

# KUMPULAN ANALISIS MODEL ESTIMASI DATA KOMODITAS PERKEBUNAN



Karet



Kopi



Kakao



Tebu



Lada



Teh



Cengkeh





**BUKU 1**  
**KUMPULAN ANALISIS MODEL**  
**ESTIMASI DATA KOMODITAS PERKEBUNAN**

**Pusat Data dan Sistem Informasi Pertanian**  
**Sekretariat Jenderal Kementerian Pertanian**

**2021**



# **BUKU 1**

## **KUMPULAN ANALISIS MODEL**

### **ESTIMASI DATA KOMODITAS PERKEBUNAN**

#### **Pengarah :**

Roby Darmawan, M.Eng.

#### **Penanggung Jawab :**

Dr. Ir. Anna Astrid Susanti, M.Si.

Rhendy Kencana Putra W, S.Si., M.AppStat.

#### **Ketua Tim :**

Ir. Mohammad Chafid, M.Si.

#### **Anggota Tim :**

1. Ir. Efi Respati, M.Si. - Pusdatin
2. Ir. Roch Widaningsih, M.Si. - Pusdatin
3. Diah Indarti, SE. MM - Pusdatin
4. Yuliawati Rohmah, SP., M.SE. - Pusdatin
5. Roydatul Zikria, SSi, M.SE - Pusdatin
6. Ir. Vera Junita S – Pusdatin
7. Ucik Mawarsari– BPS
8. Lasmiyati, SST, SAB M.Stat – BPS
9. Afifah Siti Muslikhah – BPS

#### **Desain Sampul :**

Suyati, S.Kom

#### **Diterbitkan oleh :**

Kementerian Pertanian  
Pusat Data dan Sistem Informasi Pertanian  
2021



# KATA PENGANTAR

Puji dan syukur kami panjatkan kehadiran Allah SWT atas berkat rahmat dan hidayah-Nya sehingga Buku Kumpulan Analisis Model Estimasi Data Komoditas Perkebunan dapat diselesaikan. Buku Kumpulan Analisis Model Estimasi Data Komoditas Perkebunan disusun dalam rangka melakukan pengembangan metode estimasi data perkebunan, sehingga diharapkan akan membantu penyusunan Angka Estimasi Data Perkebunan Tahun 2022. Buku ini berisi kumpulan hasil analisis estimasi komoditas utama perkebunan dengan menggunakan model-model statistik.

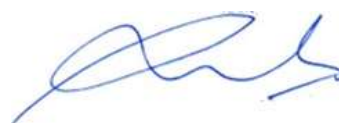
Buku Kumpulan Analisis Model Estimasi Data Komoditas Perkebunan ini disusun oleh beberapa penulis dari Fungsional Statistisi lingkup Sub Peternakan dan Perkebunan- Pusdatin, Fungsional Statistisi Lingkup Ditjen Perkebunan, dan penulis dari Direktorat Statistik Tanaman Pangan Hortikultura dan Perkebunan - BPS. Buku ini merupakan hasil kerjasama Pusdatin dengan Direktorat Jenderal Perkebunan selaku penyedia data, dan Badan Pusat Statistik sebagai pembina data. Buku ini mencakup hasil kajian estimasi produksi dan luas areal beberapa komoditas perkebunan strategis dengan pendekatan Model Arima, Model Fungsi Transfer, dan Model VAR (*Vector Auto Regression*).

Keberhasilan dalam menyusun angka estimasi data perkebunan yang lebih akurat sangat ditentukan oleh kesungguhan dan kesadaran akan pentingnya akurasi data estimasi yang dihasilkan, sehingga diharapkan dengan buku kajian analisis estimasi ini, akan menjadi acuan dan referensi dalam menyusun angka estimasi data perkebunan baik di tingkat pusat maupun di tingkat provinsi.

Kami mengucapkan terimakasih yang sebesar-besarnya kepada semua pihak yang telah membantu dalam penyusunan Buku Kumpulan Analisis Model Estimasi Data Komoditas Perkebunan ini. Kami menyadari bahwa buku ini masih perlu disempurnakan lagi, sehingga saran dan masukan untuk perbaikan buku ini ke depan sangat diharapkan.

Jakarta, Desember 2021

Kepala Pusat Data dan  
Sistem Informasi Pertanian



Roby Darmawan, M.Eng.  
NIP. 196912151991011001





# DAFTAR ISI

Halaman

<b>KATA PENGANTAR.....</b>	<b>v</b>
<b>DAFTAR ISI.....</b>	<b>vii</b>
<b>Model Estimasi Luas Areal Karet Nasional Tahun 2022-2025 Menggunakan Model ARIMA, Fungsi Transfer dan VAR (Vector Auto Regressive)</b>	
Mohammad Chafid .....	1-52
<b>Kajian Model Peramalan Luas Areal Kopi di Indonesia Pendekatan Model ARIMA, Fungsi Transfer dan VAR</b>	
Ucik Mawarsari.....	53-74
<b>Estimasi Luas Areal Tebu di Indonesia</b>	
Afifah Siti Muslikhah .....	75-88
<b>Kajian Model Peramalan Luas Areal Lada di Indonesia Pendekatan Model ARIMA, Fungsi Transfer dan VAR</b>	
Lasmiyati.....	89-110
<b>Kajian Metode Estimasi Luas Areal Teh Indonesia</b>	
Diah Indarti .....	111-152
<b>Kajian Metodologi Estimasi Data Luas Areal Cengkeh</b>	
Vera Junita Siagian .....	153-185
<b>Model Estimasi Produksi Karet Nasional 2022-2025 Menggunakan Model ARIMA, Fungsi Transfer dan VAR (Vector Auto Regressive)</b>	
Mohammad Chafid .....	187-235
<b>Kajian Model Peramalan Produksi Kopi di Indonesia Pendekatan Model ARIMA, Fungsi Transfer dan VAR</b>	
Roch Widaningsih.....	237-267
<b>Estimasi Produksi Gula Indonesia, Pendekatan Metode ARIMA, VAR Dan Fungsi Transfer</b>	
Efi Respati .....	269-289
<b>Kajian Metode Estimasi Produksi Teh Indonesia</b>	
Diah Indarti .....	291-331

**Kajian Metodologi Estimasi Data Produksi Cengkeh**

Vera Junita Siagian .....333-367

**Kajian Model Estimasi Produksi Lada di Indonesia**

Roydatul Zikria .....369-402

**Analisis Estimasi Produksi Kakao Indonesia Melalui Pendekatan**

**Model ARIMA, Fungsi Transfer dan VAR**

Yulawati Rohmah, SP, MSE .....403-438

**MODEL ESTIMASI LUAS AREAL KARET NASIONAL TAHUN 2022- 2025  
MENGUNAKAN MODEL ARIMA, FUNGSI TRANSFER DAN VAR (Vector Auto  
Regressive)**

**Mohammad Chafid**

*Center for Agricultural Data and Information System-Ministry of Agriculture  
Jalan Harsono RM No. 3 Gd D Lt 4, Ragunan-Jakarta Selatan, Indonesia  
E-mail: mohammad.hafidz1@gmail.com*

**ABSTRAK**

Komoditas karet merupakan salah satu komoditas strategis penghasil devisa negara karena sebagian besar produksi karet nasional untuk diekspor. Status Angka statistik perkebunan terdiri dari Angka Tetap, Angka Sementara dan Angka Estimasi. Tujuan penulisan ini adalah mencari model alternatif lain untuk menyusun angka estimasi luas areal karet sehingga akurasi menjadi lebih baik yang ditandai dengan semakin kecilnya MAPE baik untuk data training maupun testing.

Model yang digunakan untuk menyusun angka estimasi luas areal karet meliputi, Metode Arima, Metode Fungsi Transfer dengan peubah input harga karet sheet nasional, Metode VAR (*Vector Autoregressive*) dengan variabel luas areal, produksi, harga karet sheet dalam negeri, harga karet dunia, volume ekspor karet dan volume impor karet. Sumber data yang digunakan untuk variabel luas areal karet (1971 – 2020), harga karet nasional (1971-2020), volume ekspor dan impor karet ( 1971 – 2020) berasal dari Ditjen Perkebunan. Untuk variabel harga karet dunia (1971 - 2020) berasal dari World Bank. Run model menggunakan software RStudio.

Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1971 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan. Model ARIMA terbaik adalah ARIMA (2,1,0), menghasilkan MAPE untuk data training 1,32%, dan MAPE data testing 0,45%. Model ARIMA (1,1,1) juga menghasilkan MAPE yang cukup baik, yaitu MAPE training 1,28% dan MAPE testing 0,61%. Untuk model yang dua dengan menggunakan Fungsi Transfer ARIMA (2,1,0) dengan variabel input harga karet sheet, menghasilkan MAPE untuk data training sebesar 1,33% dan MAPE data testing 0,35%. Untuk model yang keempat model VAR(2) type ‘both’ ada pengaruh konstanta dan trend, menghasilkan MAPE data training 1,07% dan data MAPE data testing 9,51%. Model tentatif VAR adalah VAR(4) type ‘both’, menghasilkan MAPE training 0,61% dan MAPE testing 3,43%.

Berdasarkan perbandingan besarnya MAPE baik data testing maupun data training dan hasil estimasi luas areal 5 tahun kedepan, maka model terbaik yang terpilih adalah model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet sheet karena menghasilkan akurasi yang paling tinggi sehingga MAPE rata-rata data testing sebesar 0,35%. Hasil estimasi luas areal karet nasional untuk model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet sheet untuk tahun 2021 sebesar 3.747.113 hektar dan tahun 2022 sebesar 3.762.496 hektar, tahun 2023 sebesar 3.771.038 hektar, tahun 2024 sebesar 3.775.459 hektar, dan tahun 2025 sebesar 3.778.752 hektar. Laju pertumbuhan estimasi luas areal karet nasional selama 5 tahun kedepan (2021 – 2025) rata-rata 0,21% per tahun. Namun jika mempertimbangkan program yang telah dijalankan oleh direktorat teknis terkait untuk meningkatkan luas areal karet maka ada optimisme estimasi luas karet tahun 2021 sebesar 3.787.518 hektar dan tahun 2022 sebesar 3.827.540 hektar berdasarkan hasil pemodelan dengan VAR(2) type “both”.

**Kata Kunci : Luas Areal Karet, Arima, Fungsi Transfer, VAR (Vector Autoregressive)**

Data statistik perkebunan yang disajikan merupakan hasil sinkronisasi dan validasi yang dilaksanakan secara berjenjang dari tingkat kabupaten sampai tingkat nasional dengan mengacu pada Pedoman Pengelolaan Data Komoditas Perkebunan yang diterbitkan oleh Direktorat Jenderal Perkebunan. Secara periodik, data perkebunan disajikan dalam 3 status angka yaitu Angka Tetap (ATAP), Angka Sementara (ASEM) dan Angka Estimasi (AESTI). Angka Tetap merupakan angka hasil rekapitulasi dari pelaporan yang sudah tetap, sehingga tidak dilakukan estimasi, sedangkan untuk penentuan Angka Sementara dan Estimasi perlu dilakukan estimasi dengan metode estimasi yang paling relevan dan tepat (PDKP, 2013). Data Angka Tetap (ATAP) merupakan data 1 tahun yang lalu ( $n-1$ ), Angka Sementara (ASEM) merupakan data tahun berjalan ( $n$ ), dan Angka Estimasi (AESTI) merupakan data tahun satu tahun kedepan ( $n+1$ ).

Metode estimasi yang digunakan adalah Metode *Exponential Smoothing* (Peramalan Pemulusan Eksponensial) yang merupakan salah satu kategori metode time series yang menggunakan pembobotan data masa lalu secara eksponensial. Dalam kategori ini terdapat dua metode yang umum dipakai yaitu metode Pemulusan Eksponensial Tunggal (*Single Exponential Smoothing*) dan metode Pemulusan Eksponensial Ganda (*Double Exponential Smoothing*). Pemilihan model *Single Exponential Smoothing* atau *Double Exponential Smoothing* harus mempertimbangkan nilai *Mean Absolute Percentage Error (MAPE)* serta kerealistisan hasil estimasi bila dibandingkan dengan series data sebelumnya. MAPE adalah pengukur tingkat akurasi (ketepatan) nilai dugaan yang dihasilkan oleh model dalam bentuk presentase. Model yang mempunyai nilai MAPE lebih kecil dianggap sebagai model yang lebih baik. Keunggulan dari metode estimasi ini adalah dapat digunakan untuk meramalkan data yang berisi trend atau pola musiman. Namun metode estimasi ini juga memerlukan keahlian khusus dalam menginterpretasikan hasil estimasi yang diperoleh (PDKP, 2013).

Metode untuk menghasilkan angka estimasi (AESTI) yang diliris oleh Ditjen. Perkebunan pada waktu tahun berjalan ( $n$ ) perlu dikaji kembali, agar didapatkan metode yang lebih akurat, lebih objektif dan lebih baik secara statistik dari metode sebelumnya, sehingga tingkat kesalahannya lebih kecil. Guna merumuskan kebijakan produksi, pasokan dan distribusi komoditas karet sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Metode ramalan luas areal karet dapat digunakan dengan menggunakan beberapa metode statistik yang biasa digunakan untuk melakukan peramalan.

Pada analisis ini akan dikaji metode ARIMA, Fungsi Transfer dan VAR dalam melakukan pemodelan dan peramalan luas areal karet nasional.

Oleh karenanya, tujuan dari disusunnya analisis ini adalah:

- a. Melakukan analisis dan peramalan data luas areal karet nasional menggunakan model ARIMA, Fungsi Transfer dan VAR.
- b. Menentukan metode terbaik dalam meramal data luas areal karet nasional

## METODOLOGI

Data yang digunakan dalam melakukan pemodelan luas areal karet nasional adalah data series luas areal karet nasional tahun 1970 - 2020. Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1970 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan.

Peubah yang diasumsikan mempengaruhi besaran luas areal karet untuk Model Fungsi Transfer adalah harga karet sheet lokal. Harga karet lokal diasumsikan akan berpengaruh terhadap luas areal, karena jika terjadi kenaikan harga karet maka luas areal karet cenderung akan terus ditingkatkan, kondisi sebaliknya maka harga karet cenderung stagnan, pertumbuhan luas areal rendah, atau cenderung turun.

Peubah yang berpengaruh terhadap luas areal untuk pemodelan VAR adalah luas areal, harga karet domestik, harga karet dunia, volume ekspor dan impor karet. Harga karet dunia dan harga karet domestik diduga kuat berpengaruh pada tingkat luas areal. Untuk volume ekspor dan impor juga berpengaruh terhadap luas areal, dimana semakin tinggi permintaan ekspor karet maka luas areal cenderung akan semakin tinggi.

Peramalan data luas areal karet, dilakukan uji coba beberapa metode yakni mengkaji metode ARIMA, Regresi, Fungsi Transfer dan VAR (Vector Autoregression), menggunakan software *R Studio*.

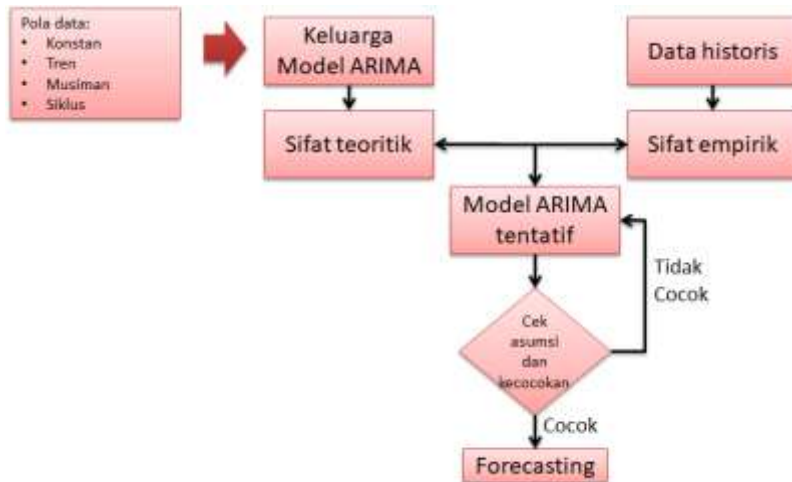
### **a. ARIMA (*Autoregressive Integrated Moving Average*)**

Metode Autoregressive Integrated Moving Average (ARIMA) atau biasa disebut juga sebagai metode Box-Jenkins merupakan metode yang secara intensif dikembangkan oleh George Box dan Gwilym Jenkins pada tahun 1970 (Iriawan, 2006).

ARIMA sering juga disebut metode runtun waktu Box-Jenkins. ARIMA sangat baik ketepatannya untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatan peramalannya kurang baik. Biasanya akan cenderung flat (mendatar/konstan) untuk periode yang cukup panjang.

Model *Autoregresif Integrated Moving Average* (ARIMA) adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari deret waktu (time series) secara statistik berhubungan satu sama lain (dependent).

Model ARIMA terdiri dari tiga langkah dasar, yaitu tahap identifikasi, tahap penaksiran dan pengujian, dan pemeriksaan diagnostik. Selanjutnya model ARIMA dapat digunakan untuk melakukan peramalan jika model yang diperoleh memadai.



Gambar 1. Prosedur Peramalan Model Arima (Box- Jenkins)

Hal yang perlu diperhatikan adalah bahwa kebanyakan deret berkala bersifat stasioner. Stasioner berarti tidak terdapat pertumbuhan atau penurunan pada data. Data secara kasarnya harus horizontal sepanjang sumbu waktu. Dengan kata lain, fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan varians dari fluktuasi tersebut pada pokoknya tetap konstan setiap waktu. Suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melakukan differencing. Yang dimaksud dengan differencing adalah menghitung perubahan atau selisih nilai observasi. Nilai selisih yang diperoleh dicek lagi apakah stasioner atau tidak. Jika belum stasioner maka dilakukan differencing lagi. Jika varians tidak stasioner, maka dilakukan transformasi logaritma.

Model Box-Jenkins (ARIMA) dibagi kedalam 3 kelompok, yaitu: model autoregressive (AR), moving average (MA), dan model campuran ARIMA (autoregressive moving average) yang mempunyai karakteristik dari dua model pertama. ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampau. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-p)

$\mu$  = suatu konstanta

$\theta_1, \theta_q, \phi_1, \phi_n$  = parameter-parameter model

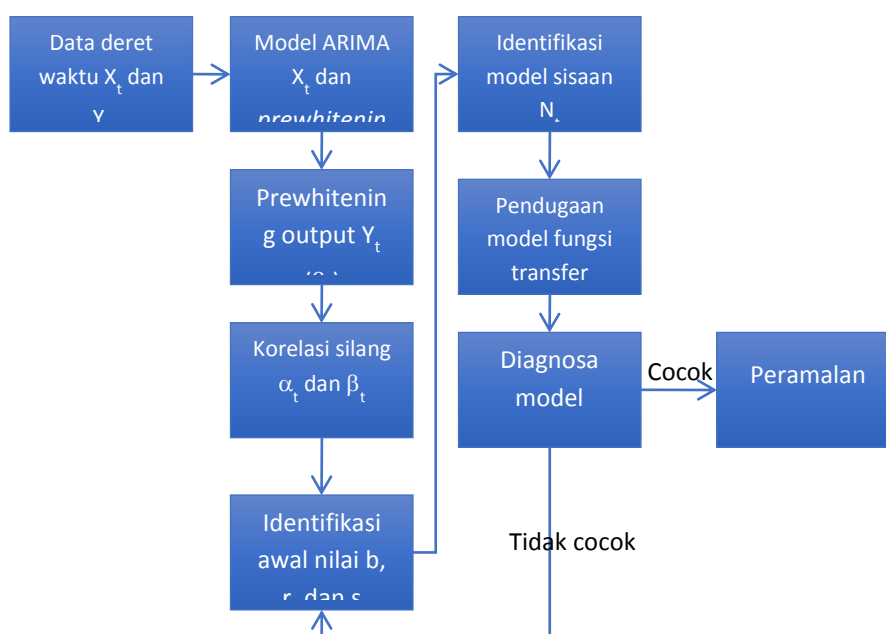
$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

Model dikatakan baik jika nilai error bersifat random, artinya sudah tidak mempunyai pola tertentu lagi. Dengan kata lain model yang diperoleh dapat menangkap dengan baik pola data yang ada. Untuk melihat kerandoman nilai error dilakukan pengujian terhadap nilai koefisien autokorelasi dari error, dengan menggunakan salah satu dari dua statistik berikut, yaitu Uji Q-Box and Pierce dan uji Ljung-Box.

## b. Model Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke- $t$ , tetapi juga pada waktu  $t+1, t+2, \dots, t+k$ . Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara peubah *input* dan peubah *output*.

Tujuan pemodelan fungsi transfer adalah untuk menetapkan model yang sederhana, yang menghubungkan deret output ( $Y_i$ ) dengan deret input ( $X_i$ ) dan gangguan/noise( $n_i$ ). Wei (1994) juga menjelaskan bahwa di dalam fungsi transfer terdapat rangkaian output yang mungkin dipengaruhi oleh rangkaian multiple input. Pada kasus single input peubah, dapat menggunakan metode korelasi silang yang dianjurkan oleh Box and Jenkins (1976). Teknik ini juga dapat digunakan ketika terdapat single input peubah yang lebih dari satu selama antar variable input tidak berkorelasi silang. Jika beberapa atau semua peubah input berkorelasi silang maka teknik *prewhitening* atau metode korelasi silang tidak dapat digunakan secara langsung. Alasan utama bagi perlunya suatu perencanaan atau peramalan adalah adanya tenggang waktu pengambilan keputusan yang dapat berkisar dari beberapa hari atau sampai beberapa tahun. Pada analisis fungsi transfer untuk peramalan deret berkala univariate, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (noise). Deret input mempengaruhi deret output melalui sebuah fungsi transfer yang mendistribusikan pengaruhnya secara dinamis melalui beberapa periode waktu yang akan datang dengan persentase tertentu yang disebut sebagai bobot respons impuls atau bobot fungsi transfer.



Gambar 2. Langkah-langkah melakukan pemodelan Fungsi Transfer

Model umum Fungsi Transfer:

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\varphi_p(B)} \varepsilon_t \quad \dots\dots\dots(7)$$

Dimana:

- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$
- $q \rightarrow$  ordo MA bagi noise  $N_t$

**c. Model Vector Autoregression (VAR)**

*Vector Autoregression* (VAR) dikemukakan pertama kali oleh Christopher Sims (1980). Sims mengembangkan model ekonometri dengan mengabaikan pengujian asumsi secara apriori. VAR dikembangkan oleh Sims sebagai kritik atas metode simultan. Jumlah peubah yang besar dan klasifikasi endogen dan eksogen pada metode simultan merupakan dasar dari kritik tersebut. Penggunaan pendekatan struktural atas pemodelan persamaan simultan biasanya menerapkan teori ekonomi di dalam usahanya untuk mendeskripsikan hubungan antar peubah yang ingin diuji. Akan tetapi sering ditemukan bahwa teori ekonomi saja ternyata tidak cukup kaya di dalam menyediakan spesifikasi yang ketat dan tepat atas hubungan dinamis antar peubah (Gujarati, 2010). Model VAR merupakan jalan keluar atas permasalahan ini melalui pendekatan non-struktural atau tidak berdasarkan pada teori ekonomi tertentu (*atheoretical*). Metode VAR memperlakukan seluruh peubah secara simetris tanpa memperlakukan peubah dependen dan independen atau dengan kata lain model ini memperlakukan seluruh peubah sebagai peubah endogen, karena pada kenyataannya suatu peubah dapat bertindak sebagai endogen atau dependen dari suatu peubah yang lainnya (Gujarati, 2010).

