIMA(1,1,2)
Log-Likelihood	-1.050.800	-1.022.680	-973.506
AIC	2.105.600	2.051.361	1.955.011
BIC	2.113.112	2.062.628	1.970.034
Ljung-Box (L1)	Signifikan	Tidak signifikan	Tidak signifikan
Sigma2	452.535	378.392	276.182

1. Log-Likelihood

- **Definisi:** Log-Likelihood mengukur seberapa baik model mencocokkan data. Nilai yang lebih tinggi (atau kurang negatif) menunjukkan kecocokan model yang lebih baik.

- **Hasil:**

ARIMA(1, 1, 2) memiliki nilai Log-Likelihood terbaik (-973.506), menunjukkan bahwa model ini paling cocok dengan data dibandingkan dua model lainnya.

2. AIC (Akaike Information Criterion)

- **Definisi:** AIC menilai kualitas model dengan mempertimbangkan keseimbangan antara kecocokan data dan kompleksitas model. Nilai yang lebih rendah lebih baik.

- **Hasil:**

ARIMA(1, 1, 2) memiliki nilai AIC terendah (1955.011), menjadikannya model terbaik di antara ketiga model.

4. Ljung-Box (L1)

- **Definisi:** Uji Ljung-Box mengevaluasi autokorelasi residu model. Jika probabilitas $(P) > 0.05$, maka tidak ada autokorelasi signifikan pada residu.

- **Hasil:**

ARIMA(1, 1, 0) memiliki $P\text{-value} < 0.05$ (signifikan), menunjukkan adanya autokorelasi residu, yang merupakan kelemahan model.

ARIMA(1, 1, 1) dan ARIMA(1, 1, 2) tidak memiliki autokorelasi residu ($P\text{-value} > 0.05$), menunjukkan bahwa residunya acak.

5. Sigma2 (Variansi Residu)

- **Definisi:** Sigma2 mengukur variansi dari kesalahan residu. Nilai yang lebih kecil menunjukkan model yang lebih stabil dan sesuai.

- **Hasil:**

ARIMA(1, 1, 2) memiliki nilai sigma2 terkecil (27.6182), menunjukkan bahwa model ini memiliki kesalahan residu yang paling kecil dibandingkan model lainnya.

3.5 Pembangunan Model ARIMA

Setelah menentukan parameter optimal untuk model ARIMA, langkah selanjutnya adalah membangun model menggunakan data latih (data_train) dengan parameter p, d, dan q. Namun, untuk mendapatkan model yang terbaik dan ideal, diperlukan evaluasi dan percobaan yang berulang agar konfigurasi p, d, dan q dapat disesuaikan secara optimal. Selama proses percobaan, diperoleh model dengan parameter p=8, d=1, dan q=2, yang menunjukkan tingkat keidealan terbaik dibandingkan dengan konfigurasi parameter lainnya.

3.5.1 Transformasi Data

Dengan data log-transformasi, model ARIMA dibangun menggunakan parameter p=8, d=1, dan q=2 yang dipilih berdasarkan analisis ACF dan PACF.

SARIMAX Results						
=====						
Dep. Variable:	log_Production	No. Observations:	317			
Model:	ARIMA(8, 1, 2)	Log Likelihood	621.018			
Date:	Fri, 10 Jan 2025	AIC	-1220.036			
Time:	14:23:13	BIC	-1178.723			
Sample:	01-01-1985	HQIC	-1203.532			
	- 05-01-2011					
Covariance Type:	opg					
=====						
	coef	std err	z	P> z	[0.025	0.975]

ar.L1	0.8739	0.072	12.133	0.000	0.733	1.015
ar.L2	-0.9250	0.089	-10.360	0.000	-1.100	-0.750
ar.L3	0.3579	0.107	3.343	0.001	0.148	0.568
ar.L4	-0.5749	0.112	-5.120	0.000	-0.795	-0.355
ar.L5	0.0305	0.108	0.283	0.778	-0.181	0.242
ar.L6	-0.1405	0.105	-1.341	0.180	-0.346	0.065
ar.L7	0.0284	0.103	0.276	0.783	-0.173	0.230
ar.L8	-0.4833	0.077	-6.258	0.000	-0.635	-0.332
ma.L1	-1.2541	0.066	-19.036	0.000	-1.383	-1.125
ma.L2	0.6918	0.064	10.860	0.000	0.567	0.817
sigma2	0.0011	8.93e-05	12.826	0.000	0.001	0.001
=====						
Ljung-Box (L1) (Q):	0.02	Jarque-Bera (JB):	8.73			