Kelebihan dalam penggunaan metode VAR (Gujarati, 2010):

- a. Kemudahan dalam penggunaan, tidak perlu mengkhawatirkan tentang penentuan peubah endogen dan peubah eksogen.
- b. Kemudahan dalam estimasi, metode *Ordinary Least Square* (OLS) dapat diaplikasikan pada tiap persamaan secara terpisah.
- c. *Forecast* atau peramalan yang dihasilkan pada beberapa kasus ditemukan lebih baik daripada yang dihasilkan oleh model persamaan simultan yang kompleks.
- d. Perangkat estimasi yang digunakan adalah *Impulse Respon Function* (IRF) untuk melacak respon dari peubah dependen dalam sistem VAR terhadap *shock* dari *error term* dan *Variance Decomposition* yang memberikan informasi mengenai pentingnya masing-masing *error term* dalam mempengaruhi peubah-peubah dalam VAR.



Di sisi lain, terdapat beberapa kritik terhadap model VAR menyangkut permasalahan berikut (Gujarati, 2010) :

- 1) Model VAR merupakan model yang *atheoritic* atau tidak berdasarkan teori, hal ini tidak seperti pada persamaan simultan. Pada persamaan simultan, pemilihan peubah yang akan dimasukkan dalam persamaan memegang peranan penting dalam mengidentifikasi model.
- 2) Pada model VAR penekanannya terletak pada *forecasting* atau peramalan sehingga model ini kurang cocok digunakan dalam menganalisis kebijakan.
- 3) Permasalahan yang besar dalam model VAR adalah pada pemilihan *lag length* atau panjang lag yang tepat. Karena semakin panjang lag, maka akan menambah jumlah parameter yang akan bermasalah pada *degrees of freedom*.
- 4) Peubah yang tergabung pada model VAR harus stasioner. Apabila tidak stasioner, perlu dilakukan transformasi bentuk data, misalnya melalui *first difference*.
- 5) Sering ditemui kesulitan dalam menginterpretasi tiap koefisien pada estimasi model VAR, sehingga sebagian besar peneliti melakukan interpretasi pada estimasi fungsi *impulse respon*.

#### d. Kriteria Pemilihan Model Terbaik

Untuk menguji kebaikan suatu model ada beberapa kriteria yang digunakan. Pada buku pedoman teknis ini kriteria yang digunakan adalah MAPE. Model time series/arima, model regresi, model fungsi transfer, dan model VAR dapat digunakan untuk melakukan estimasi variabel untuk beberapa tahun ke depan. Untuk model *time series* baik analisis ARIMA, model regresi, model fungsi transfer, dan model VAR, ukuran kelayakan model berdasarkan nilai kesalahan dengan menggunakan statistik MAPE (*mean absolute percentage error*) atau kesalahan persentase absolut rata-rata yang diformulasikan sebagai berikut:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right| \cdot 100$$

Dimana :  $X_t$  adalah data aktual

$F_t$  adalah nilai ramalan.

Semakin kecil nilai MAPE maka model yang diperoleh semakin baik, karena makin mendekati nilai aktual.

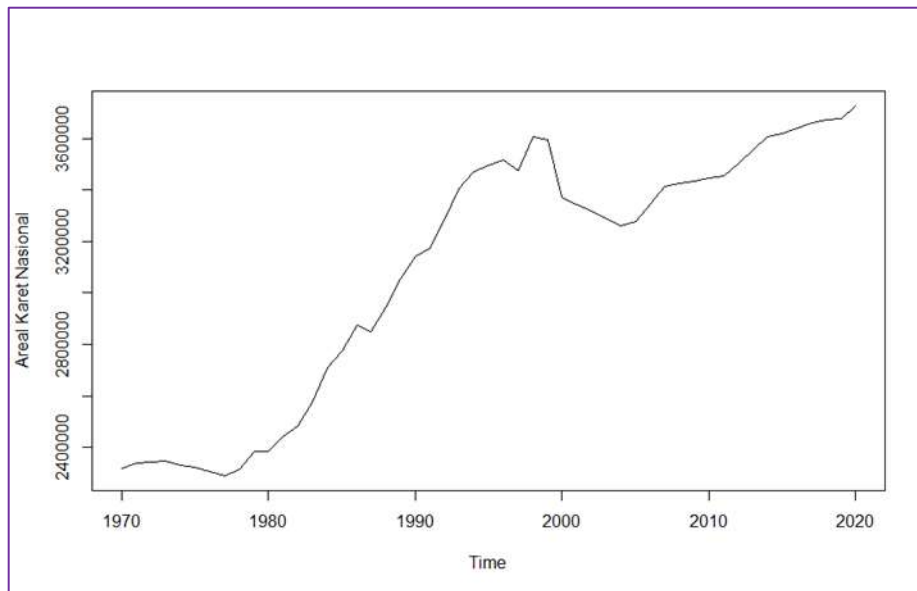
#### e. Sumber Data

Data yang digunakan untuk analisis ini bersumber dari Ditjen. Perkebunan, Kementerian Pertanian, meliputi data : produksi karet dalam bentuk karet kering satuan ton tahun 1971 – 2020, luas areal karet satuan hektar tahun 1970 – 2020, harga riil karet sheet dalam satuan Rp/kg tahun 1971 – 2020, volume ekspor dan impor karet nasional satuan ton tahun 1971 – 2020. Disamping itu juga ada data harga karet dunia yang bersumber dari World Bank, tahun 1970 – 2020.

## HASIL DAN PEMBAHASAN

### A. Model ARIMA

Eksplorasi data luas areal karet nasional dalam bentuk karet kering berupa data tahunan dari tahun 1970 sampai 2020, seperti yang terlihat pada Gambar 3. Pada Gambar 3 terlihat luas areal karet pada tahun 1970 sebesar 2,31 juta hektar, pada awalnya luas areal karet terus naik secara landai sehingga pada tahun 1980 luas areal karet mencapai 2,38 juta hektar. Setelah tahun 1980 pertumbuhan luas areal cenderung lebih cepat, pertumbuhan luas areal karet tahun 1980 – 2000 rata-rata naik sebesar 1,78%/tahun. Pada tahun 1999 sampai 2004 luas areal karet nasional cenderung terus turun, tetapi sejak tahun 2005 sampai 2020 luas areal kembali meningkat. Jika tahun 2005 luas areal karet sebesar 3,28 juta ha, maka tahun 2020 luas areal karet sudah mencapai 3,73 juta ha, atau rata-rata pertumbuhannya 0,78%/tahun atau lebih kecil pertumbuhannya dibandingkan periode sebelumnya. Berdasarkan Gambar 3 juga bisa terlihat bahwa data belum stasioner karena masih mengalami perubahan seiring perubahan waktu.



Gambar 3. Perkembangan Luas Areal Karet Nasional Tahun 1970 - 2020

Dalam melakukan pemodelan luas areal karet menggunakan model Autoregressive Integrated Average (ARIMA), data yang digunakan adalah periode tahun 1970 sampai 2020. Periode data tersebut kemudian dipisahkan menjadi data set training dan testing. Perlunya pemisahan data training dan testing adalah untuk menguji tingkat akurasi dalam melakukan peramalan. Panjang series data pada data set training adalah tahun 1970 sampai 2014, sementara dataset testing adalah periode 2015 sampai 2020 (6 titik). Dataset training digunakan untuk melakukan penyusunan model, sementara dataset testing digunakan untuk validasi model.

Uji kestasioneran data seperti yang disyaratkan apabila melakukan pemodelan ARIMA dilakukan secara visual menggunakan hasil plot data maupun uji formal statistik. Gambar 3 menunjukkan tidak ada fluktuasi yang muncul secara regular setiap bulannya, sehingga dapat

disimpulkan bahwa data belum stasioner atau nilai rata-rata dan varian dari data time series karet mengalami perubahan secara stokastik sepanjang waktu atau sebagian ahli menyatakan rata-rata dan variannya belum konstan (Narchrowi dan Hardius Usman, 2006).

Tabel 1. Hasil Uji Augmented Dickey-Fuller Luas Areal Karet

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-212809  -28723   -7243    27227   150989

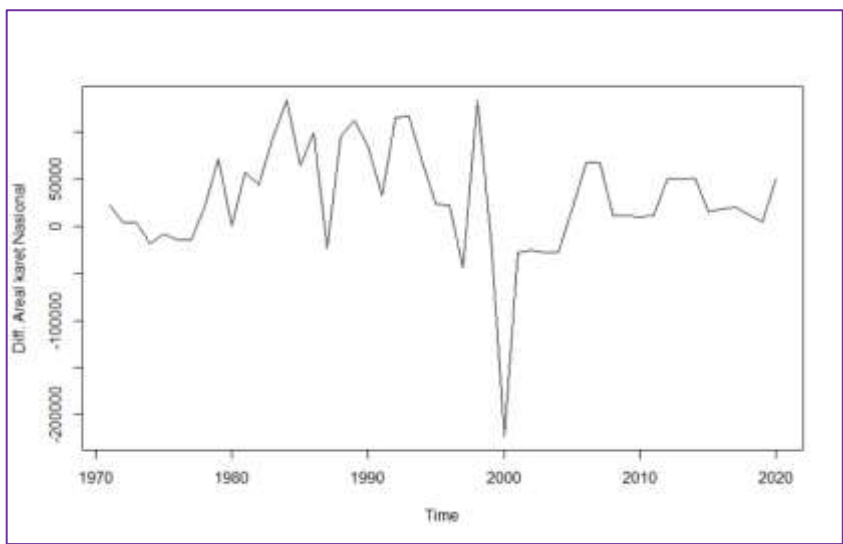
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.570e+05  9.518e+04   1.650  0.10601
z.lag.1      -5.881e-02  4.159e-02  -1.414  0.16422
tt           1.641e+03  1.440e+03   1.140  0.26048
z.diff.lag   3.798e-01  1.395e-01   2.722  0.00919 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56010 on 45 degrees of freedom
Multiple R-squared:  0.1625, Adjusted R-squared:  0.1067
F-statistic: 2.911 on 3 and 45 DF, p-value: 0.04462

Value of test-statistic is: -1.4141 2.1904 1.0815

Critical values for test statistics:
    1pct  5pct 10pct
tau3  -4.04 -3.45 -3.15
phi2   6.50  4.88  4.16
phi3   8.73  6.49  5.47
```

Hal ini diperkuat oleh hasil uji formal statistik yaitu dengan uji Augmented Dickey-Fuller yang mengindikasikan bahwa data luas areal karet adalah belum stasioner, terlihat dari hasil uji tes statistik sebesar  $-1,4141$  sementara nilai kritis pada tingkat kepercayaan  $95\% = -3,45$  (nilai tau3) atau lebih kecil dari nilai uji statistik sehingga sehingga  $H_0$  diterima, atau data luas areal karet belum stationer. Oleh karena itu, selanjutnya data luas areal karet dilakukan pembedaan (differencing) satu kali. Hasil plot setelah dilakukan differencing satu kali seperti pada Gambar 4.



Gambar 4. Data Luas Areal Karet Nasional Setelah Differencing 1

Hasil plot luas areal karet setelah dilakukan differencing 1, menunjukkan bahwa data sudah terlihat stasioner untuk rata-rata. Hal ini juga didukung dengan uji Augmented Dickey-Fuller yang mengindikasikan bahwa data luas areal karet setelah differencing 1 sudah stasioner, terlihat dari hasil uji tes statistik sebesar  $-3,097$  sementara nilai kritis pada tingkat kepercayaan 95%  $= -1,95$  ( $\tau_1$ ) dan tingkat kepercayaan 99%  $= -2,62$  ( $\tau_1$ ) atau lebih besar dari nilai uji statistik sehingga  $H_0$  ditolak, atau data luas areal karet setelah differencing 1 sudah stasioner.

Tabel 2. Hasil Uji Augmented Dickey-Fuller Luas Areal Karet Differencing 1

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-231854  -10054   6094   47298  149056

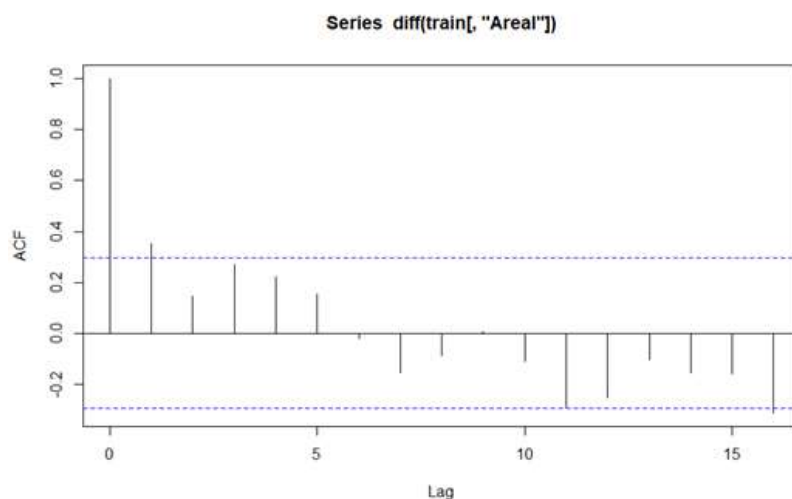
Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1      -0.4694    0.1516  -3.097  0.00333 **
z.diff.lag   -0.1082    0.1474  -0.734  0.46667
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

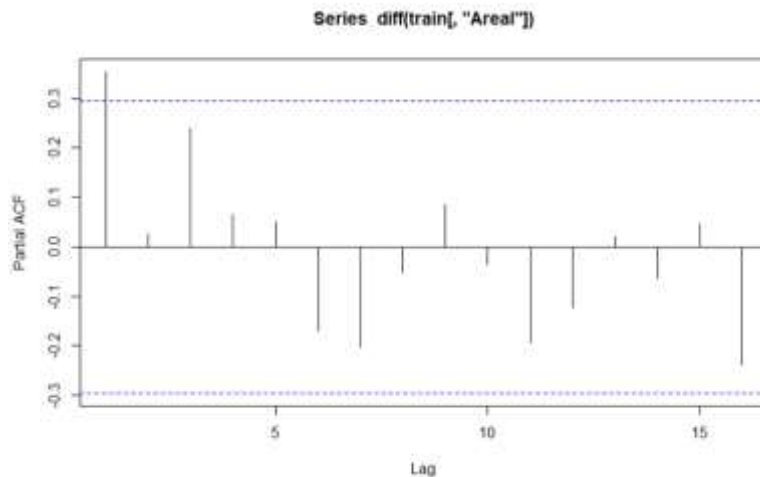
Residual standard error: 58950 on 46 degrees of freedom
Multiple R-squared:  0.2688, Adjusted R-squared:  0.237
F-statistic: 8.454 on 2 and 46 DF, p-value: 0.0007469

value of test-statistic is: -3.097

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

Identifikasi ordo AR dan MA untuk data luas areal karet sebelum ada differencing menunjukkan pola ACF *tail off* sementara pola PACF tidak ada yang signifikan. Hal ini menunjukkan bahwa benar data karet belum signifikan, sehingga perlu dilakukan differencing 1 agar pola ACF dan PACF lebih jelas terlihat.





Gambar 5. Plot ACF dan PACF Luas Areal Karet Differencing 1

Pada Gambar 5 menunjukkan bahwa baik plot ACF maupun plot ACF setelah *Differencing* 1 keduanya sudah ada *cut off*. Identifikasi ordo AR dan MA seperti tersaji pada Gambar 5 menunjukkan pola ACF *cut off* setelah lag 0 dan lag 1, sementara pola PACF menunjukkan *cut off* setelah lag 1 menunjukkan signifikan, sehingga model tentatif belum dapat ditentukan.

Tabel 3. Model Arima Tentatif Berdasarkan Automodel

```

Series: train[, "Areal"]
ARIMA(1,1,0) with drift

Coefficients:
      ar1      drift
      0.3462 29459.34
s.e.    0.1392 13127.99

sigma^2 estimated as 3.478e+09:  log likelihood=-544.81
AIC=1095.61  AICc=1096.21  BIC=1100.97

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 115.5345 56974.76 39974.45 0.0002290835 1.318027 0.7892275 -0.001754478

```

Pengamatan secara visual pada plot ACF dan PACF sulit menentukan orde ARIMA, setelah dilakukan run model dengan menggunakan *auto arima* maka orde ARIMA yang disarankan adalah ARIMA (1,1,0), artinya model ARIMA tentative terbaik untuk melakukan estimasi luas areal karet nasional adalah untuk orde AR nilai  $p=1$ , untuk orde MA nilai  $q=0$ , dan difference  $d=1$ . Berdasarkan Tabel 3 dengan menggunakan ARIMA (1,1,0) maka untuk data training, akan menghasilkan MAPE = 1,32% artinya data berdasarkan model arima akan menyimpang hasil estimasi rata-rata sekitar -1,32% sampai +1,32% dari data aktual.

Disamping metode pemilihan model Arima berdasarkan automodel, digunakan juga metode lain untuk mendapatkan orde ARIMA terbaik, yaitu dengan metode *Arima selection*. Pada metode ini akan dikeluarkan beberapa model terbaik. Setelah dilakukan pemilihan model metode terbaik tetap pada *differencing 1*. Model terbaik pertama menurut metode ini adalah ARIMA (1,1,0) menghasilkan nilai sbc paling kecil yaitu sebesar 970,43.

Tabel 4. Model Arima Tentatif Berdasarkan Arima Selection Differencing 1

	p	q	sbc
[1,]	1	0	970.4332
[2,]	0	0	971.5008
[3,]	2	0	975.0801
[4,]	3	0	977.1846
[5,]	4	0	981.3103
[6,]	1	1	984.4253
[7,]	1	2	984.4794
[8,]	0	1	984.9775
[9,]	1	3	985.5059
[10,]	0	2	985.7499

Selanjutnya dilakukan pengujian model ARIMA (1,1,0) apakah koefisien sudah signifikan dan bagaimana perbandingan data training dan data testing. Untuk model ARIMA (1,1,0) koefisien ar1 sebesar 0,4625 dan koefisien ini signifikan pada taraf alpha 0,1%. Sehingga model ARIMA (1,1,0) layak digunakan (Tabel 5).

Tabel 5. Uji Koefisien Model Arima (1,1,0)

```
Call:
arima(x = train[, "Areal"], order = c(1, 1, 0))

Coefficients:
      ar1
    0.4625
s.e.    0.1318

sigma^2 estimated as 3.626e+09:  log likelihood = -546.81,  aic = 1097.61
z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  0.46255    0.13175   3.5108 0.0004468 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tabel 6. Perbandingan MAPE untuk ARIMA (1,1,0)

```
Start = 46
End = 51
Frequency = 1
[1] 3629510 3640272 3645249 3647552 3648617 3649109
> accuracy(ramalan_arima, test[, "Areal"])
Training set  ME      RMSE      MAE      MPE      MAPE      MASE
Test set     22087.65 35408.22 25298.31 0.5959465 0.6845563 0.4994721
```

Setelah dilakukan pengujian koefisien model untuk ARIMA (1,1,0) ternyata komponen ar1 signifikan, maka dilakukan pengujian MAPE untuk data training dan data testing. Hasil menunjukkan jika menggunakan model ARIMA (1,1,0) akan menghasilkan data training sebesar 1,35%. Setelah dilakukan pengujian dengan cara meramal 6 tahun kedepan yaitu tahun 2015 – 2020, maka hasil ramalan atau data testing menghasilkan MAPE 0,68%. Hal ini

menunjukkan bahwa model ARIMA (1,1,0) jika digunakan untuk peramalan maka rata-rata hasil ramalan hanya menyimpang 0,68%.

Selanjutnya dilakukan pengujian untuk order ARIMA lainnya seperti yang terlihat pada Tabel 4. Tujuan pengujian ini adalah untuk melihat signifikansi koefisien dari masing-masing model ARIMA dan untuk mencari MAPE yang terkecil untuk data testing dan data training. Hasil pengujian tersaji pada Tabel 7.

Tabel 7. Pengujian Signifikansi Koefisien dan MAPE untuk Model ARIMA

Model	Signifikansi	MAPE Training	MAPE Testing
ARIMA (1,1,0)	Ar1 signifikan	1,352	0,684
ARIMA (2,1,0)	Ar1 signifikan Ar2 tidak signifikan	1,318	0,449
ARIMA (1,1,1)	Ar 1 signifikan Ma1 signifikan	1,275	0,613
ARIMA (1,1,2)	Ar 1 signifikan Ma1 signifikan Ma2 tidak signifikan	1,276	0,499
ARIMA (0,0,1)	Ma1 signifikan	1,457	1,179
ARIMA (1,1,3)	Ar1 signifikan Ma1 signifikan Ma2 tidak signifikan Ma3 signifikan	1,326	0,900
ARIMA (0,1,2)	Ma1 signifikan Ma2 tidak signifikan	1,430	1,099

Berdasarkan hasil pengujian pada Tabel 7 maka model tentative terbaik untuk peramalan Luas Areal Karet adalah ARIMA (1,1,1) dengan komponen ar1 signifikan, komponen ma1 signifikan. Hasil pengujian keakuratan model dalam melakukan peramalan, menghasilkan MAPE data training terkecil yaitu sebesar 1,275% dan MAPE data testing sebesar 0,613%. Model tentative lain yang perlu dipertimbangkan adalah ARIMA (2,1,0) karena menghasilkan MAPE data training sebesar 1,318% dan MAPE data testing 0,449%.