3.5.2 Analisis Model ARIMA(8, 1, 2) dengan Log-Transformasi

1. Informasi Umum Model

- **Model:** ARIMA(8, 1, 2)
- **Data Latih:** 317 observasi dari Januari 1985 hingga Mei 2011.
- **Log-Likelihood:** 621.018 (tinggi, menunjukkan kecocokan model yang baik).
- **AIC:** -1220.036 (rendah, menunjukkan model optimal dibandingkan model lain).
- **BIC:** -1178.723
- **HQIC:** -1203.532 (keseimbangan antara kecocokan data dan kompleksitas model).
- **Covariance Type:** opg (menggunakan *outer product of gradients* untuk estimasi matriks kovarians).

2. Koefisien Parameter

- **Parameter AR (Autoregressive):**
 - Koefisien ar.L1 hingga ar.L8 menunjukkan berbagai kontribusi dari lag autoregressive.
 - Koefisien signifikan ($P < 0.05$) untuk lag 1, 2, 3, 4, dan 8.
 - Lag 5, 6, dan 7 tidak signifikan ($P > 0.05$).
- **Parameter MA (Moving Average):**
 - Koefisien $ma.L1$ dan $ma.L2$ signifikan ($P < 0.05$):
 - $ma.L1$: -1.2541
 - $ma.L2$: 0.6918
- **sigma2 (Variansi Kesalahan):**
 - Nilai: 0.0011

- **Interpretasi:** Variansi residu sangat kecil, menunjukkan model mampu memprediksi data dengan akurasi tinggi.

3. Uji Diagnostik Model

- **Ljung-Box (L1) (Q):**
 - Nilai Q: 0.02
 - Probabilitas: $P > 0.05$
 - **Interpretasi:** Tidak ada autokorelasi signifikan pada residu, residu bersifat acak.
- **Jarque-Bera (JB):**
 - Nilai JB: 8.73
 - Probabilitas: $P > 0.05$
 - **Interpretasi:** Residu berdistribusi normal.
- **Heteroskedastisitas (H):**
 - Tidak ditampilkan dalam hasil, namun dapat diuji lebih lanjut jika diperlukan.

4. Interpretasi Koefisien

- **AR Parameters (Autoregressive):**
 - ar.L1, ar.L2, ar.L3, ar.L4, dan ar.L8 signifikan, menunjukkan lag ini memberikan kontribusi utama terhadap model.
 - ar.L5, ar.L6, dan ar.L7 tidak signifikan dan dapat dihapus dalam iterasi model berikutnya untuk menyederhanakan model.
- **MA Parameters (Moving Average):**
 - Kedua komponen MA signifikan dan membantu menangkap pola residu.