Tabel 8. Pengujian Koefisien Model ARIMA (1,1,1)

```
Call:
arima(x = train[, "Areal"], order = c(1, 1, 1))

Coefficients:
      ar1      ma1
 0.8574 -0.5624
s.e. 0.1200 0.1954

sigma^2 estimated as 3.415e+09:  log likelihood = -545.56,  aic = 1097.12
> library(lmtest)
> coeftest(model1)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  0.85742    0.11996   7.1478 8.816e-13 ***
ma1 -0.56244    0.19538  -2.8787 0.003993 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Untuk model ARIMA (1,1,1) koefisien  $ar_1 = 0,857$  hasil pengujian dengan z test signifikan pada tingkat 99,9%, dan koefisien  $ma_1 = -0,562$  hasil pengujian menunjukkan bahwa koefisien signifikan pada tingkat kepercayaan 99%, seperti pada Tabel 8.

Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan Ljung-Box. Hasil pengujian nilai p-value pada lag 5 sampai dengan lag 30 tidak ada yang signifikan yang ditunjukkan dengan nilai p-value tidak ada yang lebih kecil dari nilai 0.05 (kepercayaan 95%), sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (1,1,1) layak digunakan.

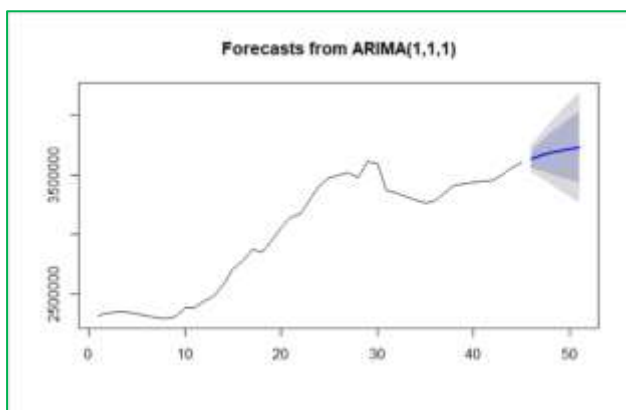
Tabel 9. Uji Sisaan Ljung Box Model Arima (1,1,1)

lags	statistic	df	p-value
5	3.553234	5	0.6153484
10	5.617459	10	0.8463140
15	11.585274	15	0.7101103
20	18.455116	20	0.5574546
25	20.685755	25	0.7099944
30	22.337020	30	0.8415053

Selanjutnya dilakukan uji kemampuan model ARIMA (1,1,1) apakah memiliki akurasi yang tinggi dalam melakukan peramalan. Untuk itu dilakukan uji coba peramalan dengan menggunakan data testing, yaitu luas areal karet tahun 2015 – 2020. Sementara data training digunakan untuk menyusun model ARIMA (1,1,1). Dari hasil pengujian pada Tabel 10, menunjukkan MAPE untuk data training sebesar 1,27%, sementara MAPE data testing 0,61%. Hal ini menunjukkan Model Arima (1,1,1) sudah cukup akurat dalam melakukan peramalan karena hasil peramalan rata-rata menyimpang tidak lebih dari 1%.

Tabel 10. Hasil Pengujian Data Training dan Testing Arima (1,1,1)

```
Time Series:
Start = 46
End = 51
Frequency = 1
[1] 3635660 3660881 3682506 3701047 3716945 3730577
> accuracy(ramalan_arima,test[, "Areal"])
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 10680.95 57786.55 39012.13 0.3775327 1.2749226 0.7702283
Test set    -22463.50 25198.21 22463.50 -0.6134803 0.6134803 0.4435037
```



Gambar 6. Plot Hasil Peramalan Data Testing untuk ARIMA (1,1,1)



Selanjutnya dilakukan pengepasan model untuk seluruh data. Untuk Model ARIMA (1,1,1) koefisien ar1 sebesar 0,862 dan koefisien ma1= -0,572. Jika melakukan run model ARIMA (1,1,1) untuk seluruh data yaitu dari tahun 1967 – 2020 maka akan dihasilkan MAPE sebesar 1,16%. Hal ini menunjukkan bahwa antara data estimasi dengan data actual akan berbeda rata-rata berkisar antara -1,16% sampai +1,16%. Untuk metode estimasi dengan bias masih dibawah 2% dianggap masih cukup baik dan akurat.

Tabel 8. Model Arima (1,1,1) untuk Seluruh Data

```

Series: karet[, "Areal"]
ARIMA(1,1,1)

Coefficients:
      ar1      ma1
    0.8620 -0.5718
s.e.  0.1093  0.1805

sigma^2 estimated as 3.173e+09:  log likelihood=-617.07
AIC=1240.14  AICC=1240.66  BIC=1245.88

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set 9662.453 54646.04 35846.81 0.3391246 1.163567 0.7631749

```

Dengan menggunakan model ARIMA (1,1,1) menghasilkan angka estimasi luas areal karet untuk 5 tahun ke depan. Hasil Estimasi dengan model ARIMA ini pada tahun 2021 luas areal karet nasional sebesar 3,745 juta ha. Pada tahun 2022 luas areal karet diestimasi akan naik sebesar 0,45% menjadi 3,762 juta. Pada tahun 2023 sampai 2025 juga menunjukkan luas areal karet nasional mengalami kenaikan secara perlahan. Model Arima rata-rata menghasilkan angka estimasi yang cenderung terus naik atau terus turun. Estimasi rata-rata pertumbuhan luas areal karet tahun 2021 – 2025 rata-rata sebesar 0,36% per tahun. Jika dibandingkan pertumbuhan luas areal karet selama 5 tahun terakhir (tahun 2015 -2020) dengan menggunakan Angka Tetap rata-rata sebesar 0,59% per tahun, sementara hasil estimasi lima tahun kedepan rata-rata pertumbuhan hanya 0,36% per tahun atau lebih rendah dari data historisnya. Hal ini terjadi karena beberapa tahun terakhir harga karet dunia terus turun, sehingga perluasan areal sangat kecil pertumbuhannya, jika harga karet dunia meningkat maka pertumbuhan luas areal karet diduga akan lebih besar.

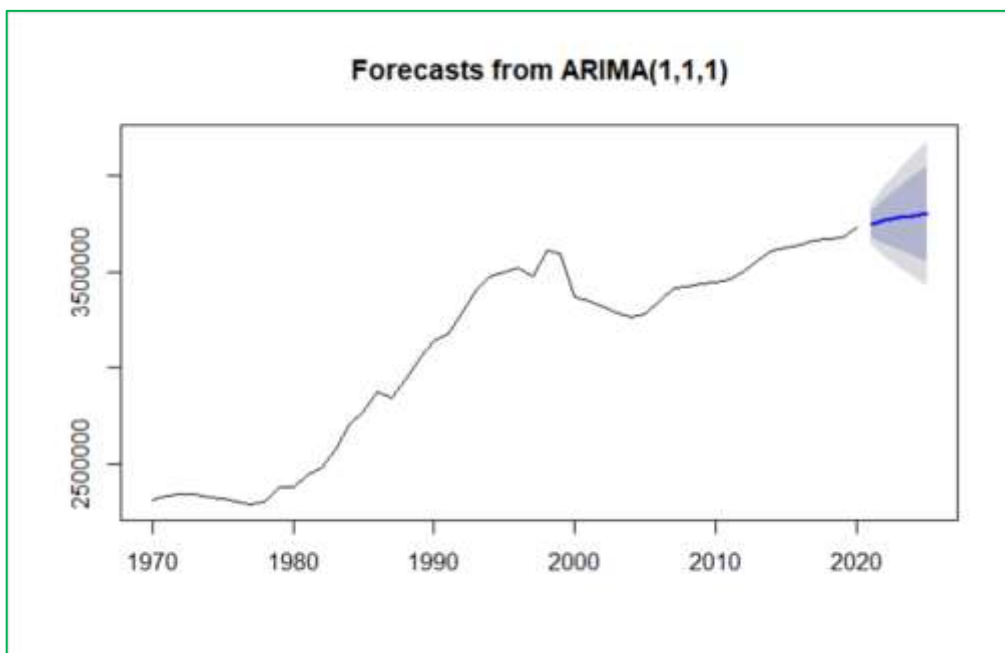
Tabel 9. Ouput Peramalan Model Arima (1,1,1) untuk Luas Areal Karet

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	3745592	3673405	3817779	3635192	3855993
2022	3762331	3644497	3880164	3582120	3942542
2023	3776758	3614747	3938770	3528984	4024533
2024	3789194	3583493	3994896	3474601	4103788
2025	3799914	3550998	4048830	3419229	4180598

Tabel 10. Hasil Estimasi Luas Areal Karet dengan Model ARIMA (1,1,1)

	Model ARIMA		
	Pengujian MAPE	ARIMA (1,1,1)	(%)
	MAPE Training	1.275	
	MAPE Testing	0.613	
ATAP	2016	3,639,048	
	2017	3,659,090	0.55
	2018	3,671,387	0.34
	2019	3,676,035	0.13
	2020	3,726,173	1.36
Angka Estimasi (AESTI)	2021	3,745,592	0.52
	2022	3,762,331	0.45
	2023	3,776,758	0.38
	2024	3,789,194	0.33
	2025	3,799,914	0.28
Rata-rata Pertumbuhan	ATAP 2016 - 2020		0.59
	AESTI 2021 - 2025		0.36

Pada Tabel 10, menunjukkan bahwa luas areal karet selama 5 tahun terakhir pertumbuhan masih positif, dengan kisaran pertumbuhan 0,13% sampai 1,36%. Hasil estimasi dengan Model ARIMA (1,1,1) pada tahun 2021 – 2025 luas areal karet masih tumbuh dengan kisaran 0,28% sampai 0,52%. Estimasi luas areal Karet nasional pada tahun 2022 sebesar 3.762.331 hektar.



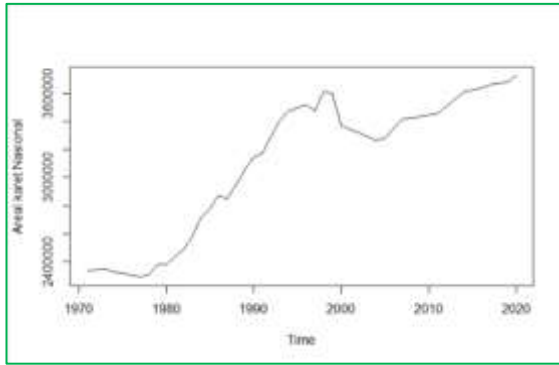
Gambar 6. Hasil Estimasi Luas Areal Karet Tahun 2021 – 2025 Model Arima (1,1,1)

## B. Model Fungsi Transfer

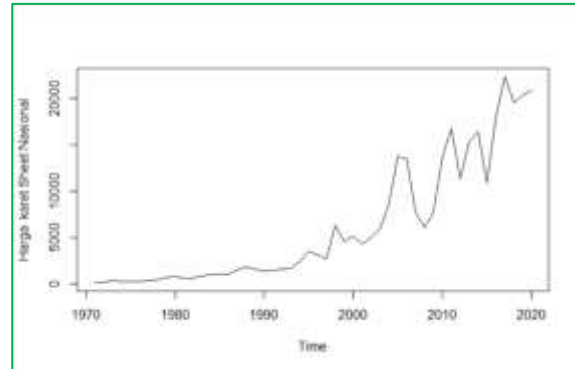
Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (*Autoregressive Integrated Moving Average*). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variabel lain sebagaimana pada model regresi. Model fungsi transfer merupakan suatu model yang menggambarkan nilai prediksi dari suatu time series (deret output atau  $Y_t$ ) berdasarkan nilai-nilai dari deret itu sendiri ( $Y_t$ ) dan berdasarkan pula pada data time series yang mempunyai hubungan (deret input atau  $X_t$ ) dengan deret output. Model fungsi transfer yang bersifat dinamis berpengaruh tidak hanya pada hubungan linier antara waktu ke- $t$  input  $X_t$  dan waktu ke- $t$  output  $Y_t$ , tetapi juga saat input  $X_t$  dengan saat  $t, t+1, \dots, t+k$  pada output  $Y_t$  (Fitriani et al., 2013). Dalam melakukan estimasi dengan fungsi transfer diperlukan dua variabel, yaitu output  $Y_t$  dan input  $X_t$ . Input  $X_t$  merupakan variabel yang diduga paling berpengaruh terhadap output  $Y_t$ . Pada analisis ini dicontohkan estimasi untuk luas areal karet (output  $Y_t$ ) yang diduga dipengaruhi oleh harga karet dunia.

Dalam bahasan ini, model fungsi transfer variabel output luas areal karet akan melibatkan variabel harga karet sheet nasional sebagai variabel input. Langkah awal dalam mencari model fungsi transfer menggunakan software RStudio adalah penyiapan data, baik output  $Y_t$  (luas areal karet) maupun input  $X_t$  (harga karet sheet). Luas areal karet dalam satuan hektar bersumber dari Direktorat Jenderal Perkebunan, Kementerian Pertanian, sedangkan harga karet sheet dalam satuan Rp/Kg bersumber dari Ditjen Perkebunan. Banyaknya series data yang digunakan baik untuk luas areal maupun harga karet dunia harus seragam. Pada bahasan ini, series data yang digunakan yaitu 1971-2020 atau sebanyak 50 observasi.

Eksplorasi data dilakukan dengan menampilkan plot data luas areal maupun harga karet dunia. Berdasarkan plot data dapat diketahui pola data series 50 tahun yang akan digunakan untuk pemodelan. Berdasarkan Gambar 10 dan Gambar 11, terlihat bahwa terdapat data luas areal karet nasional memiliki tren meningkat dari tahun ke tahun sedangkan harga karet sheet cenderung berfluktuasi. Harga karet sheet cenderung turun pada sekitar 10 tahun terakhir, akibat persaingan harga dengan produk karet sintetis. Luas areal karet nasional maupun harga karet sheet terindikasi tidak stasioner berdasarkan plotnya.



Gambar 10. Plot Data Luas areal Karet, 1971-2020



Gambar 11. Plot Data Harga Sheet, 1971-2020

Tahapan penyusunan model Fungsi Transfer luas areal karet dengan variabel input harga karet dunia adalah sebagai berikut :

- a. Pembagian series data awal menjadi series data *training* dan *testing*
- b. Pemeriksaan kestasioneran
- c. Pencarian model tentatif untuk variabel input
- d. *Prewhitening* dan korelasi silang
- e. Identifikasi model noise
- f. Pengepasan model
- g. Peramalan berbasis fungsi transfer

Data luas areal dan harga karet tahun 1971-2020 sebanyak 50 series akan dibagi menjadi series data *training* untuk periode 1971-2014 dan series data *testing* untuk periode 2015-2020. Selanjutnya dilakukan uji kestasioneran data untuk data input  $X_t$  yaitu harga karet dunia menggunakan uji Augmented Dickey-Fuller (ADF).

Hipotesis pada uji ADF ini adalah:

- $H_0$  : data tidak stasioner  
 $H_1$  : data stasioner

Nilai test-statistik yang lebih besar dari critical values (nilai tau3), baik untuk taraf 1%, 5% maupun 10% menunjukkan bahwa  $H_0$  gagal ditolak, atau series data harga karet dunia belum stasioner. Oleh karena itu akan dilakukan pembedaan/differencing satu kali dan kemudian dilakukan uji ADF. Output hasil uji Augmented Dickey-Fuller sebelum dilakukan *diffrencing* menunjukkan belum stationer. Setelah dilakukan Differencing 1 hasil uji menunjukkan sudah stationer seperti pada Tabel 16.

Tabel 16. Hasil uji Kestationeran untuk Harga Karet Sheet Setelah Differencing 1

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-3854.6  -72.6   83.4   485.7  5462.6

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.3117     0.2143  -6.12 3.53e-07 ***
z.diff.lag    0.3692     0.1571   2.35 0.0239 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

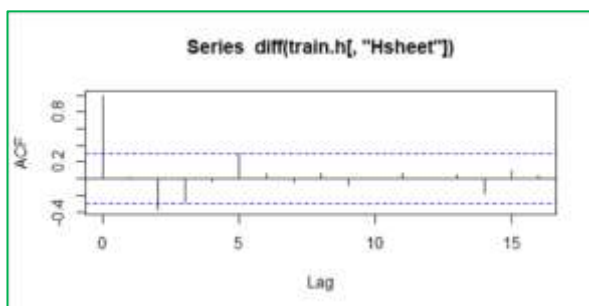
Residual standard error: 2046 on 39 degrees of freedom
Multiple R-squared:  0.5399,    Adjusted R-squared:  0.5163
F-statistic: 22.88 on 2 and 39 DF,  p-value: 2.669e-07

Value of test-statistic is: -6.1197

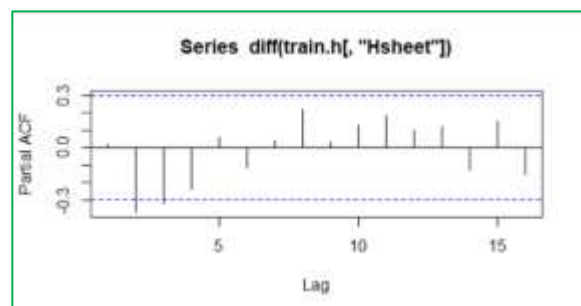
Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

Uji ADF pada data yang telah dilakukan *differencing* satu kali menunjukkan bahwa nilai *test-statistic* yaitu -6,1197 lebih kecil dari *critical values* (nilai tau1) baik 1% maupun 5% menunjukkan bahwa  $H_0$  ditolak yang berarti data harga karet sheet nasional sudah stasioner.

Stasioneritas data dapat juga dilihat dari plot ACF dan PACF. Jika lag pada ACF maupun PACF menunjukkan *tail off* maka data tidak stasioner. Berikut adalah output untuk menunjukkan plot ACF dan PACF harga karet dunia setelah differencing 1.



Gambar 12. Plot ACF Harga Karet Sheet dengan *Differencing* satu kali



Gambar 13. Plot PACF Harga Karet Sheet dengan *Differencing* satu kali

Plot ACF cenderung cut off dimana hanya lag 0 yang keluar dari *confidence interval*. Plot PACF menunjukkan cut off pada lag 2, ditunjukkan berada di luar garis *confidence interval*. Hal tersebut memperkuat hasil uji ADF yaitu data harga karet sheet telah stasioner.

Selain itu, plot ACF dan PACF juga dapat digunakan untuk mengidentifikasi ordo AR dan MA, dengan cara melihat lag yang nyata atau keluar dari garis *confidence interval*. Namun

demikian, pada plot ACF diatas, terlihat nyata pada lag=0 yang sering terjadi pada data riil, sehingga sulit untuk mengidentifikasi ordo AR dan MA yang tepat melalui plot ACF dan PACF.

Pencarian model tentatif variabel input harga karet sheet dilakukan melalui penelusuran menggunakan model ARIMA. Model terbaik dapat dipilih menggunakan script *auto.arima* yang tersedia pada RStudio. Data yang digunakan untuk memilih model terbaik adalah series data *training*.

Tabel 17. Model Auto Arima Harga Karet Sheet Lokal

```
Series: train.h[, "Hsheet"]
ARIMA(0,1,0)

sigma^2 estimated as 4347036:  log likelihood=-389.64
AIC=781.28  AICc=781.38  BIC=783.05

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 369.4796 2061.126 1153.525 6.459251 23.65776 0.9772747 0.01742706
```

Berdasarkan pemilihan orde ARIMA menggunakan *automodel* menyarankan bahwa model terbaik untuk harga karet sheet adalah ARIMA (0,1,0) dengan MAPE data training yang masih sangat tinggi yaitu sebesar 23,7%. Model ARIMA (0,1,0) untuk data harga karet sheet tersebut menghasilkan nilai AIC = 781,28 sebagaimana output yang disajikan pada Tabel 17. Oleh karena model tentatif ARIMA (0,1,0) maka tidak dapat dilakukan pengujian koefisien ar maupun ma karena order AR dan order MA nilainya 0. Hasil estimasi jika menggunakan ARIMA (0,1,0) akan cenderung konstan, karena yang berpengaruh hanya differencing. Oleh karena model ARIMA (0,1,0) hasil permalan kurang memuaskan maka perlu dicari model tentative lainnya. Pada Tabel 18 output yang dihasilkan untuk memilih model tentative terbaik untuk model ARIMA faktor input  $X_t$  yaitu harga karet sheet nasional, dengan *option arima selection*.

Tabel 19. Model Arima Tentatif untuk Harga Karet Sheet (Differencing 1)

```

p q      sbc
[1,] 0 0 656.8293
[2,] 3 0 659.3247
[3,] 2 0 659.4997
[4,] 1 0 661.5872
[5,] 4 0 661.7675
[6,] 5 0 665.7938
[7,] 0 3 670.5481
[8,] 0 2 671.3443
[9,] 1 2 672.7522
[10,] 5 2 673.8003
```

Hasil otput R-Studio akan menunjukkan sepuluh model tentatif dimana idealnya model terbaik adalah model yang memiliki nilai SBC terkecil. Model ARIMA yang direkomendasikan ditunjukkan dari nilai p,d,q. Sebagai contoh model pertama dengan nilai p=0 dan q=0. Karena data harga karet sheet telah dilakukan *differencing* satu kali berarti d=1, artinya model yang

direkomendasikan adalah ARIMA (0,1,0). Model tentative lainnya adalah ARIMA (1,1,0), untuk mengetahui apakah model ARIMA (1,1,0) lebih baik dari hasil *auto.arima* yaitu ARIMA (0,1,2) maka dapat dibandingkan dari nilai AIC dan koefisien ar1.

Selanjutnya dilakukan pemilihan model ARIMA terbaik untuk faktor input Harga Karet Sheet. Pengujian dilakukan dengan memperhatikan tingkat signifikansi koefisien ar atau ma dan pengujian dengan menggunakan data training dan data testing untuk peubah harga karet sheet nasional. Hasil pengujian model ARIMA dengan differencing 1 menunjukkan bahwa koefisien AR pada umumnya signifikan pada ar2, dan koefisien ma signifikan pada ma2. Pada pengujian data training menunjukkan data training relative masih cukup besar dengan kisaran MAPE sebesar 20,9 sampai 23,7. Sementara untuk MAPE data testing pada kisaran yang lebih besar yaitu antara 17,1 sampai 28,9. Jika dilihat dari perbandingan MAPE maka model ARIMA terbaik untuk Harga Karet Sheet adalah ARIMA (5,1,2) dengan nilai MAPE training dan Testing yang paling kecil yaitu masing-masing 20,9 dan 17,1. Hasil pengujian model ATIMA tenataif pada differencing1 terlihat pada Tabel 19.

Tabel 19. Model Tentatif Harga Karet Sheet pada Diiferencing 1

<b>Model ARIMA</b>	<b>Signifikansi</b>	<b>MAPE Training</b>	<b>MAPE Testing</b>
<b>ARIMA (3,1,0)</b>	ar2 signifikan	22,06	28,64
<b>ARIMA (2,1,0)</b>	ar2 signifikan	22,79	27,88
<b>ARIMA (1,1,0)</b>	Tidak ada	23,73	24,03
<b>ARIMA (4,1,0)</b>	ar2 signifikan	22,46	28,37
<b>ARIMA (5,1,0)</b>	Tidak ada	21,75	23,27
<b>ARIMA (0,1,3)</b>	ma2 signifikan	22,85	28,84
<b>ARIMA (0,1,2)</b>	ma2 signifikan	22,88	28,92
<b>ARIMA (1,1,2)</b>	ma2 signifikan	22,86	27,88
<b>ARIMA (5,1,2)</b>	ar5 signifikan	20,89	17,05

Hasil pengujian model terbaik untuk harga karet sheet, pada differencing 1 masih menghasilkan MAPE data training dan testing yang masih cukup besar, maka dilakukan pengujian kembali untuk diiferencing 2. Hasil ouput model tentative seperti pada Tabel 20.

Tabel 20. Model Arima Tentatif untuk Harga Karet Sheet (Differencing 2)

	p	q	sbc
[1,]	4	2	654.7657
[2,]	4	0	655.3789
[3,]	5	2	658.0513
[4,]	4	3	659.2606
[5,]	4	4	659.8181
[6,]	5	0	659.9451
[7,]	5	3	660.4148
[8,]	4	1	661.9082
[9,]	4	5	662.3915
[10,]	5	4	662.4589

Selanjutnya dilakukan pemilihan model ARIMA terbaik untuk faktor input Harga Karet Sheet pada differencing 2. Hasil pengujian model ARIMA dengan differencing 2 menunjukkan bahwa koefisien AR pada umumnya signifikan pada ar1 sampai ar4, dan koefisien ma signifikan pada ma1 sampai ma3. Pada pengujian data training menunjukkan data training relative lebih baik dibandingkan differencing 1 dengan kisaran MAPE sebesar 16,8 sampai 23,2. Sementara untuk MAPE data testing pada kisaran yang lebih besar yaitu antara 14,1 sampai 33,6. Jika dilihat dari perbandingan MAPE maka model ARIMA terbaik untuk Harga Karet Sheet adalah ARIMA (4,2,1) dengan nilai MAPE Training dan Testing yang paling kecil yaitu masing-masing 19,59 dan 14,1. Hasil pengujian model ARIMA tentatif pada differencing 2 terlihat pada Tabel 21.

Tabel 21. Model Tentatif Harga Karet Sheet pada Diiferencing 2

Model ARIMA	Signifikansi	MAPE Training	MAPE Testing
<b>ARIMA (4,2,2)</b>	ar1 ar2 ar3 ar4 signifikan	20,05	15,86
<b>ARIMA (4,2,0)</b>	ar1 ar2 ar3 ar4 signifikan	23,21	33,59
<b>ARIMA (5,2,2)</b>	ar2 ar3 ar4 signifikan	20,02	15,69
<b>ARIMA (4,2,3)</b>	ar1 ar3 ma1 ma2 ma3 signifikan	17,62	16,57
<b>ARIMA (4,2,4)</b>	ar1 ar2 ar4 ma1 ma2 ma3 ma4 signifikan	16,80	17,94
<b>ARIMA (5,2,0)</b>	ar1 ar2 ar3 ar4 signifikan	23,09	32,17
<b>ARIMA (5,2,3)</b>	ar1 ar2 ar3 ar4 ma1 ma3 signifikan	18,98	23,19
<b>ARIMA (4,2,1)</b>	ar1 ar2 ar3 ar4 ma1 signifikan	19,59	14,08
<b>ARIMA (4,2,5)</b>	ar1 ar2 ar3 ma1 ma2 signifikan	17,25	17,23



Tabel 22. Pengujian Koefisien Model Arima (4,2,1) Harga Karet Sheet

```
Call:
arima(x = train.h[, "Hsheet"], order = c(4, 2, 1))

Coefficients:
      ar1      ar2      ar3      ar4      ma1
-0.2805  -0.5250  -0.3779  -0.3748  -0.8632
s.e.    0.1554   0.1572   0.1645   0.1881   0.0806

sigma^2 estimated as 3044781:  log likelihood = -375.13,  aic = 762.26
> library(lmtest)
> coeftest(model1)

z test of coefficients:

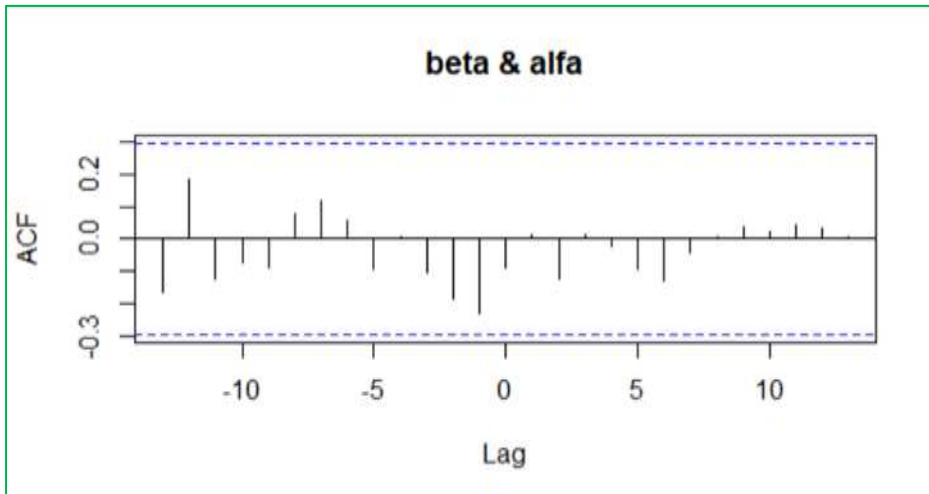
      Estimate Std. Error z value Pr(>|z|)
ar1 -0.280545   0.155450  -1.8047 0.0711165 .
ar2 -0.525001   0.157219  -3.3393 0.0008399 ***
ar3 -0.377937   0.164519  -2.2972 0.0216063 *
ar4 -0.374789   0.188052  -1.9930 0.0462599 *
ma1 -0.863243   0.080552 -10.7166 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model ARIMA (4,2,1) menghasilkan nilai AIC = 762,26 relatif kecil dibandingkan model ARIMA pada order lainnya. Berdasarkan Tabel 22, model ARIMA (4,2,1) menunjukkan bahwa ar1 signifikan pada taraf 10%, ar2 dan ma1 signifikan pada taraf 1%, ar3 dan ar4 signifikan pada taraf 5%. Hal ini menunjukkan semua komponen pada model ARIMA (4,2,1) signifikan. Oleh karena ARIMA (4,2,1) memiliki MAPE training dan MAPE Testing yang terkecil seperti pada Tabel 23, maka selanjutnya yang digunakan untuk peramalan harga karet sheet lokal sebagai faktor input untuk meramalkan luas areal karet nasional menggunakan ARIMA (4,2,1).

Tabel 23. Pengujian MAPE untuk Faktor Input Harga Karet Sheet

```
Time Series:
Start = 45
End = 50
Frequency = 1
[1] 16907.37 18786.07 18120.88 18766.69 20056.09 20939.13
> accuracy(ramalan_arima,test.h[, "Hsheet"])
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 318.6544 1704.811  968.7743  2.356552 19.58877 0.8207526
Test set    -238.7062 3048.535 2026.1498 -5.864928 14.08351 1.7165687
      ACF1
Training set -0.05524081
Test set      NA
```

Tahap selanjutnya untuk penyusunan model fungsi transfer ini adalah *prewhitening* dan korelasi silang. Korelasi silang menggambarkan struktur hubungan antara  $X_t$  dengan  $Y_t$ . Untuk mengidentifikasi pengaruh  $X_t$  terhadap  $Y_t$  maka deret  $X_t$  harus stasioner atau sudah distasionerkan. Dalam konteks pemodelan  $X_t$  terhadap  $Y_t$ , untuk membuat  $X_t$  stasioner tidak dengan pembedaan (*differencing*) namun dengan mengambil komponen *white noise* dari  $X_t$  (*prewhitening*). *Prewhitening* dilakukan terhadap deret input  $X_t$  yang didefinisikan sebagai alfa serta deret input  $Y_t$  yang didefinisikan sebagai beta. Hasil output untuk *prewhitening* dan korelasi silang berupa grafik ACF untuk beta dan alfa.



Gambar 14. Plot korelasi silang Luas Areal Karet dengan Harga Karet Sheet Lokal

Hasil plot korelasi silang digunakan untuk mengidentifikasi ordo  $r$ ,  $s$ , dan  $b$ . Ordo  $r$  adalah panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$ , ordo  $s$  adalah panjang lag  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$ , dan ordo  $b$  adalah panjang jeda pengaruh  $X_t$  terhadap  $Y_t$ . Identifikasi ordo  $r$ ,  $s$  dan  $b$  hanya dilihat pada lag yang positif.

Plot korelasi silang diatas menunjukkan bahwa tidak ada yang signifikan ditunjukkan dengan tidak ada yang keluar dari garis signifikansi, maka nilai  $b=0$ . Kemudian, tidak ada tambahan lagi lag yang signifikan maka nilai  $s=0$ . Mengingat data luas areal karet dan harga karet sheet merupakan data tahunan yang tidak mengandung musiman maka diasumsikan nilai  $r=0$ . Nilai  $b=0$  menunjukkan tidak ada jeda pengaruh antara harga karet sheet pada waktu  $t$  terhadap luas areal karet pada waktu  $t$ . Nilai  $s=0$  berarti ada korelasi antara luas areal dan harga karet sheet pada tahun yang sama. Dengan kata lain, dampak dari harga karet sheet terhadap luas areal dirasakan pada waktu yang sama ( $t$ ).

Hasil pengujian dengan Identifikasi Awal nilai  $r=0$   $s=0$  dan  $b=0$ , menghasilkan nilai MAPE yang cukup baik yaitu sebesar 10,82. Hal ini menunjukkan bahwa dengan ARIMA (0,0,0) sudah cukup baik untuk meramalakan luas areal karet, seperti pada Tabel 23.

Tabel 23. Fungsi Transfer Identifikasi Awal ( $r = 0$ ,  $s=0$ , dan  $b=0$ ) ARIMA (0,0,0)

```
Series: train.h[, "Areal"]
Regression with ARIMA(0,0,0) errors

Coefficients:
  intercept      xreg
 2751094      63.5437
s.e.          71386    10.7948

sigma^2 estimated as 1.318e+11:  log likelihood=-624.7
AIC=1255.41  AICC=1256.01  BIC=1260.76

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
Training set -1.270625e-10  354637.3  314546.4 -1.559505  10.82309  6.130481
              ACF1
Training set  0.8964068
```

Untuk menghasilkan orde yang paling tepat untuk menentukan orde Arima fungsi transfer dengan melakukan identifikasi model noise. Untuk menghasilkan model terbaik dengan menggunakan *auto-arima* pada R Studio, model maka noise yang disarankan adalah ARIMA (0,1,0). Tahap selanjutnya dilakukan pengepasan model, dengan model noise. Hasil pengujian fungsi transfer dengan nilai  $r=0$ ,  $s=0$ , dan  $b=0$  dengan model noise ARIMA (0,1,0) menghasilkan nilai MAPE yang cukup besar yaitu 115,12%, sehingga perlu dicari model alternatif lain. Pemilihan model noise tentative, terlihat seperti Tabel 25.

Tabel 24. Pengujian Orde  $r=0$ ,  $s=0$ , dan  $b=0$  ARIMA Noise (0,1,0) Luas areal Karet

```
Series: res
ARIMA(0,1,0)

sigma^2 estimated as 2.181e+10: log likelihood=-572.83
AIC=1147.67  AICC=1147.77  BIC=1149.43

Training set error measures:
Training set  ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
5299.447  145989.7  99786.78  21.37614  115.1229  0.9773658  0.1097942
```

Tabel 25. Model Residual/Noise Tentatif ARIMA Untuk Luas areal Karet

```
      p q      sbc
[1,] 0 0 1024.593
[2,] 1 0 1028.868
[3,] 2 0 1031.286
[4,] 3 0 1035.086
[5,] 5 0 1035.313
[6,] 4 0 1039.907
[7,] 5 1 1042.181
[8,] 5 2 1043.099
[9,] 5 5 1045.083
[10,] 0 1 1047.030
```

Oleh karena model *autoarima* disarankan differencing tingkat 1, maka solusinya akan dicari model alternative noise pada differencing 1. Model alternatif yang diberikan untuk *model noise* seperti Tabel 25. Model tentatif pertama yang disarankan untuk data residual adalah ARIMA (0,1,0). Selanjutnya model tersebut didefinisikan sebagai *modelres* atau *model residual* dan dilihat signifikansi AR. Oleh karena model noise ARIMA (0,1,0) sudah diujicobakan dan menghasilkan MAPE yang tinggi, maka selanjutnya dilakukan uji untuk model noise untuk residual dengan Arima (1,1,0), Arima (2,1,0), dan Arima (0,1,1). Untuk model noise ARIMA (1,1,0) menghasilkan komponen ar1 yang tidak signifikan (Tabel 23).

Tabel 26. Pemilihan Model Arima Residual untuk Fungsi Transfer

Model Residual	ARIMA	Signifikansi ARIMA	Signifikansi Xreg	MAPE Training
ARIMA (1,1,0)		ar1 signifikan	Xreg tidak signifikan	1,36
ARIMA (2,1,0)		ar1 signifikan	Xreg tidak signifikan	1,33
ARIMA (0,1,1)		ma1 signifikan	Xreg tidak signifikan	1,46

Berdasarkan perbandingan nilai MAPE untuk Residual pada model fungsi transfer maka model terbaik adalah model ARIMA (2,1,0) karena menghasilkan MAPE data training yang paling kecil untuk Fungsi Transfer. Berdasarkan identifikasi ini maka model fungsi transfer untuk Luas Areal Karet Nasional adalah pada nilai  $r=0$ ,  $b=0$ , dan  $b=0$  dengan model Noise ARIMA (2,1,0) dan komponen Xreg adalah harga karet sheet nasional.

Tabel 27. Model Residual Tentatif ARIMA (2,1,0) Untuk Fungsi Transfer Luas Areal Karet

```
Series: res
ARIMA(2,1,0)
Coefficients:
      ar1      ar2
      0.1333  -0.2241
s.e.  0.1473  0.1480

sigma^2 estimated as 2.14e+10:  log likelihood=-571.46
AIC=1148.92  AICC=1149.54  BIC=1154.21
> coeftest(modelres)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  0.133333  0.14731  0.9051  0.3654
ar2 -0.22410  0.14800 -1.5142  0.1300
```

Oleh karena model noise ARIMA (2,1,0) menghasilkan komponen ar1 dan ar2 yang tidak signifikan maka perlu dilakukan pengujian untuk model noise yang lain. Setelah dilakukan pengujian terhadap seluruh model tentatif, ternyata sebagian besar tidak signifikan baik komponen ar maupun komponen ma, seperti pada ARIMA (1,1,0) komponen ar1 tidak signifikan. Model noise ARIMA (0,1,1) menghasilkan komponen ma1 yang signifikan (Tabel 24).

Tabel 28. Model Residual Tentatif ARIMA (1,1,0) Untuk Luas Areal Karet

```
Series: res
ARIMA(1,1,0)

Coefficients:
      ar1
      0.1086
s.e.    0.1498

sigma^2 estimated as 2.205e+10:  log likelihood=-572.57
AIC=1149.15  AICC=1149.45  BIC=1152.67
> coeftest(modelres)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  0.10859    0.14979  0.7249  0.4685
```

Tabel 29. Model Residual Tentatif ARIMA (0,1,1) Untuk Luas Areal Karet

```
Series: res
ARIMA(0,1,1)

Coefficients:
      ma1
      0.1654
s.e.    0.1736

sigma^2 estimated as 2.189e+10:  log likelihood=-572.42
AIC=1148.84  AICC=1149.14  BIC=1152.36
> coeftest(modelres)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ma1  0.16544    0.17356  0.9532  0.3405
```

Tahap berikutnya adalah dilakukan pengepasan model dengan noise. Pengepasan model dengan noise dilakukan dengan menggabungkan nilai  $r=0$ ,  $s=0$ ,  $b=0$  dan hasil identifikasi model noise yaitu ARIMA(2,1,0). Selanjutnya diperoleh output model fungsi transfer ARIMA (2,1,0) seperti pada Tabel 30.

Tabel 30. Model Fungsi Transfer ARIMA (2,1,0) Untuk Luas Areal Karet

```
Series: train.h[, "Areal"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1    ar2    xreg
      0.4146  0.1004  0.0302
s.e.    0.1498  0.1494  4.0288

sigma^2 estimated as 3.936e+09:  log likelihood=-534.6
AIC=1077.19  AICC=1078.24  BIC=1084.24

Training set error measures:
Training set 14694.04  ME 59815.86  RMSE 40835.16  MAE 0.5001509  MPE 1.328636  MAPE 0.7958735
Training set -0.08419297  ACF1
> coeftest(tf.arima1)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  0.414630    0.149767  2.7685 0.005631 **
ar2  0.100416    0.149447  0.6719 0.501635
xreg 0.030215    4.028782  0.0075 0.994016
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model fungsi transfer luas areal karet dengan ARIMA(2,1,0) menghasilkan MAPE yang sangat baik karena nilainya sangat kecil yaitu sebesar 1,32%. Untuk mengetahui *performance* dari model fungsi transfer tersebut dilakukan pengujian *coeftest* untuk mengetahui tingkat signifikansi dari masing- masing koefisien. Berdasarkan hasil uji Z tes untuk fungsi transfer ARIMA (2,1,0), untuk koefisien ar1 signifikan dengan tingkat kepercayaan 99%, dan komponen fungsi transfer xreg yaitu variabel harga karet sheet tidak signifikan.

Oleh karena itu komponen xreg tidak signifikan, maka akan dicobakan model fungsi transfer dengan model tentatif komponen noise ARIMA (0,1,1). Hasil output untuk fungsi transfer dengan model noise ARIMA (0,1,1) seperti pada Tabel 31.

Tabel 31. Model Tentatif Fungsi Transfer ARIMA (1,1,1) Untuk Luas Areal Karet

```
Series: train.h[, "Areal"]
Regression with ARIMA(0,1,1) errors

Coefficients:
      ma1      xreg
      0.4336  -0.2901
s.e.    0.1354   4.4287

sigma^2 estimated as 4.039e+09:  log likelihood=-535.66
AIC=1077.32  AICC=1077.93  BIC=1082.6

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 20458.58 61350.74 45037.06 0.6863286 1.465371 0.8777682
      ACF1
Training set -0.06133488
> coeftest(tf.arima1)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ma1  0.43359   0.13539   3.2026 0.001362 **
xreg -0.29008   4.42874  -0.0655 0.947776
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MAPE yang dihasilkan dari fungsi transfer ARIMA (0,1,1) adalah 1,47 lebih tinggi dari ARIMA (2,1,0), jadi lebih baik dari MAPE fungsi transfer ARIMA(2,1,0). Selanjutnya dilakukan pengujian *coeftest* untuk fungsi transfer ARIMA(0,1,1) untuk melihat signifikansi dari komponen ma. Dari hasil pengujian di atas, ditemukan bahwa komponen ma1 signifikan pada tingkat kepercayaan 99%, sebaliknya komponen fungsi transfer xreg tidak signifikan. Berdasarkan hasil penelusuran model dan uji yang telah dilakukan maka disimpulkan bahwa model fungsi transfer luas areal karet ARIMA (2,1,0) lebih baik dibandingkan model ARIMA fungsi transfer (0,1,1).

### Peramalan Berbasis Fungsi Transfer

Berdasarkan model fungsi transfer dengan model noise ARIMA (2, 1, 0), dilakukan peramalan berbasis nilai aktual dimana luas areal karet diestimasi menggunakan data aktual harga karet sheet periode 2015-2020. Meskipun data aktual luas areal karet periode 2015-2020 telah ada, dilakukan peramalan luas areal karet untuk mengecek *performance/kelayakan* model fungsi transfer.

Tabel 32. Pengujian MAPE Model Fungsi Transfer ARIMA (2,1,0) Dengan Faktor Input Harga Karet Sheet Data Aktual tahun 2015 - 2020

```
Series: test.h[, "Areal"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
 0.4146  0.1004  0.0302
s.e.  0.0000  0.0000  0.0000

sigma^2 estimated as 3.936e+09: log likelihood=-57.4
AIC=116.79  AICC=118.13  BIC=116.4

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 12958.5 20863.34 13794.34 0.3508772 0.3736146 0.6564295
      ACF1
Training set -0.2348525
```

Peramalan luas areal karet periode 2015-2020 menggunakan fungsi transfer ARIMA (1,1,1) dengan input harga karet dunia **data aktual** menghasilkan MAPE testing 0,37%.

Tujuan melakukan pemodelan fungsi transfer adalah untuk mendapatkan nilai ramalan periode ke depan, yakni luas areal karet tahun 2021 - 2025. Karena data series input harga karet dunia tersedia hingga tahun 2020, maka perlu dilakukan peramalan harga karet dunia terlebih dahulu atau dengan kata lain peramalan luas areal dilakukan berbasis nilai ramalan harga karet.

Oleh karenanya, untuk pengujian data testing terlebih dahulu dilakukan estimasi harga karet dunia periode 2021 - 2025 menggunakan model ARIMA (2,1,0) sebagaimana yang telah diperoleh dari tahap pencarian model tentatif untuk variabel input, sebagai variabel input harga karet sheet. Pemilihan variabel input harga karet sheet karena secara tidak langsung berpengaruh juga pada luas areal karet nasional. Selanjutnya dilakukan peramalan luas areal karet dengan fungsi transfer ARIMA (2, 1, 0) sebagai model terbaik berdasarkan tahapan pengepasan model dengan noise. Peramalan luas areal dengan fungsi transfer ARIMA (2,1,0) menggunakan **nilai ramalan** harga karet sheet yang telah diramal dengan model ARIMA (4,2,1).

Tabel 33. Pengujian MAPE Model Fungsi Transfer ARIMA (2,1,0) Dengan Faktor Input Harga Karet Sheet Nilai Ramalan tahun 2015 – 2020.

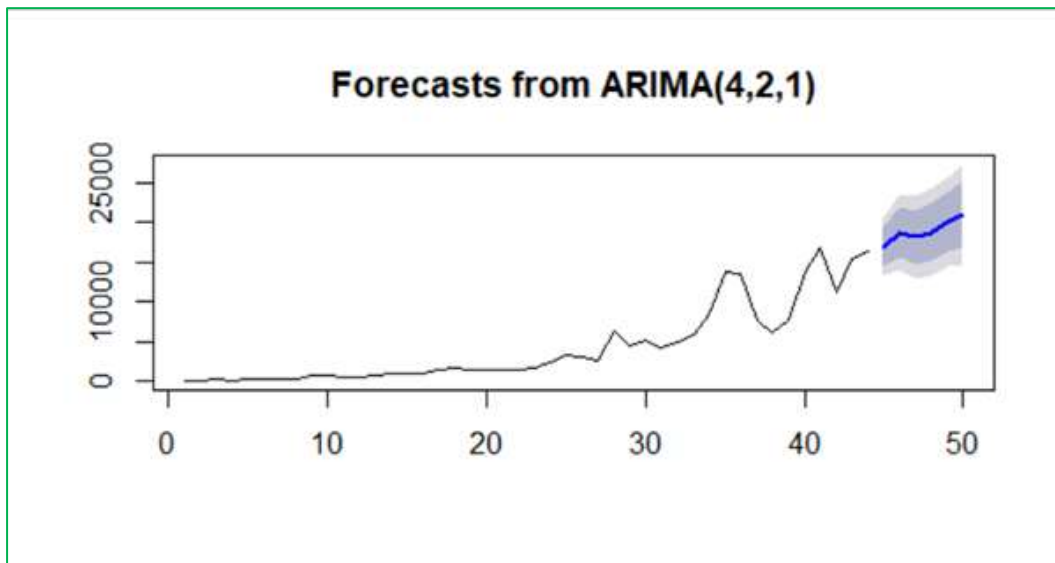
```
Series: test.h[, "Areal"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
 0.4146  0.1004  0.0302
s.e.  0.0000  0.0000  0.0000

sigma^2 estimated as 3.936e+09: log likelihood=-57.4
AIC=116.8  AICC=118.14  BIC=116.41

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 12967.47 20887.04 13799.51 0.3511277 0.3737621 0.6566757
      ACF1
Training set -0.2333819
```

Estimasi luas areal karet berbasis fungsi transfer dengan model fungsi transfer ARIMA (2,1,0) selama 6 tahun terakhir (2015-2020) menggunakan input harga karet sheet hasil **angka ramalan** ARIMA (4,2,1) menghasilkan MAPE testing 0,37%. Nilai MAPE yang kurang dari 1% ini untuk penyusunan estimasi sudah sangat baik, karena kesalahannya sangat kecil.