3.6 Evaluasi Model

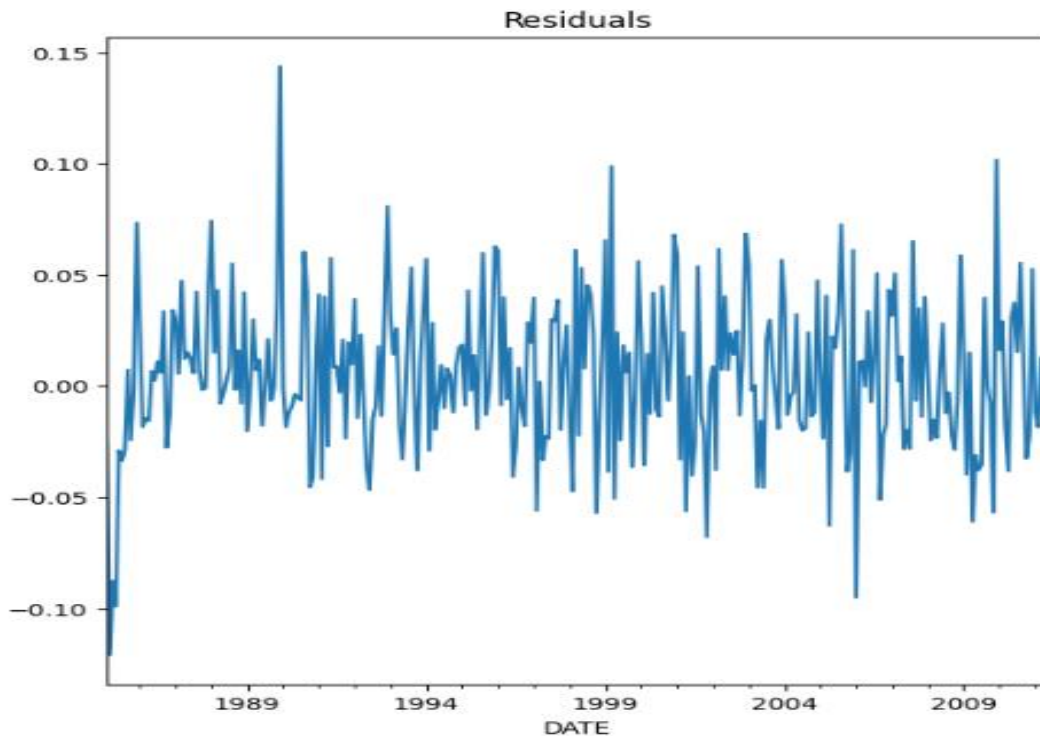
Setelah membangun model deret waktu dengan pendekatan ARIMA, langkah berikutnya adalah melakukan evaluasi untuk memastikan bahwa model yang dibangun dapat menggambarkan data dengan baik dan memberikan prediksi yang akurat. Evaluasi ini meliputi uji signifikansi parameter, pengujian terhadap kesalahan model, dan uji goodness-of-fit dengan menggunakan berbagai metrik.

Salah satu cara untuk mengevaluasi kualitas model adalah dengan menganalisis residuals. Residuals adalah selisih antara nilai yang diprediksi oleh model dan nilai aktual dalam data. Residuals yang baik seharusnya terdistribusi secara acak di sekitar nol, tanpa pola yang jelas, dan mengikuti distribusi normal. Untuk itu, dalam bagian ini, dilakukan analisis terhadap residuals model menggunakan dua jenis visualisasi.

3.6.1 Variasi Residual

Variasi residual mengacu pada fluktuasi atau penyebaran nilai residuals (selisih antara nilai yang diprediksi oleh model dan nilai observasi aktual) sepanjang waktu. Evaluasi variasi residual penting karena memberikan wawasan tentang seberapa baik model menangkap pola dalam data.

Pada dasarnya, jika model yang dibangun sudah sesuai, residuals seharusnya terdistribusi secara acak tanpa pola yang jelas. Artinya, tidak ada hubungan yang dapat diprediksi antara residuals di berbagai waktu, dan model tidak meninggalkan pola yang signifikan dalam data. Sebaliknya, jika residuals menunjukkan pola tertentu, ini bisa menjadi indikasi bahwa model belum sepenuhnya mengcapture informasi penting dalam data, seperti tren musiman atau non-linearitas yang tidak ditangani dengan baik.