Gambar 15. Hasil Pemalan Data Testing Faktor Input Harga Karet Sheet 2015 - 2020

Setelah dilakukan peramalan luas areal karet baik menggunakan input (harga karet dunia) data aktual maupun ramalan, tahapan berikutnya adalah pengepasan model arima output. Pengepasan model ARIMA output dimaksudkan untuk membandingkan hasil ramalan luas areal baik berdasarkan data training (1970-2014) maupun data testing (2015-2020). Model ARIMA biasa, tanpa fungsi transfer, terbaik untuk variabel luas areal yang terpilih adalah ARIMA (1,1,0). Model ARIMA (1,1,0) tersebut merupakan ARIMA biasa dan bukan hasil dari fungsi transfer. Hasil ramalan model ARIMA tanpa fungsi transfer dan model ARIMA dengan fungsi transfer selanjutnya akan dibandingkan.



Tabel 29. Pengujian MAPE Model ARIMA (2,1,0) Tanpa Fungsi Tranfer Untuk Faktor Output (Luas Areal Karet) tahun 2015 – 2020.

```

Series: train.h[, "Areal"]
ARIMA(2,1,0)

Coefficients:
      ar1      ar2
    0.4147  0.1004
s.e.  0.1496  0.1494

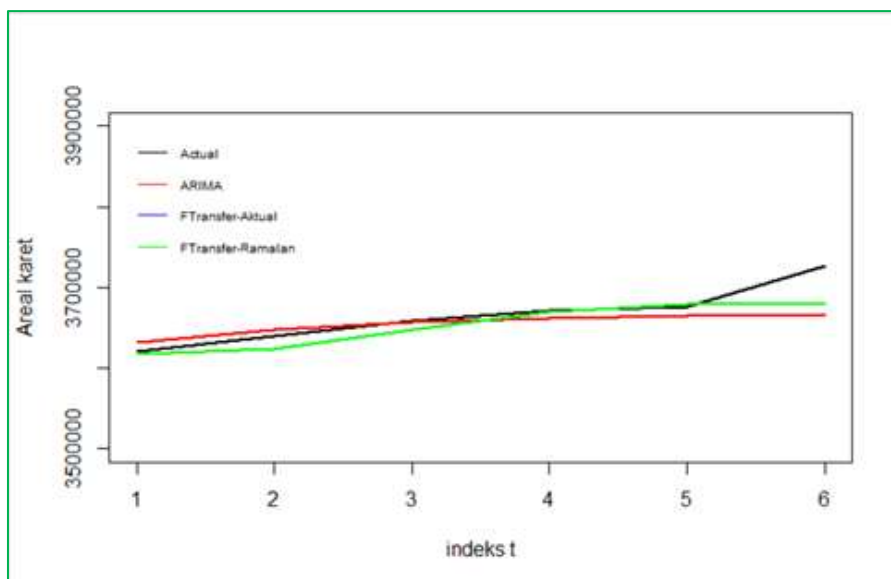
sigma^2 estimated as 3.84e+09:  log likelihood=-534.6
AIC=1075.19  AICC=1075.81  BIC=1080.47

Training set error measures:
Training set      ME      RMSE      MAE      MPE      MAPE      MASE
14698.77  59815.86  40834.03  0.5002917  1.328598  0.7958515
ACF1
Training set -0.08426372
> ramalan.output <- forecast(model.output,6)
> accuracy(ramalan.output,test.h[, "Areal"])
Training set      ME      RMSE      MAE      MPE      MAPE      MASE
14698.77  59815.86  40834.03  0.5002917  1.328598  0.7958515
Test set      10100.45  25341.78  16705.90  0.2701098  0.452125  0.3255965
ACF1
Training set -0.08426372
Test set      NA

```

Hasil peramalan luas areal karet dengan ARIMA biasa atau ARIMA tanpa fungsi transfer menunjukkan bahwa hasil ramalan jika menggunakan data training menghasilkan MAPE = 1,32%. Di sisi lain, MAPE yang dihasilkan dari data testing yaitu 0,45%. Artinya jika model yang digunakan hanya ARIMA tanpa fungsi transfer maka MAPE data testing cenderung lebih tinggi dibandingkan dengan MAPE fungsi transfer ARIMA (2,1,0).

Untuk membandingkan ketepatan model estimasi, dilakukan perbandingan hasil estimasi terhadap data aktual luas areal karet pada tahun 2015-2020 (data testing). Hasil ramalan yang dibandingkan yaitu ramalan dengan model ARIMA(2,1,0), model fungsi transfer ARIMA(2,1,0) dimana input harga karet sheet yang digunakan adalah data aktual maupun ramalan. Selanjutnya hasil ramalan dibandingkan dalam sebuah grafik seperti Gambar 15.



Gambar 16. Perbandingan Hasil Ramalan Luas Areal Karet Tahun 2015-2020

Dari grafik di atas terlihat jika dibandingkan dengan data aktual luas areal karet 2015-2020 (warna hitam), maka hasil peramalan dengan model ARIMA tanpa fungsi transfer (warna merah) tidak mengikuti pola data aktualnya dan cenderung *underestimate*, karena hasil peramalan selalu lebih kecil dari nilai aktualnya / nilai sebenarnya. Sebaliknya, peramalan dengan fungsi transfer khususnya jika input harga karet sheet yang digunakan adalah data aktual maka hasil ramalan luas arealnya (warna biru) sangat menyerupai pola data luas areal sebenarnya selama 6 tahun terakhir. Jika input harga karet sheet yang digunakan adalah hasil ramalan, maka estimasi luas arealnya (warna hijau) juga menghasilkan angka estimasi yang sangat baik karena makin mendekati nilai luas areal sebenarnya (data aktual), hal ini juga terlihat di grafik data hasil ramalan dan data actual hampir berimpit pada tahun ke-1 sampai ke-5, dan agak berbeda pada tahun ke-6.

Hasil identifikasi model untuk meramalkan luas areal karet dengan ke-3 cara diatas menghasilkan nilai MAPE masing-masing sebagai berikut:

Tabel 30. Perbandingan MAPE untuk Arima biasa dan Arima Fungsi Transfer

No	Model	MAPE Training	MAPE Testing
1	ARIMA (2,1,0) (tanpa Fungsi transfer)	1,32%	0,45%
2	Fungsi Transfer dengan ARIMA (2,1,0) input data aktual	1,32%	0,37%
3	Fungsi Transfer dengan ARIMA(2,1,0) input data ramalan	1,32%	0,37%

Berdasarkan nilai MAPE yang dihasilkan maka model fungsi transfer ARIMA(2,1,0) dengan input data aktual hampir sama jika dibandingkan dengan model fungsi transfer ARIMA (2,1,0) dengan input data ramalan maupun model ARIMA(2,1,0) tanpa fungsi transfer. Namun demikian, karena tujuan penyusunan model adalah untuk mendapatkan angka ramalan tahun 2021 - 2025, maka digunakan model Fungsi Transfer berbasis input data ramalan, yakni dengan terlebih dahulu dilakukan estimasi harga karet sheet lima tahun ke depan menggunakan model ARIMA (4,2,1). Hasil estimasi harga karet dunia tersebut kemudian menjadi input untuk meramalkan luas areal karet lima tahun ke depan menggunakan model fungsi transfer ARIMA (2,1,0) dan sebagai variabel input harga karet sheet. Hasil ramalan luas areal karet disajikan pada Tabel 26 dan Tabel 27.

Selain mencari model terbaik untuk meramalkan luas areal karet, akan diestimasi juga luas areal karet lima tahun ke depan (2021-2025) menggunakan fungsi transfer ARIMA (2,1,0) dengan menggunakan seluruh data. Setelah dilakukan run model Fungsi Transfer dengan ARIMA (2,1,0) untuk seluruh data (tidak lagi menggunakan data training), maka dihasilkan MAPE untuk seluruh data sebesar 1,21%, nilai ini juga relative kecil, sehingga estimasi lima tahun kedepan cukup akurat.

Tabel 31. Hasil Peramalan Luas Areal Karet Model Fungsi Transfer Arima (2,1,0) Tahun 2021 – 2025

```

Series: dataestimasi[, "Areal"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
    0.4134  0.1044  0.4367
s.e.  0.1411  0.1402  2.8981

sigma^2 estimated as 3.475e+09:  log likelihood=-606.35
AIC=1220.7  AICc=1221.6  BIC=1228.26

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 13843.05 56542.47 37455.83 0.4640413 1.21035 0.7889852 -0.0840228
Time Series:
Start = 51
End = 55
Frequency = 1
[1] 3747113 3762496 3771038 3775459 3778752
    
```

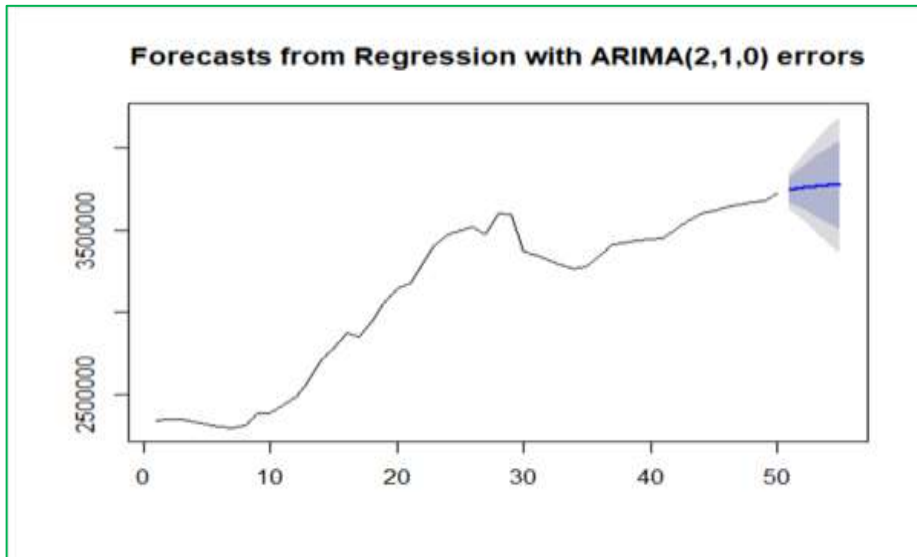
Tabel 32. Hasil Peramalan Luas Areal Karet Tahun 2021-2025 Model Fungsi Transfer ARIMA (2,1,0)

Tahun	Luas Areal (Ha)	Pertumbuhan (%)
<b>2020</b>	3,726,173	
<b>2021</b>	3,747,113	0.56
<b>2022</b>	3,762,496	0.41
<b>2023</b>	3,771,038	0.23
<b>2024</b>	3,775,459	0.12
<b>2025</b>	3,778,752	0.09
<b>Rata-rata Pertumbuhan 2021 -2025</b>		0.21

Tahun 2020 : Angka Tetap (Ditjenbun)

Tahun 2021 – 2025 : Estimasi Berdasarkan Model Fungsi Transfer

Pada tahun 2020 Angka Tetap (ATAP) untuk luas areal karet sebesar 3.726.173 hektar (Ditjen Perkebunan, 2021). Hasil peramalan dengan fungsi transfer ARIMA (2,1,0) menunjukkan bahwa selama lima tahun ke depan luas areal karet akan meningkat rata-rata 0,21% per tahun. Tahun 2021 luas areal karet diestimasi sebesar 3.747.113 hektar atau naik 0,56% dibandingkan tahun 2020. Hasil estimasi tersebut hampir sama dibandingkan Angka Estimasi Ditjenbun untuk luas areal karet tahun 2021 adalah sebesar 3.692.351 hektar atau lebih rendah 54 ribu hektar atau 1,46% . Luas areal karet diramalkan terus meningkat hingga tahun 2025, dengan pertumbuhan yang semakin melambat, hingga pada tahun 2025 luas arealnya diestimasi sebesar 3.778.752 hektar. Pada Gambar 17 menunjukkan bahwa hasil peramalan selama lima tahun ke depan (2021 – 2025) luas areal karet masih tumbuh positif meskipun pertumbuhan kecil.



Gambar 17. Plot Peramalan Luas Areal Karet Nasional Menggunakan Fungsi Transfer

### C. Model VAR (*Vector Auto Regressive*)

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen. Model VAR berlaku pada saat nilai setiap variabel dalam sebuah system tidak hanya bergantung pada lag-nya sendiri, namun juga pada nilai lag variabel lain.

Tahapan yang dilakukan dalam penyusunan model VAR adalah sebagai berikut: persiapan data, pembagian data training dan testing, pemilihan lag dan type, pengajuan asumsi, ramalan data training, testing, penghitungan MAPE, dan plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Disamping itu dibahas Interpretasi Impulse Response Function (IRF) dan Variance Decomposition.

Variabel yang digunakan untuk estimasi model VAR adalah luas areal (areal) dalam satuan hektar, produksi (Produksi) dalam satuan ton, harga karet sheet (HRsheet) dalam satuan Rp/kg, harga karet dunia (HKaretworld) dalam satuan US\$/kg, volume ekspor karet (volekkaret) dalam satuan ton, dan volume impor (volimkaret) dalam satuan ton. Data luas areal karet, luas areal, volume ekspor karet, volume impor karet, dan harga karet sheet diperoleh dari publikasi Ditjen Perkebunan, sementara data variabel harga karet dunia diperoleh dari World Bank. Series masing-masing variabel berbeda karena keterbatasan ketersediaan data. Series data luas areal karet adalah dari tahun 1970-2020, series data harga sheet adalah dari tahun 1971-2020, series data harga karet dunia adalah dari tahun 1970-2020, sementara series data volume ekspor dan impor karet adalah dari tahun 1970-2020. Format data yang digunakan bisa dalam bentuk excell (CSV).

## **Pembagian Data Training dan Testing**

Series data yang digunakan adalah series tahun 1971 – 2020 akan dibagi menjadi 2 set data yakni set data training (tahun 1971-2014) atau 44 titik dan set data testing (2015-2020) atau 6 titik.

## **Pemilihan Lag (p) dan Type**

Dalam permodelan VAR karet ini digunakan enam variabel, yaitu luas areal karet (areal), produksi karet (Produksi), harga sheet (HRsheet), harga karet dunia (HKaretworld), volume ekspor karet (volekkaret), dan volume impor (volimkaret). Selain komposisi variabel tersebut, komponen konstanta dan trend juga perlu dilakukan pengujian untuk mengetahui signifikan atau tidak keberadaannya. Jika keduanya signifikan, maka komponen tersebut harus dimasukkan ke dalam model VAR dengan type “both”. Jika hanya konstanta yang signifikan, maka trend perlu dikeluarkan dari model VAR dengan model VAR type “const”. Jika hanya trend yang signifikan maka konstanta dikeluarkan dari model menggunakan model VAR type “trend”, dan jika keduanya tidak signifikan, maka type yang digunakan model VAR adalah “none”.

Keberadaan konstanta dan trend dapat dideteksi dari plot data awal, namun terkadang hal tersebut sulit dilakukan. Untuk mengantisipasi hal tersebut, sebaiknya dilakukan uji coba/trial and error menggunakan model VAR dari lag  $p=1$  sampai dengan lag  $p=5$  dengan type “both” untuk mengetahui signifikan atau tidak keberadaannya. Untuk panjang lag maksimum bisa dilakukan *trial error* sampai tidak memungkinkan untuk dilakukan permodelan.

Untuk data karet ini, setelah dilakukan running model VAR dengan lag  $p=1$  type “both”, diperoleh informasi komponen konstanta dan trend tidak signifikan, jumlah variabel yang signifikan ada 2 variabel dari total 6 variabel dalam system atau jumlah yg jumlah variabel yang signifikan (33,3%). Selanjutnya dilakukan uji VAR(2) type both, model ini menghasilkan konstanta dan trend yang signifikan, dan ada 5 variabel yang signifikan dari total 12 variabel dalam model (41,7%). Untuk model VAR(3) type both, model ini menghasilkan trend yang signifikan, dan ada 3 variabel yang signifikan dari total 18 variabel dalam model (16,7%). Untuk model VAR(4) type both, model ini menghasilkan konstanta dan trend yang signifikan, dan ada 10 variabel yang signifikan dari total 24 variabel dalam model (41,7%). Selanjutnya model VAR(5) type both, model ini menghasilkan konstanta dan trend yang tidak signifikan, dan hanya ada 5 variabel yang signifikan dari total 30 variabel dalam model (16,7%). Oleh karena jumlah variabel yang signifikan semakin sedikit sehingga penelusuran berhenti di VAR(5) type both (Tabel 33).

Pemilihan lag p ditentukan dengan melihat banyaknya variabel yang signifikan dalam lag tersebut sekaligus memastikan harus ada variabel/peubah yang signifikan pada lag terpilih dimaksud. Setelah dilakukan run model ternyata ada 2 kandidat model VAR terbaik yaitu VAR (2) type “both” dan VAR(4) type “both“. Untuk menguji dua model VAR yang terbaik maka dilakukan pengujian MAPE untuk data training dan MAPE data testing.

Tabel 33. Hasil Pengujian Model VAR pada Beberapa Tingkat Lag p dan Type

Lag (p)	Type	Signifikansi Type	Jumlah Variabel Signifikan	Jumlah Total Variabel	Adj-R <sup>2</sup>
p=1	both	Const dan Trend Tidak Signifikan	2	6	98,47%
p=2	both	Const dan Trend Signifikan	5	12	98,75%
p=3	both	Trend signifikan	3	18	98,49%
p=4	both	Const dan Trend Tidak Signifikan	10	24	99,17%
p=5	both	Const dan Trend Tidak Signifikan	5	30	98,77%

Hasil pengujian besaran MAPE untuk 3 kandidat model VAR terbaik seperti pada Tabel 34, menunjukkan bahwa untuk data training yang memiliki MAPE terkecil adalah VAR(4) type “both”, sedangkan untuk MAPE data testing yang paling kecil adalah VAR(4) type “both”. Oleh karena tujuan peramalan adalah untuk menghasilkan angka yang akurat maka model terbaik yang terpilih adalah model VAR(4) type both, karena menghasilkan MAPE data training dan testing terkecil, meskipun demikian perlu juga dicoba untuk VAR(2) type = both karena menghasilkan MAPE Yang cukup kecil baik untuk data testing maupun training.

Tabel 34. Perbandingan MAPE Pengujian Model VAR

Lag (p)	Type	MAPE Training	MAPE Testing
p=1	both	1,36	4,95
p=2	both	1,07	9,51
p=4	both	0,61	3,42

### Model VAR(2) Type=both

Untuk model VAR kandidat terbaik adalah Model VAR (p=2) type=both. Model VAR(1) type=both bukan termasuk kandidat terbaik karena komponen constanta dan trend tidak signifikan. Hasil model VAR(2) type both, untuk mengestimasi luas areal (t) karet menunjukkan bahwa koefisien constanta menunjukkan hasil yang signifikan pada tingkat kepercayaan 90% dan koefisien trend menunjukkan signifikan pada tingkat 99%. Untuk variabel lain yang signifikan untuk mengestimasi luas areal (t) antara lain luas areal lag1 (signifikan 99%), volume impor karet lag 1 (95%), harga karet dunia lag 2 (99%), dan volume ekspor karet lag 2 (90%), volume impor karet lag 2 (95%).

Model VAR (2) type both ini menghasilkan nilai Adjusted R Square = 98,75%, artinya keragaman luas areal dipengaruhi oleh variabel-variabel penjelasnya sebesar 98,75%. Nilai F hitung = 250,9, sehingga nilai p-value untuk model luas areal ini sangat kecil atau jauh lebih kecil dari 0,05, sehingga model layak untuk digunakan.

Tabel 35. Output Model VAR(2) type=both

```

VAR Estimation Results:
=====
Endogenous variables: Areal, Produksi, Hsheet, HKaretworld, volekkaret, volimkaret
Deterministic variables: both
Sample size: 42
Log Likelihood: -2324.643
Roots of the characteristic polynomial:
0.9789 0.8525 0.8525 0.8022 0.8022 0.6531 0.6531 0.6186 0.6186 0.3787 0.3787 0.2979
Call:
VAR(y = karet[1:44, c(2, 3, 4, 6, 7, 8)], p = 2, type = "both")

Estimation results for equation Areal:
=====
Areal = Areal.l1 + Produksi.l1 + Hsheet.l1 + HKaretworld.l1 + volekkaret.l1 + volimkaret.l1 + Areal.l2 + Produksi.l2 + Hsheet.l2 + HKaretworld.l2 + volekkaret.l2 + volimkaret.l2 + const + trend

      Estimate Std. Error t value Pr(>|t|)
Areal.l1      7.028e-01  2.059e-01  3.413  0.00198 **
Produksi.l1   1.333e-01  2.127e-01  0.627  0.53588
Hsheet.l1     -3.690e-01  2.391e+00 -0.154  0.87848
HKaretworld.l1 -4.186e+04  2.731e+04 -1.533  0.13656
Volekkaret.l1  3.523e-02  1.651e-01  0.213  0.83254
Volimkaret.l1 -4.538e+00  2.175e+00 -2.087  0.04614 *
Areal.l2      2.071e-01  2.099e-01  0.987  0.33213
Produksi.l2   -6.491e-02  1.767e-01 -0.367  0.71603
Hsheet.l2     -7.931e-01  2.681e+00 -0.296  0.76955
HKaretworld.l2  8.805e+04  2.867e+04  3.071  0.00471 **
Volekkaret.l2 -2.754e-01  1.410e-01 -1.953  0.06084 .
Volimkaret.l2 -4.837e+00  2.097e+00 -2.307  0.02866 *
const         2.556e+05  1.352e+05  1.891  0.06907 .
trend         1.157e+04  4.164e+03  2.778  0.00965 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 52020 on 28 degrees of freedom
Multiple R-Squared:  0.9915,    Adjusted R-squared:  0.9875
F-statistic: 250.9 on 13 and 28 DF,  p-value: < 2.2e-16

```

### Pengujian Asumsi VAR(2) type 'both'

Selanjutnya dilakukan pengujian asumsi non autokorelasi, normalitas, dan homoskedastisitas pada sisaan model VAR terbaik. Untuk data karet akan dilakukan pengujian sisaan pada dua model terbaik VAR (2) type 'both'.

Pemeriksaan autokorelasi residual model menggunakan fungsi "serial.test" yang di dalamnya dilakukan pengujian Portmanteau-and Breusch-Godfrey test. Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi. Pengujian Jarque-Bera tests untuk menguji kenormalan, hasil pengujian menunjukkan Nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal. Pemeriksaan heteroskedastisitas model menggunakan fungsi "arch.test" yang di dalamnya dilakukan pengujian ARCH-LM tests. Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.

Tabel 36. Ouput Pengujian Asumsi VAR(2) type=both

```

Portmanteau Test (asymptotic)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 533.75, df = 504, p-value = 0.1736

$JB
      JB-Test (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 100.07, df = 12, p-value = 5.551e-16

$Skewness
      Skewness only (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 19.011, df = 6, p-value = 0.004145

$Kurtosis
      Kurtosis only (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 81.054, df = 6, p-value = 2.109e-15

ARCH (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 756, df = 2646, p-value = 1

```

### Ramalan Data Training, Testing, Penghitungan MAPE, dan Plot

Selanjutnya dilakukan peramalan data, baik untuk data training maupun untuk data testing sekaligus dilakukan penghitungan MAPE. MAPE (Mean Absolute Percentage Error) merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan data hasil peramalan. Untuk menguji suatu model lebih baik dengan model yang lain, maka dilakukan pengujian model dengan membandingkan Nilai MAPE baik untuk data training maupun data testing. Data Testing hasil ramalan luas areal dengan VAR(2) type=both, menghasilkan MAPE =9,51%. Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data aktual dengan data hasil ramalan pada data testing adalah 9,51%. Data training hasil ramalan luas areal dengan nilai p=2 type=both menghasilkan MAPE =1,07%. Model VAR ini menunjukkan ketika menggunakan data training sangat baik, terlihat dari MAPE yang kecil yaitu hanya sebesar 1,07%, namun ketika digunakan untuk melakukan estimasi maka MAPE melonjak menjadi 9,51%, artinya kemampuan dalam meramalkan tidak sebaik data training.



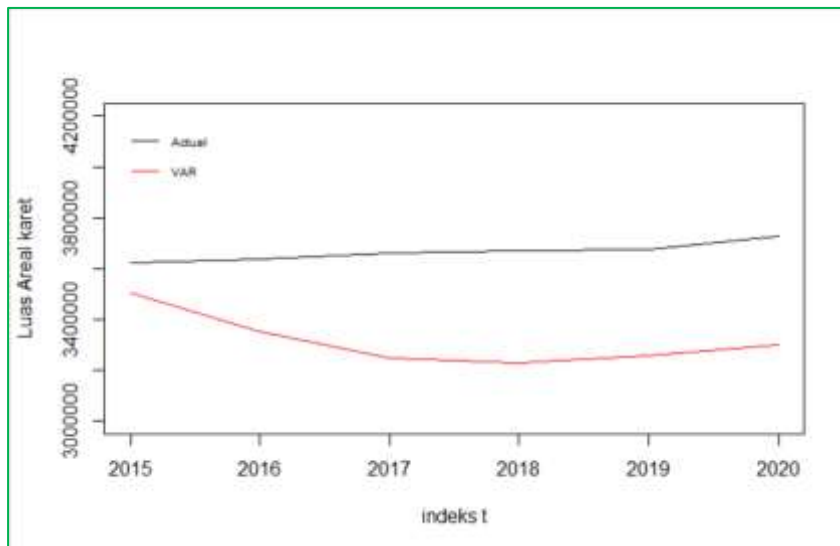
Tabel 37. Pengujian Nilai MAPE untuk Model VAR(2) type=both

MAPE TESTING					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
3.203	8.701	11.245	9.509	11.377	12.088

MAPE TRAINING					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.04557	0.43832	0.85441	1.06639	1.61114	2.96076

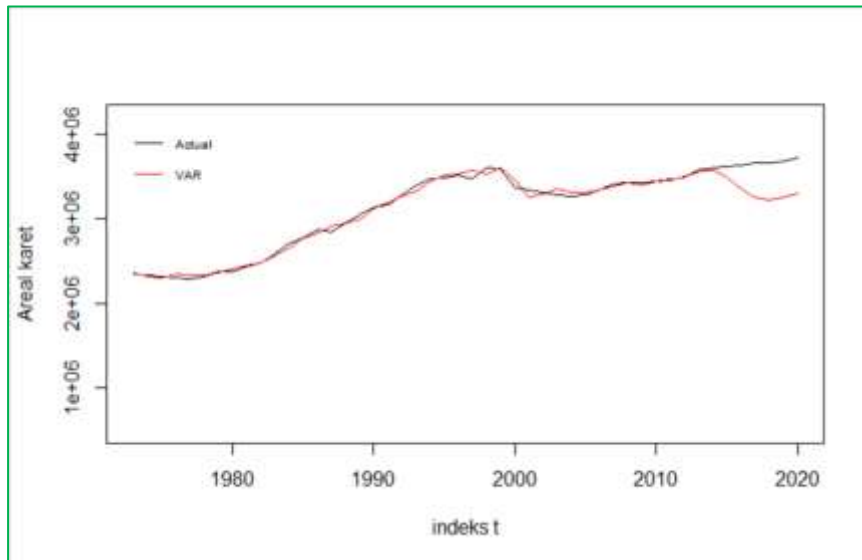
Dari grafik di bawah dapat dilihat bahwa pergerakan ramalan pada data testing tidak mengikuti pergerakan data asli/aktual. Sehingga mungkin model VAR (2) type “both” kemampuan dalam meramalkan perlu dipertimbangkan. Hasil peramalan dengan model VAR (2) ini cenderung lebih rendah dari data aktual. Pada tahun 2015 data actual sedikit lebih tinggi dari data estimasi dengan VAR (2), pada tahun 2016 beda antara data aktual dan data estimasi dengan VAR(2) ini semakin lebar, begitu juga tahun 2017 dan 2018, semakin jauh perbedaan antara data aktual dan hasil estimasi. Untuk tahun 2019 dan 2020 perbedaan antara data estimasi dan data aktual semakin lebar lagi. Untuk data testing ini rata-rata penyimpangannya adalah sebesar 9,59%. Dari segi besaran MAPE sebenarnya model ini sudah cukup baik karena MAPE masih dibawah 10%.



Gambar 18. Plot Ramalan dan Aktual Data Testing VAR (2) Type “both”

Jika plot antara data testing dan data training digabungkan maka bentuk plotnya seperti Gambar 16. Untuk data tahun 1971 - 2015 atau data training plot sangat baik, karena antara data aktual dan estimasi dengan model VAR (2) hampir selalu berimpit plotnya, sehingga MAPE akan kecil. MAPE hasil pengujian untuk data training adalah sebesar 1,07%, suatu nilai yang cukup kecil karena rata-rata penyimpangan hanya 1,07%, artinya model cukup akurat. Namun plot tahun 2015 – 2020 menunjukkan data aktual dengan data estimasi dengan Model VAR (2) ini mulai agak renggang, dimana hasil estimasi rata-rata lebih rendah dari data aktual.

Hasil MAPE data testing ini, menunjukkan nilai yang lebih besar dari MAPE training yaitu sebesar 9,51%.



Gambar 19. Plot Ramalan dan Aktual Data Training dan Testing VAR (2) Type “both”

### Model VAR(4) Type=both

Untuk model VAR kandidat terbaik lainnya adalah Model VAR (4) type=both. Hasil model VAR(4) type both, untuk mengestimasi luas areal (t) karet menunjukkan bahwa koefisien constanta dan trend tidak signifikan karena nilai probability  $|t| > 0.05$ . Untuk variabel lain yang signifikan untuk mengestimasi luas areal (t) adalah variabel : luas areal lag1 (signifikan 90%), volume impor karet lag 1 (99%), produksi karet lag-2 (99%), ahrga karet sheet nasional lag-2 (90%), harga karet dunia lag-2 (99%), produksi karet lag-3 (99%), harga karet dunia lag-3 (99%), volume impor karet lag-3 (95%), produksi karet lag-4 (99%) dan harga karet dunia lag-4 (99%). Dari total jumlah variabel sebanyak 24, ada 10 variabel diantara signifikan, atau sebesar 41,7%.

Model VAR (4) type both ini menghasilkan nilai Adjusted R Square = 99,14%, artinya keragaman luas areal dipengaruhi oleh variabel-variabel penjelasnya sebesar 99,14%. Nilai F hitung = 187,9, sehingga nilai p-value untuk model luas areal ini sangat kecil atau jauh lebih kecil dari 0,05, sehingga model layak untuk digunakan.

Tabel 38. Output Model VAR(4) type=both

```

VAR Estimation Results:
=====
Endogenous variables: Areal, Produksi, Hsheet, HKaretworld, Volekkaret, Volimkaret
Deterministic variables: both
Sample size: 40
Log Likelihood: -2000.988
Roots of the characteristic polynomial:
1.08 1.08 1.025 1.025 1.014 1.014 1.01 1.01 0.9777 0.9777 0.977 0.9588 0.9588 0
.9521 0.9521 0.8808 0.8808 0.8046 0.8046 0.7949 0.7949 0.7575 0.5472 0.5472
Call:
VAR(y = karet[1:44, c(2, 3, 4, 6, 7, 8)], p = 4, type = "both")

Estimation results for equation Areal:
=====
Areal = Areal.l1 + Produksi.l1 + Hsheet.l1 + HKaretworld.l1 + Volekkaret.l1 + volim
karet.l1 + Areal.l2 + Produksi.l2 + Hsheet.l2 + HKaretworld.l2 + Volekkaret.l2 + Vo
limkaret.l2 + Areal.l3 + Produksi.l3 + Hsheet.l3 + HKaretworld.l3 + volekkaret.l3 +
Volimkaret.l3 + Areal.l4 + Produksi.l4 + Hsheet.l4 + HKaretworld.l4 + Volekkaret.l4
+ volimkaret.l4 + const + trend

      Estimate Std. Error t value Pr(>|t|)
Areal.l1      4.067e-01 2.081e-01  1.954 0.07094 .
Produksi.l1   3.954e-01 2.568e-01  1.540 0.14595
Hsheet.l1     -3.989e+00 3.354e+00 -1.189 0.25420
HKaretworld.l1 -3.966e+04 6.326e+04 -0.627 0.54086
Volekkaret.l1  1.120e-01 1.814e-01  0.618 0.54681
Volimkaret.l1 -1.367e+01 3.320e+00 -4.118 0.00105 **
Areal.l2      4.567e-01 3.322e-01  1.375 0.19080
Produksi.l2   -1.147e+00 3.836e-01 -2.989 0.00976 **
Hsheet.l2     -8.344e+00 3.965e+00 -2.105 0.05388 .
HKaretworld.l2  2.501e+05 8.258e+04  3.029 0.00902 **
Volekkaret.l2  1.388e-01 1.780e-01  0.780 0.44852
Volimkaret.l2 -4.487e+00 2.745e+00 -1.635 0.12443
Areal.l3      -3.772e-02 3.365e-01 -0.112 0.91233
Produksi.l3    1.550e+00 4.490e-01  3.453 0.00388 **
Hsheet.l3      4.982e-01 3.411e+00  0.146 0.88597
HKaretworld.l3 -3.248e+05 1.001e+05 -3.245 0.00587 **
Volekkaret.l3  9.170e-02 2.077e-01  0.441 0.66562
Volimkaret.l3  1.061e+01 3.682e+00  2.882 0.01205 *
Areal.l4      1.859e-01 2.542e-01  0.731 0.47662
Produksi.l4   -1.172e+00 3.914e-01 -2.996 0.00963 **
Hsheet.l4     -7.091e+00 4.270e+00 -1.661 0.11900
HKaretworld.l4  3.622e+05 8.962e+04  4.042 0.00121 **
Volekkaret.l4 -7.983e-02 1.788e-01 -0.446 0.66211
Volimkaret.l4 -3.178e+00 2.403e+00 -1.322 0.20721
const        -1.338e+05 2.547e+05 -0.525 0.60748
trend         7.253e+03 7.463e+03  0.972 0.34761
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 40690 on 14 degrees of freedom
Multiple R-Squared: 0.997, Adjusted R-squared: 0.9917
F-statistic: 187.9 on 25 and 14 DF, p-value: 7.978e-14

```

### Ramalan Data Training, Testing, Penghitungan MAPE, dan Plot

Selanjutnya dilakukan peramalan data, baik untuk data training maupun untuk data testing sekaligus dilakukan penghitungan MAPE. MAPE (Mean Absolute Percentage Error) merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan data hasil peramalan. Untuk menguji suatu model lebih baik dengan model yang lain, maka dilakukan pengujian model dengan membandingkan Nilai MAPE baik untuk data training

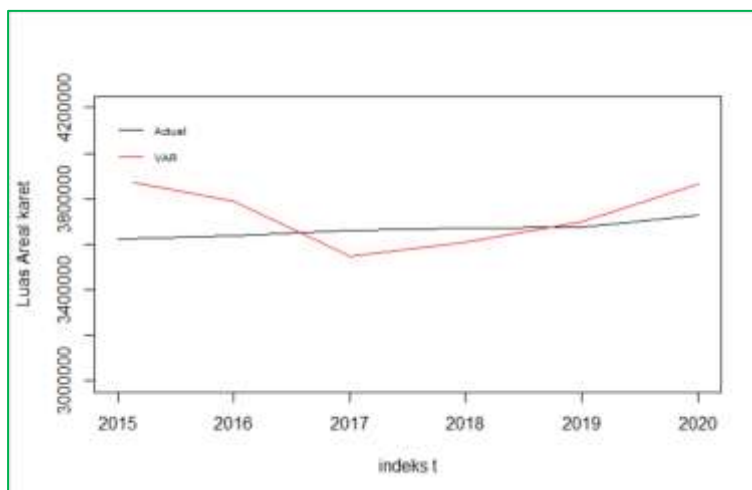
maupun data testing. Data Testing hasil ramalan luas areal dengan VAR(4) type=both, menghasilkan MAPE =9,42%. Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data aktual dengan data hasil ramalan pada data testing adalah 3,42%. Data training hasil ramalan luas areal dengan nilai p=4 type=both menghasilkan MAPE =0,61%. Model VAR ini menunjukkan ketika menggunakan data training sangat baik, terlihat dari MAPE yang kecil yaitu hanya sebesar 0,61%, namun ketika digunakan untuk melakukan estimasi maka MAPE meningkat menjadi 3,41%. Jika dibandingkan dengan model VAR(2) type=bpth, maka model VAR(4) type=both menghasilkan MAPE yang lebih kecil untuk data training maupun data testing, sepertinya model VAR(4) both ini lebih baik.

Tabel 39. Pengujian Nilai MAPE untuk Model VAR(4) type=both

MAPE TESTING					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.6571	2.0380	3.3929	3.4246	4.0204	7.2888

MAPE TRAINING					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.02011	0.25413	0.51583	0.61159	0.78617	2.69311

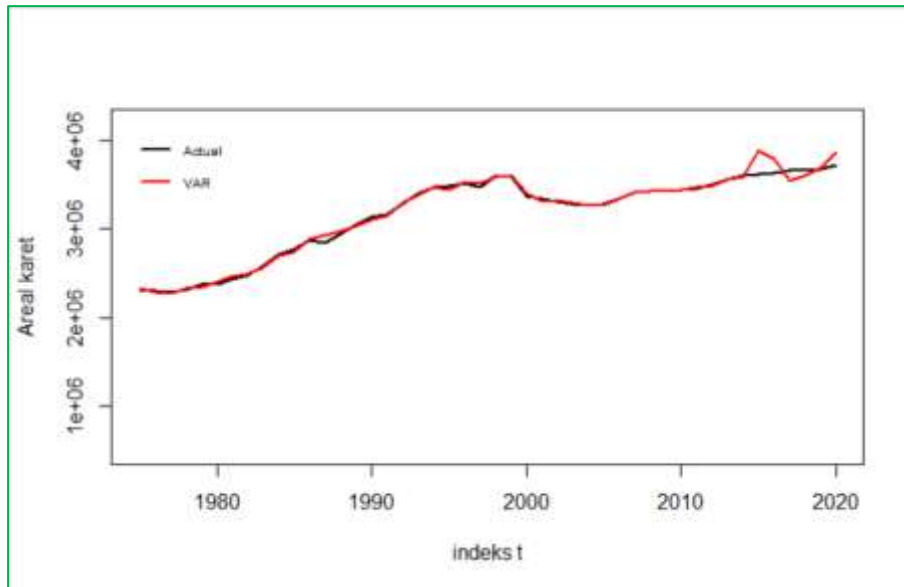
Dari grafik di bawah dapat dilihat bahwa pergerakan ramalan pada data testing mengikuti pergerakan data asli/actual, data actual bisa lebih tinggi atau bisa juga lebih lebih dari nilai ramalan, maka sempat terjadi perpotongan antara data actual dan data ramalan (Gambar 20). Oleh karena menghasilkan MAPE yang cukup kecil untuk data testing maka model VAR (4) type “both” untuk peramalan perlu dipertimbangkan. Hasil peramalan dengan model VAR (2) ini bisa lebih rendah atau lebih tinggi dari data aktual. Pada tahun 2015 dan 2016 data aktual sedikit lebih rendah dari data estimasi dengan VAR (4) both, pada tahun 2017 data aktual sedikit lebih tinggi dari data estimasi, pada tahun 2018 dan 2019 perbedaan data aktual dan hasil estimasi semakin kecil atau hampir berimpit grafiknya, dan tahun 2020 hasil estimasi dengan model VAR(4) both lebih tinggi dari data aktualnya.



Gambar 20. Plot Ramalan dan Aktual Data Testing VAR (4) Type “both”

Jika plot antara data testing dan data training digabungkan maka bentuk plotnya seperti Gambar 21. Untuk data tahun 1971 - 2015 atau data training plot sangat baik, karena antara

data aktual dan estimasi dengan model VAR (4) hampir selalu berimpit plotnya, sehingga MAPE akan kecil. MAPE hasil pengujian untuk data training adalah sebesar 0,61%, suatu nilai yang cukup kecil karena rata-rata penyimpangan hanya 0,61%, artinya model cukup akurat. Pada saat plot tahun 2015 – 2020 menunjukkan data aktual dengan data estimasi dengan model VAR (4) ini mulai agak renggang, dimana hasil estimasi dapat lebih rendah atau lebih tinggi dari data aktual. Hasil MAPE data testing ini, menunjukkan nilai yang lebih besar dari MAPE training yaitu sebesar 3,42%.



Gambar 21. Plot Ramalan dan Aktual Data Training dan Testing VAR (4) Type “both”

### **Pengepasan Model untuk Seluruh Data dan Peramalannya**

Ada 2 Model VAR tentative terbaik dari hasil perbandingan nilai MAPE yaitu model VAR (2) type “both” dan VAR(4) type “both, sehingga dapat dilanjutkan pengepasan model pada keseluruhan data. Setelah dilakukan pengepasan model, maka dilakukan estimasi luas areal karet untuk 5 tahun kedepan yaitu tahun 2021 -2025.

Peramalan dengan model VAR (2) type “both” menggunakan keseluruhan data menunjukkan luas areal karet tahun 2021 sebesar 3,79 juta ha, naik kembali pada tahun 2022 menjadi 3,83 juta ha, terus turun meningkat sehingga pada tahun 2025 luas areal karet diperkirakan mencapai 4,03 juta ha. Hasil run model untuk VAR(4) type “both” menghasilkan angka estimasi yang fluktuatif, maka tahun 2021 luas areal karet diperkirakan 3,88 juta ha, tahun 2022 naik sedikit menjadi 3,89 juta ha, tahun 2023 turun menjadi 3,88 juta hektar, tahun 2024 cenderung tetap, dan akhirnya tahun 2025 kembali naik menjadi 3,94 juta ha.

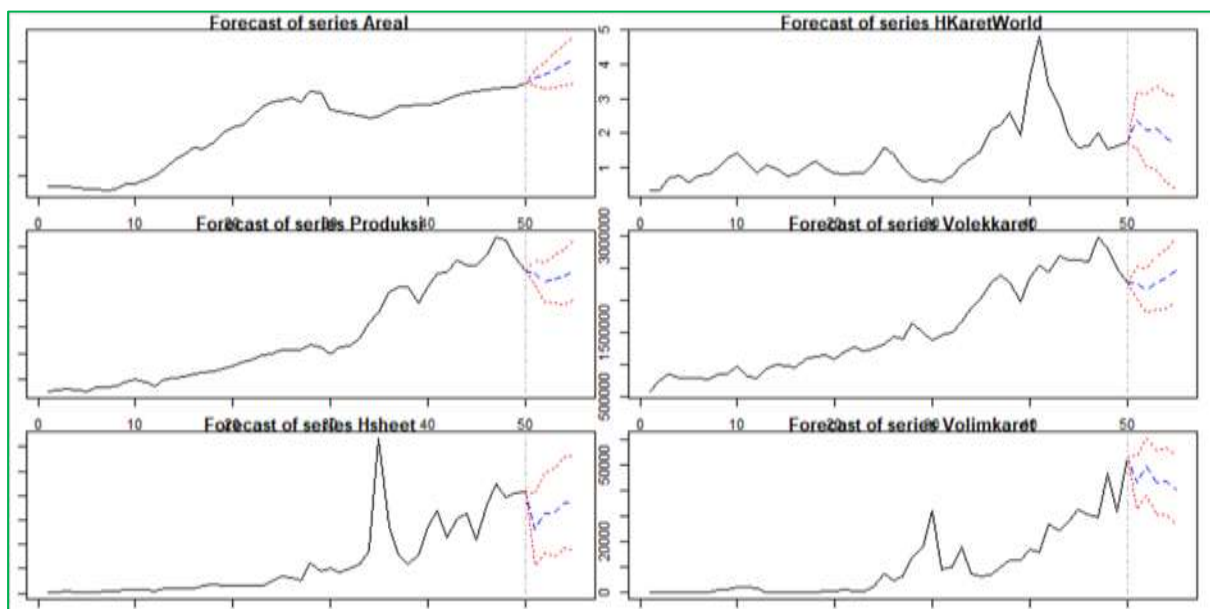
Tabel 40. Hasil Peramalan Luas areal Karet Tahun 2021-2025, Model VAR(2) Type=both

[1] 3787518 3827540 3896807 3967502 4034359

Tabel 41. Hasil Peramalan Luas areal Karet Tahun 2021-2025, Model VAR(4) Type=both

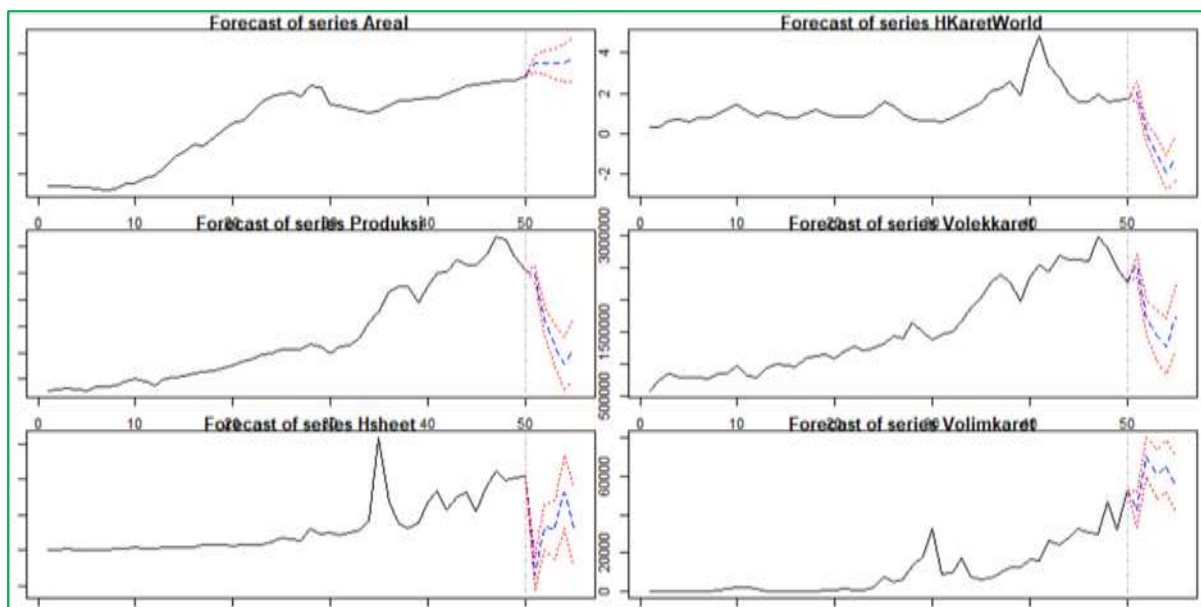
[1] 3882613 3891241 3885618 3888491 3944120

Keunggulan model VAR ini pada saat dilakukan peramalan maka semua variabel dalam system akan ikut diramalkan. Hasil peramalan bukan hanya target variabel, tetapi semua variabel yang masuk dalam sistem. Untuk Model VAR (2) both, hasil peramalan 5 tahun ke depan untuk variabel luas areal menunjukkan luas areal akan terus meningkat secara linier. Sebaliknya untuk produksi akan turun dalam 2 tahun ke depan, kemudian meningkat kembali tahun 2023 – 2025. Sebaliknya harga karet sheet dalam negeri cenderung akan turun, sementara harga karet dunia diperkirakan akan perlahan terus meningkat kemudian sedikit menurun kembali. Volume ekspor karet cenderung turun pada tahun 2021 -2022 tetapi meningkat kembali setelahnya, sebaliknya volume impor karet berfluktuasi tetapi cenderung turun.