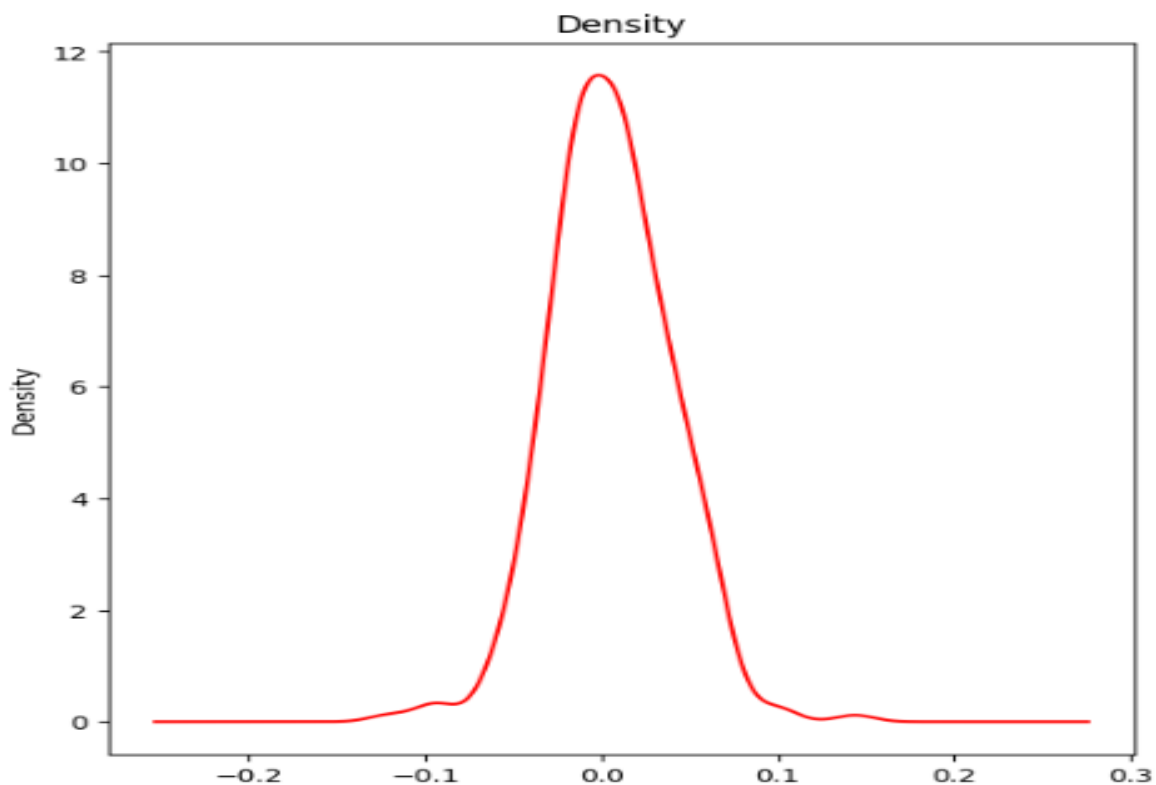
Dalam grafik ini, terlihat bahwa residuals tidak membentuk pola yang jelas. Hal ini karena grafik menunjukkan fluktuasi acak di sekitar garis nol, yang mengindikasikan bahwa model telah berhasil menangkap pola utama dalam data. Tidak adanya pola yang terstruktur pada residuals menunjukkan bahwa kesalahan model tersebar secara acak dan tidak mengandung informasi yang belum tertangkap oleh model, yang merupakan indikasi bahwa model yang dibangun cukup baik dan sesuai untuk data yang dianalisis.

3.6.2 Kernel Density Estimation (KDE)

Kernel Density Estimation (KDE) adalah metode statistik yang digunakan untuk memperkirakan fungsi kepadatan probabilitas (probability density function, PDF) dari data yang tidak terdistribusi secara normal atau data yang tidak teratur. Dalam konteks evaluasi model deret waktu, KDE digunakan untuk menganalisis distribusi dari residuals, yang merupakan selisih antara nilai yang diprediksi oleh model dan nilai aktual.

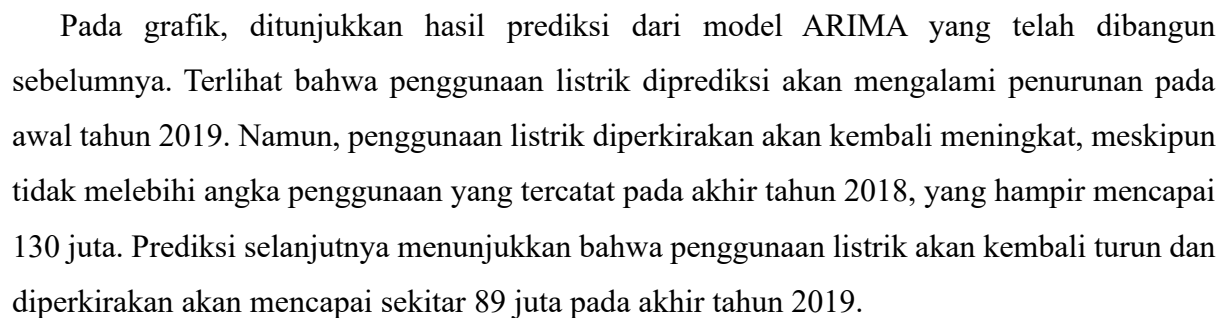
KDE bekerja dengan menggambarkan sebaran data dalam bentuk kurva halus, yang dapat memberikan gambaran yang lebih jelas mengenai distribusi residuals dibandingkan dengan histogram biasa. Pada grafik KDE, sumbu-x menunjukkan nilai residuals, sementara sumbu-y menunjukkan estimasi kepadatan atau kemungkinan distribusi residuals tersebut. Jika model yang dibangun telah menangkap pola data dengan baik, maka residuals diharapkan memiliki distribusi yang mendekati distribusi normal. Dengan kata lain, grafik KDE dari residuals