Gambar 22. Hasil Estimasi untuk Semua Variabel dalam Sistem Model VAR(2) Both

Untuk Model VAR (4) both, hasil peramalan 5 tahun ke depan untuk variabel luas areal menunjukkan luas areal akan cenderung meningkat secara linier. Sebaliknya untuk produksi akan turun secara tajam dalam 5 tahun ke depan. Sebaliknya harga karet sheet dalam negeri diperkirakan berfluktuasi tetapi ada kecenderungan meningkat, sementara harga karet dunia diperkirakan masih terus menurun, kembali meningkat pada tahun ke-5. Volume ekspor karet cenderung turun karena harga karet dunia yang masih turun, sebaliknya volume impor karet berfluktuasi tetapi cenderung meningkat.



Gambar 23. Hasil Estimasi untuk Semua Variabel dalam Sistem Model VAR(4) Both

Untuk model VAR (2) type “both”, pergerakan hasil ramalan luas areal karet cenderung mengalami peningkatan dengan rata-rata pertumbuhan luas areal tahun 2021 -2025 sebesar 1,59% per tahun. Untuk model VAR(4) type “both” pergerakan hasil ramalan luas areal karet meningkat cukup tinggi pada tahun 2021 sebesar 4,20%. Dari data historisnya pertumbuhan luas areal karet setiap tahun biasanya dibawah 1 % saja atau paling tinggi 1,5%. Pertumbuhan di atas 4% akan sulit dicapai mengingat kondisi saat ini tidak ada penambahan dari ekstensifikasi yang cukup signifikan. Pertumbuhan luas areal karet hasil peramalan tahun 2021 – 2025 model VAR(4) both, rata-rata masih tumbuh sebesar 0,40%. Rata-rata pertumbuhan lebih kecil dibandingkan dengan Model VAR (2) both.

Jika berdasarkan nilai MAPE, maka model terbaik yang terpilih untuk peramalan data luas areal karet adalah model VAR (4) type “both” karena memberikan nilai MAPE training dan testing yang lebih kecil sekaligus memberikan plot ramalan yang lebih lambat pertumbuhannya. Model Var(2) both meskipun menghasilkan MAPE data training dan testing yang lebih besar tetapi hasil peramalan tidak menunjukkan lonjakan pertumbuhan luas areal yang terlalu besar, sehingga pertumbuhan luas areal lebih realistis.

Jika dibandingkan dengan tahun 2020 maka luas areal karet hasil estimasi tahun 2021 akan naik sebesar 1,65%, sehingga menjadi sekitar 3,79 juta hektar. Pada tahun 2022 luas areal karet masih meningkat sebesar 1,06% menjadi 3,83 juta hektar. Hal ini didukung dari realitas luas areal karet yang hanya sedikit saja mengalami peningkatan, karena harga karet dunia yang masih cenderung turun. Peningkatan luas areal karet hasil estimasi selama 5 tahun kedepan adalah sekitar 0,40% sampai 1,59% per tahun. Meskipun terjadi peningkatan tetapi sangat lambat, hal ini bisa saja terjadi karena beberapa tahun belakangan ini harga karet alam dunia cenderung turun dan stabil di harga rendah. Beberapa petani karet telah mengkonversi lahannya ke tanaman yang lebih menguntungkan seperti kelapa sawit atau tanaman perkebunan lainnya.



Tabel 43. Hasil Estimasi Luas Areal Karet Nasional dengan Model VAR(2) dan VAR(4) Both

Tahun	Model VAR(2) Both	Pertumbuhan (%)	Model VAR(4) Both	Pertumbuhan (%)
2020	3.726.173		3.726.173	
2021	3.787.518	1,65	3.882.613	4,20
2022	3.827.540	1,06	3.891.241	0,22
2023	3.896.807	1,81	3.885.618	(0,14)
2024	3.967.502	1,81	3.888.491	0,07
2025	4.034.359	1,69	3.944.120	1,43
<b>Rata-rata Petumbuhan</b>		1,59		0,40

Keterangan : Tahun 2020 Angka Tetap Ditjenbun

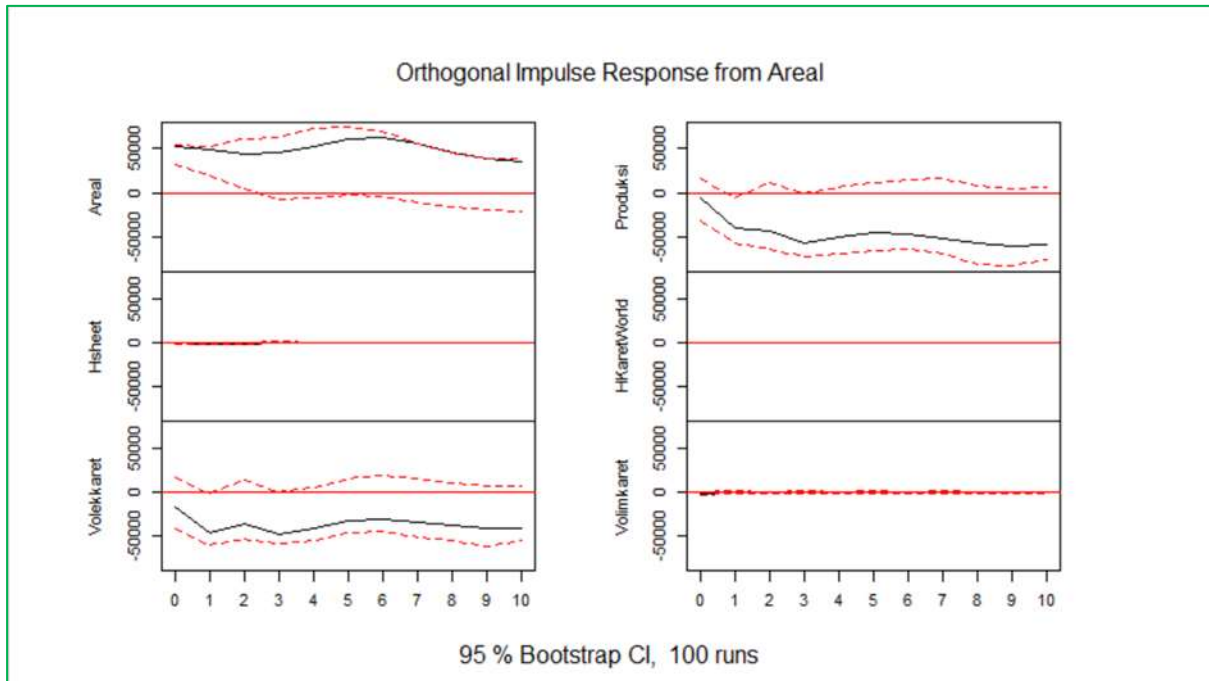
Tahun 2021 – 2025 : Estimasi Berdasarkan Model

### Interpretasi Impulse Response Function (IRF) dan Variance Decomposition

Setelah diperoleh model terbaik, untuk permodelan VAR dapat dilakukan interpretasi lebih mendalam terkait Impulse Response Function (IRF) dan Variance Decomposition. Berikut interpretasi ketiga hal di atas untuk model terbaik VAR (2) type “both”. **Impulse Response Function** akan menjelaskan bagaimana perubahan atau gejolak yang terjadi pada suatu variabel/peubah di tahun tertentu akan berdampak pada variabel/peubah lain di tahun tertentu dan tahun-tahun setelahnya.

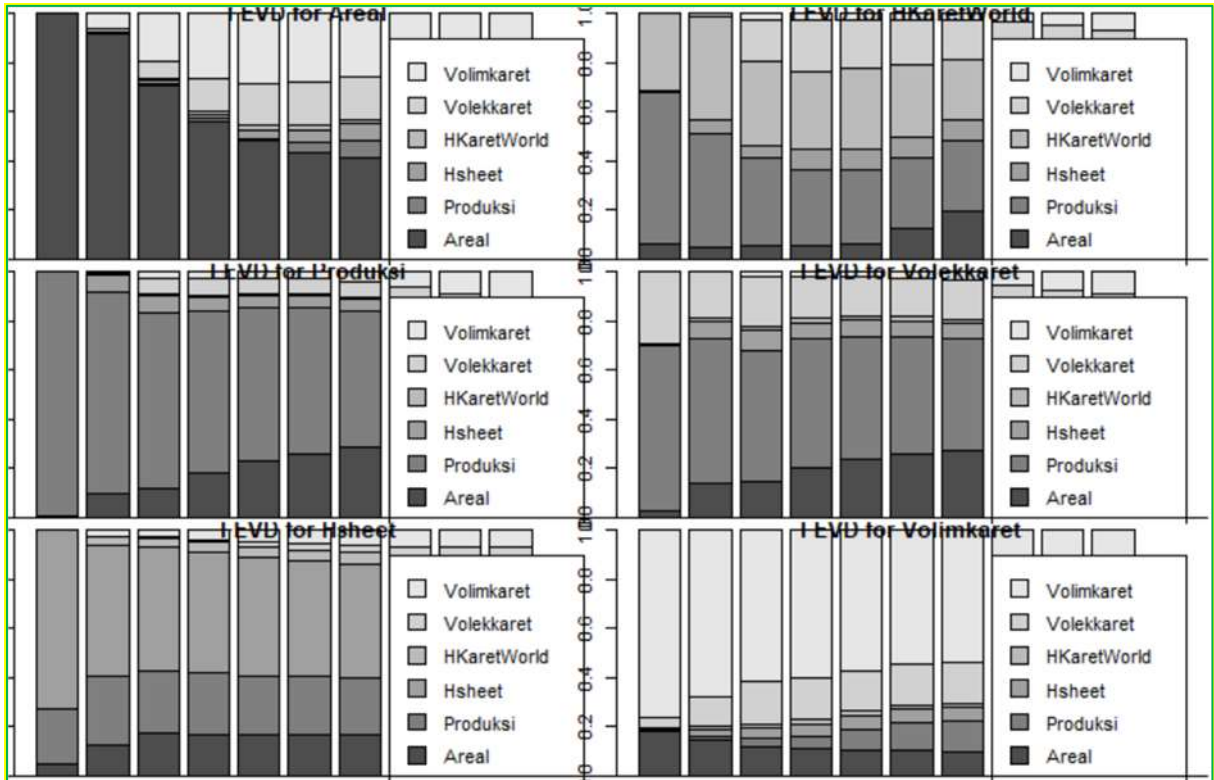
Dari grafik Impulse Response Function luas areal model terbaik VAR (2) type “both”, dapat dilihat bahwa jika terjadi perubahan pada luas areal di tahun tertentu maka akan berdampak pada luas areal itu sendiri sampai 2 tahun ke depan. Dampak tersebut akan hilang setelah 2 tahun. Sedangkan dampak perubahan luas areal tidak berdampak pada harga sheet, harga karet dunia, dan volume impor karet, namun masih sedikit berdampak pada volume ekspor kedepan dan produksi. Perubahan luas areal berdampak pada volume ekspor sampai dengan satu tahun dan akan hilang dampaknya setelah melewati tahun pertama tersebut, sedangkan pada produksi dampaknya terjadi bersamaan.





Gambar 24. Impuls Respon Beberapa Variabel Terhadap Luas areal Model VAR(2) Both

**Variance Decomposition** atau dekomposisi keragaman digunakan untuk melihat variabel apa saja yang mempengaruhi komposisi keragaman suatu variabel/peubah. Dari grafik dekomposisi keragaman model terbaik VAR (2) type “both” di atas dapat dilihat bahwa komposisi luas areal pada tahun pertama dipengaruhi sepenuhnya oleh luas areal itu sendiri. Pada tahun kedua, komposisi luas areal 90% dipengaruhi oleh luas areal itu sendiri, sekitar 8% dipengaruhi oleh volume ekspor, dan sekitar 2% dipengaruhi oleh produksi. Pada tahun ketiga, komposisi luas areal dipengaruhi sekitar 70% oleh luas areal itu sendiri, sekitar 20% dipengaruhi oleh volume ekspor, dan 5% dipengaruhi oleh harga karet dunia dan sekitar 5% oleh produksi karet. Semakin bertambahnya tahun, pengaruh luas areal karet terhadap keragaman luas areal karet sendiri semakin berkurang diikuti makin tingginya pengaruh volume ekspor, volume impor, harga karet dunia, harga karet sheet local dan produksi.



Gambar 20. Dekomposisi Keragaman Beberapa Variabel Untuk Model VAR(2) Both

#### D. Pemilihan Model Terbaik Estimasi Luas Areal Karet Nasional

Salah satu dasar penentuan model terbaik dapat dilakukan dengan membandingkan nilai MAPE untuk data testing dan training yaitu dengan memilih nilai MAPE yang paling kecil, terutama untuk data testing. Selain MAPE yang terkecil, pola pergerakan ramalan juga harus diperhatikan. Pilihlah plot yang paling berhimpit/bersesuaian dengan data asli/aktual atau dengan kata lain performa hasil ramalan seiring dengan data historisnya.

Berdasarkan data historis yang ada luas areal karet nasional berfluktuasi, luas areal tahun 2017 sebesar 3,66 juta ha atau naik 0,55%. Pada tahun 2018 dan 2019 luas areal karet nasional meningkat masing-masing sebesar 0,34% dan 0,13%, sehingga luas areal karet tahun 2018 menjadi sebesar 3,67 juta ha. Pada tahun 2019 luas areal karet nasional kembali naik sebesar 0,13%, kemudian pada tahun 2020 kembali naik sebesar 1,36%. Rata-rata pertumbuhan luas areal karet nasional selama 5 tahun terakhir atau tahun 2016 – 2020 sebesar 0,59% per tahun.

Tabel 40. Luas Areal Karet Nasional Tahun 2016 – 2020

Tahun	Luas Areal (Ha)	Pertumbuhan (%)
2016	3.639.048	
2017	3.659.090	0,55
2018	3.671.387	0,34
2019	3.676.035	0,13
2020	3.726.173	1,36
<b>Rata-rata Petumbuhan</b>		0,59

Sumber : Direktorat Jenderal Perkebunan

Untuk menyusun angka estimasi luas areal karet telah dilakukan uji coba dengan 4 (empat) model. Model yang pertama adalah model time series atau ARIMA, model terbaik untuk ARIMA adalah pada orde ARIMA (1,1,1) dan ARIMA (2,1,0). Untuk model estimasi luas areal karet nasional dengan ARIMA (1,1,1) menghasilkan MAPE untuk data training sebesar 1,28% dan MAPE untuk data testing sebesar 0,61%. Hal ini menunjukkan bahwa kemampuan untuk melakukan estimasi dengan model ARIMA ini rata-rata akan mengalami kesalahan sekitar 0,61% lebih tinggi atau 0,61% lebih rendah. Hasil estimasi dengan model ARIMA(1,1,1) pertumbuhan 5 tahun kedepan relatif lambat, yaitu hanya 0,36%/tahun. Hal ini mendekati dengan data 5 tahun ke belakang (2016 – 2020), dimana pertumbuhan luas areal mencapai 0,59% per tahun.

Model ARIMA (2,1,0) patut dipertimbangkan juga sebagai model tentative terbaik karena menghasilkan MAPE untuk data training sebesar 1,32% dan MAPE untuk data testing sebesar 0,45%. Model ARIMA (2,1,0) menghasilkan MAPE data testing yang lebih kecil dari ARIMA (1,1,1) artinya kemampuan untuk mengestimasi lebih akurat. Hasil estimasi dengan model ARIMA(2,1,0) pertumbuhan 5 tahun kedepan relatif lebih lambat, yaitu hanya 0,20%/tahun. Pertumbuhan ini lebih rendah dibandingkan dengan data 5 tahun ke belakang (2016 – 2020), dimana pertumbuhan luas areal mencapai 0,59% per tahun.

Metode estimasi yang kedua adalah dengan model fungsi transfer, untuk melakukan estimasi luas areal karet dengan variabel bebas adalah harga karet sheet nasional. Untuk model fungsi transfer ini menghasilkan MAPE data training 1,33%, sementara untuk MAPE data testing sebesar 0,35%. Model fungsi transfer ini menghasilkan MAPE yang lebih kecil dibandingkan model ARIMA meskipun perbedaan tidak terlalu signifikan, sehingga model fungsi transfer lebih akurat dalam melakukan estimasi. Hasil estimasi juga menunjukkan angka yang lebih realistis, dengan angka estimasi tahun 2021 sebesar 3,747 juta hektar, atau naik 0,56%. Disamping itu untuk estimasi 5 tahun kedepan angka pertumbuhan rata-rata sebesar 0,21%/tahun, sementara angka pertumbuhan 5 tahun sebelumnya sebesar 0,59%/tahun, pertumbuhan ini lebih rendah pertumbuhan data historisnya (5 tahun kebelakang).

Tabel 41. Perbandingan Hasil Estimasi dan MAPE Model Arima, Fungsi Transfer dan VAR

	Pengujian MAPE	Model ARIMA				Fungsi Transfer		Model VAR			
		ARIMA (1,1,1)	Growth (%)	ARIMA (2,1,0)	Growt h (%)	Arima (2,1,0) Xreg=Hsheet	Growt h (%)	VAR (2) type=both	Growth (%)	VAR (4) type=both	Growth (%)
	MAPE Training	1,28		1,32		1,33		1,07		0,61	
	MAPE Testing	0,61		0,45		0,35		9,51		3,43	
ATAP	2016	3.639.048		3.639.048		3.639.048		3.639.048		3.639.048	
	2017	3.659.090	0,55	3.659.090	0,55	3.659.090	0,55	3.659.090	0,55	3.659.090	0,55
	2018	3.671.387	0,34	3.671.387	0,34	3.671.387	0,34	3.671.387	0,34	3.671.387	0,34
	2019	3.676.035	0,13	3.676.035	0,13	3.676.035	0,13	3.676.035	0,13	3.676.035	0,13
	2020	3.726.173	1,36	3.726.173	1,36	3.726.173	1,36	3.726.173	1,36	3.726.173	1,36
Angka Estimasi (AESTI)	2021	3.745.592	0,52	3.747.437	0,57	3.747.113	0,56	3.787.518	1,65	3.882.613	4,20
	2022	3.762.331	0,45	3.761.485	0,37	3.762.496	0,41	3.827.540	1,06	3.891.241	0,22
	2023	3.776.758	0,38	3.769.526	0,21	3.771.038	0,23	3.896.807	1,81	3.885.618	(0,14)
	2024	3.789.194	0,33	3.774.326	0,13	3.775.459	0,12	3.967.502	1,81	3.888.491	0,07
	2025	3.799.914	0,28	3.777.154	0,07	3.778.752	0,09	4.034.359	1,69	3.944.120	1,43
Rata-rata	ATAP 2016 - 2020		0,59		0,59		0,59		0,59		0,59
Pertumbuhan	AESTI 2021 - 2025		0,36		0,20		0,21		1,59		0,40

Untuk model estimasi yang terakhir adalah dengan model VAR (Vector Auto Regressive). Untuk model VAR ini menggunakan 6 variabel yaitu luas areal, produksi, harga karet sheet, harga karet dunia, volume ekspor dan volume impor karet. Model yang terbaik untuk Model VAR ada dua yaitu adalah nilai  $p=2$  dan  $type="both"$ ,  $p=2$  artinya menggunakan variabel bebas sampai lag-2 dan ada konstanta dan trend. Estimasi luas areal karet dengan menggunakan model VAR(2) both ini menghasilkan ketelitian yang cukup tinggi yaitu MAPE untuk data training 1,07% dan MAPE untuk data testing 9,51%. MAPE untuk data testing ini model VAR lebih tinggi dibandingkan dengan model Fungsi Transfer atau Model Arima. Jika dibandingkan angka pertumbuhan luas areal karet model VAR(2) both antara hasil estimasi 5 tahun kedepan dengan rata-rata pertumbuhan 1,59% per tahun, berbeda dengan angka pertumbuhan 5 tahun terakhir yaitu sebesar 0,59% per tahun. Angka hasil estimasi untuk luas areal karet nasional tahun 2021 sebesar 3,79 juta hektar, sementara untuk angka tetap tahun 2020 sebesar 3,73 juta hektar atau naik sebesar 1,65%.

Model VAR yang kedua yang perlu dipertimbangkan adalah VAR(4) both. Model VAR(4) both menghasilkan MAPE data training sebesar 0,61%, dan MAPE data testing 3,43%. Angka pertumbuhan rata-rata selama 5 tahun ke depan (2021 -2025) untuk model ini sebesar 0,40%/tahun atau paling mendekati dengan data historisnya (2016 – 2020) dengan rata-rata pertumbuhan 0,59%/tahun. Kelemahan dari model VAR(4) both ini hasil estimasi luas areal pada tahun pertama (2021) yang mencapai 3,88 jta hektar atau naik 4.20% dibandingkan tahun 2020. Estimasi ini diduga over estimasi karena dari data historisnya pertumbuhan luas areal karet hanya sekitar 1% saja per tahun.

Berdasarkan Tabel 41 diatas, untuk data training dan data testing yang paling baik adalah yang memiliki MAPE terkecil. Untuk data training yang paling kecil adalah model VAR(4) both, sedangkan untuk data testing yang paling kecil adalah Fungsi Transfer ARIMA (2,1,0) xreg= Hsheet, dengan MAPE 0,35%. Oleh karena tujuan penyusunan model menghasilkan angka estimasi dengan kesalahan yang paling kecil, maka model yang paling kecil MAPE Testing adalah Model Fungsi Transfer ARIMA (2,1,0) dengan faktor input harga karet sheet merupakan model yang terbaik untuk menyusun angka Estimasi Luas Areal Karet nasional. Disamping faktor MAPE, hasil estimasi 5 tahun kedepan (2021 – 2025) menunjukkan

pertumbuhan rata-rata luas areal karet nasional 0,21% per tahun, paling mendekati dari data aktual 5 tahun kebelakang yaitu pertumbuhan luas areal karet tahun 2016 – 2020 sebesar 0,59% per tahun. Berdasarkan MAPE testing terkecil dan angka pertumbuhan luas areal maka model Fungsi Transfer menjadi model terbaik untuk meramalkan luas areal karet nasional.