seharusnya berbentuk kurva lonceng yang simetris di sekitar nilai nol, yang mengindikasikan bahwa kesalahan model terdistribusi secara acak dan tidak mengandung pola yang signifikan.



Dalam grafik ini, terlihat bahwa distribusi residuals mengikuti bentuk kurva lonceng yang simetris di sekitar nilai nol, yang mengindikasikan bahwa residuals terdistribusi secara normal. Hal ini menunjukkan bahwa model yang dibangun berhasil menangkap pola utama dalam data, dan kesalahan model tersebar secara acak tanpa pola yang dapat diprediksi. Oleh karena itu, model ini dapat dianggap cocok untuk data yang dianalisis, karena distribusi residuals yang normal merupakan tanda bahwa tidak ada informasi penting yang terlewatkan dalam proses pemodelan.

Pada tahap ini, model yang telah dibangun dan dievaluasi digunakan untuk melakukan prediksi terhadap data di masa depan. Periode prediksi yang digunakan adalah 12 bulan atau 1 tahun, yang bertujuan untuk memberikan gambaran mengenai perilaku data di masa yang akan datang berdasarkan pola yang telah teridentifikasi oleh model.



BAB IV PENUTUP

4.1 Kesimpulan

Pada tugas besar ini, digunakan dataset yang berisi data produksi listrik dari tahun 1985 hingga 2018. Awalnya, dataset tersebut tidak stasioner, namun setelah diterapkan metode differencing, data berhasil dijadikan stasioner. Langkah selanjutnya adalah menentukan model yang paling sesuai untuk melakukan prediksi terhadap produksi listrik di masa depan. Berdasarkan analisis yang dilakukan, model ARIMA (AutoRegressive Integrated Moving Average) dipilih sebagai model yang ideal untuk memprediksi produksi listrik, mengingat kemampuannya dalam menangani data deret waktu yang memiliki komponen tren dan musiman.

Hasil prediksi menunjukkan adanya penurunan pada awal tahun 2019, diikuti oleh kenaikan yang tidak melebihi angka penggunaan pada akhir 2018, dan diakhiri dengan penurunan pada akhir tahun 2019. Model ARIMA yang dibangun memberikan gambaran yang cukup baik mengenai perkembangan produksi listrik dalam 12 bulan ke depan, yang dapat digunakan untuk perencanaan atau pengambilan keputusan yang lebih baik di masa yang akan datang.

4.2 Saran

Berdasarkan analisis dan hasil prediksi yang telah dilakukan dalam tugas besar ini, beberapa saran yang dapat diberikan untuk pengembangan lebih lanjut adalah sebagai berikut:

1. Peningkatan Akurasi Model: Meskipun model ARIMA telah memberikan hasil yang cukup baik, masih terdapat ruang untuk meningkatkan akurasi prediksi.
2. Penggunaan Data Tambahan: Untuk meningkatkan kualitas prediksi, disarankan untuk mempertimbangkan penggunaan data tambahan yang relevan, seperti faktor cuaca, kebijakan energi, atau perubahan teknologi yang dapat mempengaruhi konsumsi listrik.
3. Menambah Tahun Prediksi: Untuk meningkatkan manfaat dari prediksi yang dihasilkan, disarankan untuk memperpanjang periode prediksi tidak hanya 12 bulan, tetapi juga untuk jangka waktu yang lebih panjang, seperti 24 bulan atau lebih. Hal ini dapat memberikan wawasan yang lebih komprehensif bagi pemangku kepentingan dalam merencanakan kebijakan energi dan investasi di sektor kelistrikan.