Berdasarkan hasil kajian ini model terbaik berdasarkan MAPE testing terkecil yaitu Model Fungsi Transfer ARIMA (2,1,0). Namun demikian ada pertimbangan lain untuk memilih hasil estimasi yang paling tepat, misalnya intervensi program intensifikasi dan ekstensifikasi. Berdasarkan pertimbangan program yang telah dilakukan maka ada optimisme bahwa hasil peramalan dengan model VAR(2) “both” menjadi model estimasi terbaik dengan angka estimasi tahun 2021 sebesar 3,79 juta hektar dan estimasi tahun 2022 sebesar 3,83 juta hektar.

## KESIMPULAN

Untuk meningkatkan akurasi dalam penyusunan angka estimasi, maka dilakukan pengembangan metode estimasi luas areal karet nasional. Metode estimasi data perkebunan selama ini menggunakan model *Single Smoothing Exponential (SSE)* atau menggunakan *Double Smoothing Exponential (DSE)*. Meskipun dua metode tersebut dapat menghasilkan angka estimasi yang cukup baik, namun masih perlu melakukan pengembangan model alternatif yang diharapkan lebih akurat.

Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1971 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan. Untuk estimasi luas areal karet alternatif model pertama adalah Model ARIMA. Model ARIMA terbaik adalah ARIMA (2,1,0), menghasilkan MAPE untuk data training 1,32%, dan MAPE data testing 0,45%. Model ARIMA (1,1,1) juga menghasilkan MAPE yang cukup baik, yaitu MAPE training 1,28% dan MAPE testing 0,61%. Untuk model yang dua dengan menggunakan Fungsi Transfer ARIMA (2,1,0) dengan variabel input harga karet sheet, menghasilkan MAPE untuk data training sebesar 1,33% dan MAPE data testing 0,35%. Untuk model yang keempat model VAR(2) type ‘both’ ada pengaruh konstanta dan trend, menghasilkan MAPE data training 1,07% dan data MAPE data testing 9,51%. Model tentatif VAR adalah VAR(4) type ‘both’, menghasilkan MAPE training 0,61% dan MAPE testing 3,43%.

Berdasarkan perbandingan besarnya MAPE baik data testing maupun data training dan hasil estimasi luas areal 5 tahun kedepan, maka model terbaik yang terpilih adalah model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet sheet karena menghasilkan akurasi yang paling tinggi sehingga MAPE rata-rata data testing sebesar 0,35%. Hasil estimasi luas areal karet nasional untuk model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet sheet untuk tahun 2021 sebesar 3.747.113 hektar dan tahun 2022 sebesar 3.762.496 hektar, tahun 2023 sebesar 3.771.038 hektar, tahun 2024 sebesar 3.775.459 hektar, dan tahun 2025 sebesar 3.778.752 hektar. Laju pertumbuhan estimasi luas areal karet nasional selama 5 tahun kedepan (2021 – 2025) rata-rata 0,21% per tahun. Namun jika mempertimbangkan program yang telah dijalankan oleh direktorat teknis terkait untuk meningkatkan areal karet maka ada optimisme estimasi luas karet tahun 2021 sebesar 3.787.518 hektar dan tahun 2022 sebesar 3.827.540 hektar seperti hasil pemodelan dengan VAR(2) type “both”.

## DAFTAR PUSTAKA

- Anonim. 2013. Pedoman Pelaksanaan Pengelolaan Data Komoditas Perkebunan (PDKP). Dirjen Perkebunan – Kementerian Pertanian.
- Enders, W. 2010. Applied Econometric Time Series. USA: University of Alabama. Wiley, Third Edition.
- Fitriani, D.R, Darsyah, M.Y., & Wasono, R. 2013. Peramalan Fungsi Transfer pada Harga Emas Pasar Komoditi. Seminar Nasional Pendidikan Sains dan Teknologi, Fakultas MIPA, Universitas Muhammadiyah Semarang.
- Guha, B and Bandyopadhyay, G. 2016. Gold Price Forecasting Using ARIMA Model. Journal of Advanced Management Science Vol. 4, No. 2, March 2016
- Gujarati, D.N. and D.C. Porter, 2010. Dasar-dasar Ekonometrika. Jakarta: Penerbit Salemba Empat, Buku 2, Edisi 5.
- Kementerian Pertanian. 2019. Statistik Perkebunan Indonesia 2018-2020 (Karet). Jakarta: Direktorat Jenderal Perkebunan.
- M. Firdaus 2020. Aplikasi Ekonometrika dengan E-Views, Stata dan R. Institut Pertanian Bogor (IPB) Press.
- Montgomery DC, Johnson LA & Gardiner JS. 1990. Forecasting and Time Series Analysis. Singapore:Mc-Graw Hill.
- Myers R. 1994. Classical And Modern Regression with Applications. Boston: PWS – KENT Publishing Company.
- Ryan TP. 1997. Modern Regression Methods. New York,USA: John Wiley & Sons, INC.

# **KAJIAN MODEL PERAMALAN LUAS AREAL KOPI DI INDONESIA: PENDEKATAN MODEL ARIMA, FUNGSI TRANSFER, DAN VAR**

Ucik Mawarsari

*Direktorat Statistik Tanaman Pangan, Hortikultura dan Perkebunan-Badan Pusat Statistik*

*Jalan Dr. Sutomo no. 6-8 Gd 4 Lt 4, Pasar Baru-Jakarta Pusat, Indonesia*

*E-mail: ucik@bps.go.id*

## **ABSTRAK**

Kopi merupakan salah satu komoditas perkebunan yang memegang peranan cukup penting dalam perekonomian Indonesia. Kontribusi penting dari komoditas kopi bagi perekonomian nasional tercermin pada kinerja perdagangan dan peningkatan nilai tambahnya. Selain peluang ekspor yang semakin terbuka, pasar kopi di dalam negeri masih cukup besar. Pengembangan lahan usaha budidaya kopi perlu dilakukan dengan mempertimbangkan aspek berkelanjutan. Guna merumuskan kebijakan terkait usaha budidaya kopi, pasokan dan distribusi komoditas kopi sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Penelitian ini akan mengkaji tiga metode yaitu *Auto-Regressive Integrated Moving Average* (ARIMA), fungsi transfer, dan *Vector Auto Regression* (VAR) dengan menggunakan *software* RStudio guna melakukan pemodelan luas areal kopi di Indonesia. Metode yang dipilih guna melakukan peramalan data luas areal kopi adalah metode yang terbaik dikaji dari nilai *Mean Absolute Percentage Error* (MAPE) terkecil dan kerealistikan hasil peramalan. Berdasarkan dari hasil perbandingan MAPE disimpulkan model terbaik adalah model ARIMA yang memiliki MAPE terkecil yaitu 0,40%.

## **ABSTRACT**

*Coffee is a commodity that plays an important role in the Indonesian economy. The important contribution of the coffee commodity to the national economy is reflected in trade performance and increased value added. Apart from being an export commodity, domestic coffee market also has considerable opportunities. The development of coffee land area needs to be carried out by considering sustainable aspects. In order to formulate policies related to coffee cultivation, supply and distribution of coffee, the availability of up-to-date data, and even forecasts for the next few periods, is urgently required. This study will examine three methods, namely the Auto-Regressive Integrated Moving Average (ARIMA), transfer function, and Vector Auto Regression (VAR) using RStudio software to model the area of coffee in Indonesia. The method chosen to forecast coffee area data is the best method studied from the smallest Mean Absolute Percentage Error (MAPE) value and the reliability of the forecasting results. Based on the results of the MAPE comparison, it can be concluded that the best model is the ARIMA model having the smallest MAPE, namely 0.40%.*

## PENDAHULUAN

Kopi merupakan salah satu komoditas perkebunan yang memegang peranan cukup penting dalam perekonomian Indonesia. Kontribusi penting dari komoditas kopi bagi perekonomian nasional tercermin pada kinerja perdagangan dan peningkatan nilai tambahnya. Sebagai produk ekspor, komoditas kopi dapat memberikan kontribusi berupa penghasil devisa dan pendapatan negara, sumber pendapatan petani, penciptaan lapangan kerja, pendorong pertumbuhan sektor agribisnis dan agroindustri, pengembangan wilayah serta pelestarian lingkungan. Selain peluang ekspor yang semakin terbuka, pasar kopi di dalam negeri masih cukup besar.

Pengembangan lahan usaha budidaya kopi perlu dilakukan dengan mempertimbangkan aspek berkelanjutan. Kegiatan budidaya/pasca panen/pengolahan kopi harus sesuai *Standard Operating Procedure* (SOP), sesuai *Good Agricultural Practices* (GAP), dan memperhatikan analisis resiko terhadap lingkungan. Guna merumuskan kebijakan terkait usaha budidaya kopi, pasokan dan distribusi komoditas kopi sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Makalah ini akan mengkaji beberapa metode statistik yang telah dikembangkan oleh para ahli, diantaranya metode *Auto-Regressive Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Auto Regression* (VAR) dalam melakukan pemodelan dan peramalan luas areal kopi Indonesia.

Oleh karenanya, tujuan dari disusunnya makalah ini adalah:

- c. Melakukan analisis dan peramalan data luas areal kopi di Indonesia menggunakan model ARIMA, Fungsi Transfer dan VAR .
- d. Membandingkan metode tersebut dalam memperoleh ramalan data luas areal komoditas kopi.
- e. Menentukan metode terbaik dalam meramal data luas areal komoditas kopi di Indonesia.

## BAHAN DAN METODE

Data yang digunakan dalam melakukan pemodelan luas areal kopi Indonesia adalah data series tahun 1980 sampai 2020. Peubah yang diasumsikan mempengaruhi besaran luas areal kopi untuk Fungsi Transfer adalah volume ekspor kopi, sedangkan untuk pemodelan VAR luas areal kopi, peubah yang diasumsikan mempengaruhi adalah produksi, volume ekspor, dan harga kopi dunia. Adapun harga kopi dunia yang digunakan adalah harga kopi robusta karena merupakan jenis kopi mayoritas di Indonesia. Peramalan data luas areal kopi dilakukan melalui pengujian beberapa metode yakni metode ARIMA, Fungsi Transfer, dan VAR menggunakan *software Rstudio*.

Metode yang digunakan dalam melakukan peramalan data produksi dan luas areal kopi di Indonesia adalah sbb.

- f. *Autoregressive Integrated Moving Average* (ARIMA)

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:



$Y_t$	= data <i>time series</i> sebagai variable dependen pada waktu ke-t
$Y_{t-p}$	= data <i>time series</i> pada kurun waktu ke $(t-P)$
$\mu$	= suatu konstanta
$\theta_1\theta_q\phi_1\phi_n$	= parameter-parameter model
$\varepsilon_{t-q}$	= nilai sisaan pada waktu ke- $(t-q)$

g. Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut.

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t$$

Dimana:

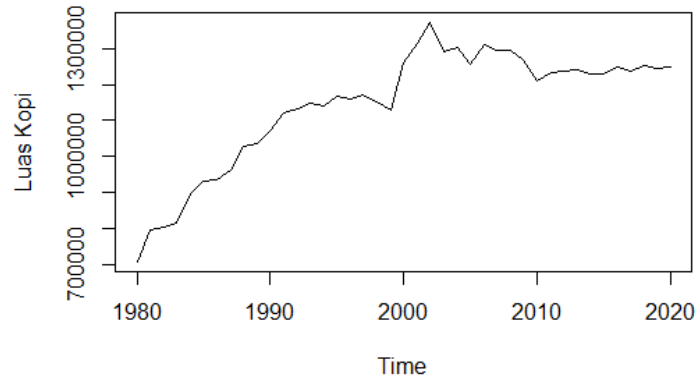
- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang *lag*  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$

h. VAR (*Vector Auto Regression*)

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen, karena pada kenyataannya suatu variabel dapat bertindak sebagai endogen atau dependen dari suatu variabel yang lainnya (Gujarati, 2010).

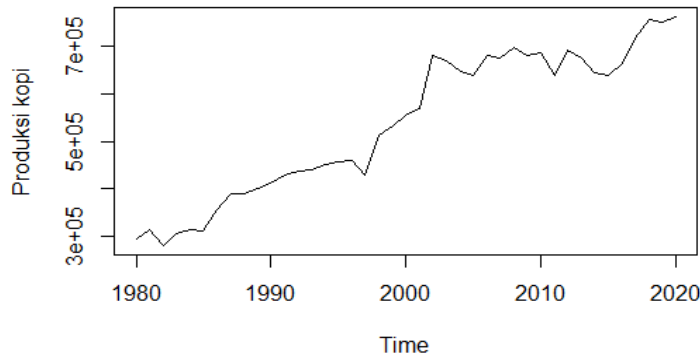
## HASIL DAN PEMBAHASAN

### a. Eksplorasi Data



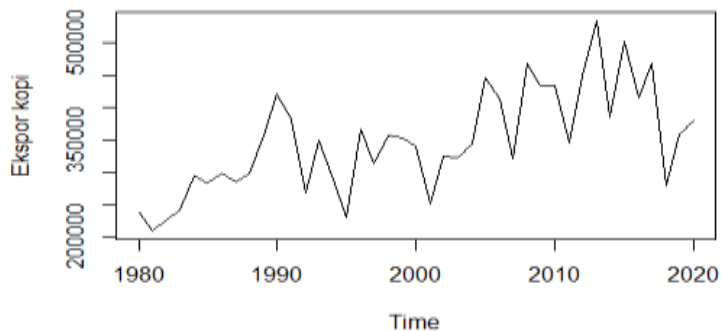
**Gambar 1.** Perkembangan Luas Areal Kopi Indonesia Tahun 1980-2020

Luas areal kopi cukup berfluktuasi dengan kecenderungan meningkat hingga tahun 2002 dan menurun setelahnya (Gambar 1). Berkurangnya luas areal kopi disebabkan oleh banyak faktor, di antaranya alih fungsi lahan atau penggantian jenis tanaman.



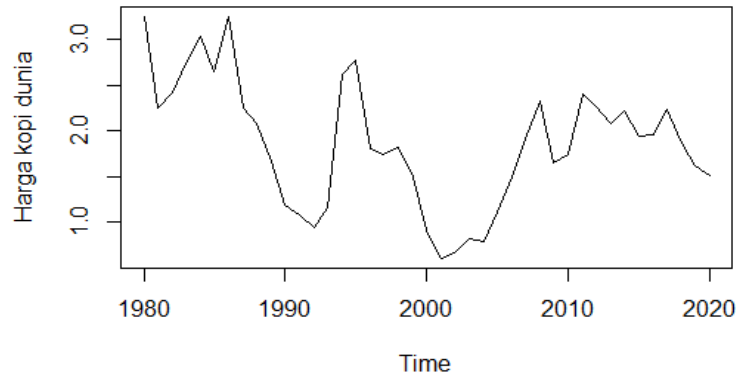
**Gambar 2.** Perkembangan Produksi Kopi Indonesia Tahun 1980-2020

Produksi tahunan kopi di Indonesia periode 1980-2020 berfluktuasi setiap tahunnya disebabkan berbagai macam faktor namun cenderung meningkat dengan rata-rata pertumbuhan sebesar 2,58% (Gambar 2)



**Gambar 3.** Perkembangan Volume Ekspor Kopi Indonesia tahun 1980-2020

Volume Ekspor Kopi Indonesia periode 1980-2020 sangat berfluktuasi setiap tahunnya (Gambar 3). Bahkan sulit untuk memprediksi arah kecenderungan volume ekspor kopi disebabkan oleh fluktuasi yang sangat besar.



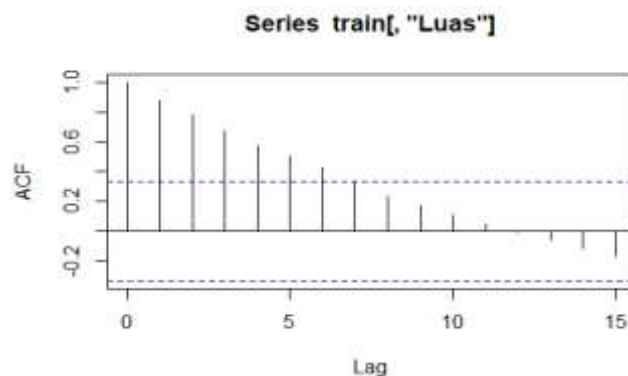
**Gambar 4.** Perkembangan Harga Kopi Dunia tahun 1980-2019

Harga kopi robusta dunia secara tahunan periode 1980-2020 sangat berfluktuasi setiap tahunnya disebabkan berbagai macam faktor (Gambar 4). Harga kopi dunia diperkirakan turut memengaruhi minat petani dalam menanam dan melakukan budidaya tanaman kopi.

#### b. ARIMA

Dalam melakukan pemodelan luas areal kopi menggunakan model *Autoregressive Integrated Moving Average* (ARIMA), data yang digunakan adalah periode tahun 1980 hingga 2020. Periode data tersebut kemudian dipisahkan menjadi data set *training* dan *testing*. Panjang series data pada data set *training* adalah tahun 1980 hingga 2014, sementara dataset *testing* adalah periode tahun 2015-2020. Dataset *training* digunakan untuk melakukan penyusunan model, sementara dataset *testing* digunakan untuk validasi model.

Uji kestasioneran data seperti yang disyaratkan apabila melakukan pemodelan ARIMA dilakukan secara visual menggunakan hasil plot data maupun uji formal statistik. Gambar 5 menunjukkan plot *Auto Correlation Function* (ACF) dari data *training*. Terlihat bahwa plot ACF data *training* memiliki pola tertentu sehingga diperkirakan data yang digunakan tidak stasioner. Hal ini diperkuat oleh hasil uji formal statistik yaitu dengan uji *Augmented Dickey-Fuller* yang mengindikasikan bahwa data luas areal kopi adalah tidak stasioner, terlihat dari nilai absolut dari statistik uji (1,5042) yang lebih kecil dari nilai absolut dari nilai kritis taraf 5% yakni 3,50 (Tabel 1).



**Gambar 5.** Plot *Auto Correlation Function* (ACF) Data *Training* Luas areal Kopi

**Tabel 1. Hasil Uji Augmented Dickey-Fuller Luas Areal Kopi**

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-64842 -18627   -342   10884 115173

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.313e+05  6.287e+04   2.088  0.0441 *
z.lag.1     -1.013e-01  6.735e-02  -1.504  0.1415
tt          -7.524e+01  9.435e+02  -0.080  0.9369
z.diff.lag  -1.017e-01  1.591e-01  -0.639  0.5268
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 36050 on 35 degrees of freedom
Multiple R-squared:  0.1694,    Adjusted R-squared:  0.09816
F-statistic: 2.379 on 3 and 35 DF,  p-value: 0.08639

Value of test-statistic is: -1.5042 3.5627 3.5667

Critical values for test statistics:
            1pct  5pct 10pct
tau3      -4.15  -3.50 -3.18
phi2       7.02   5.13  4.31
phi3       9.31   6.73  5.61
```

Untuk mengatasi masalah ketidak-stasioneran, dilihat dari pola grafiknya maka metode yang dilakukan adalah melakukan *differencing*. Hasil pengujian *Augmented Dickey-Fuller test* pada Tabel 2, diperoleh hasil bahwa data *differencing* luas areal kopi telah stasioner. Hal tersebut diketahui dari nilai absolut dari statistik uji (3,4133) yang lebih besar dari nilai absolut dari nilai kritis taraf 5% yakni 1,95.

**Tabel 2. Hasil Uji Augmented Dickey-Fuller Difference Luas Areal Kopi**

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-95408 -10075   2769  23684 138796

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1     -0.7221    0.2116  -3.413  0.0016 **
z.diff.lag  -0.1977    0.1532  -1.291  0.2050
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 39700 on 36 degrees of freedom
Multiple R-squared:  0.4732,    Adjusted R-squared:  0.444
F-statistic: 16.17 on 2 and 36 DF,  p-value: 9.757e-06

Value of test-statistic is: -3.4133

Critical values for test statistics:
            1pct  5pct 10pct
tau1      -2.62  -1.95 -1.61
```

Setelah mendapatkan data luas areal kopi yang stasioner, maka langkah selanjutnya adalah menentukan ordo AR dan MA menggunakan fungsi auto arima yang terdapat pada RStudio. Dari hasil pendugaan auto arima pada Tabel 3, maka model ARIMA yang terpilih adalah ARIMA (0,2,1) yang berarti model ARIMA yang optimal untuk mengestimasi data yang diolah adalah ARIMA dengan ordo *Auto Regressive* (AR) = 0, ordo *differencing* = 2, dan ordo *Moving Average* (MA) = 1.

**Tabel 3.** Hasil Fungsi Auto ARIMA Luas Areal Kopi

---

```

Series: train[, "Luas"]
ARIMA(0,2,1)

Coefficients:
      ma1
    -0.8753
s.e.    0.0825

sigma^2 estimated as 1.871e+09: log likelihood=-399.31
AIC=802.63 AICC=803.03 BIC=805.62

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -11133.78 41359.79 30549 -1.061169 2.695648 0.9300218 -0.1234001

```

---

Namun model ARIMA yang dipilih oleh fungsi autoarima dirasa kurang tepat karena pada model terpilih, ordo *differencing* = 2 sementara data luas areal kopi menurut perhitungan uji ADF telah stasioner pada ordo *differencing* = 1. Oleh karena itu, dilakukan *overfitting* menggunakan fungsi *armaselect* untuk mencoba ordo lain agar dapat menemukan model ARIMA terbaik yang dapat digunakan untuk melakukan peramalan luas areal kopi Indonesia.

**Tabel 4.** Ordo ARIMA Optimum menggunakan Fungsi *Armaselect* (*diff*)

---

	p	q	sbc
[1,]	5	5	705.7470
[2,]	4	5	713.4449
[3,]	3	5	717.0580
[4,]	<b>0</b>	<b>5</b>	<b>718.4803</b>
[5,]	1	5	719.2170
[6,]	2	5	721.3790
[7,]	0	0	724.5414
[8,]	1	0	725.7172
[9,]	2	0	729.7920
[10,]	4	0	731.3535

---

Dari hasil pencarian ordo ARIMA optimum menggunakan fungsi *armaselect* (*diff*), maka diperoleh beberapa kombinasi AR dan MA. Dari beberapa alternatif tersebut diperoleh kandidat model terbaik yaitu ARIMA (0,1,5).

**Tabel 5.** Hasil pengujian ARIMA (0,1,5) untuk Luas Areal Kopi

---

```

arma(x = train[, "Luas"], order = c(0, 1, 5))

Coefficients:
      ma1      ma2      ma3      ma4      ma5
0.2053  0.1019 -0.1873  0.3436  0.0341
s.e.    0.1898  0.1924  0.1567  0.1703  0.2359

sigma^2 estimated as 1.641e+09: log likelihood = -409.41, aic = 830.82

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF
Training set 10199.172 39924.166 26688.065 1.0336691 2.4012587 0.8124810 -0.116415
Test set     -3879.626  6688.925  4966.895 -0.3144797 0.4012693 0.1512102      N

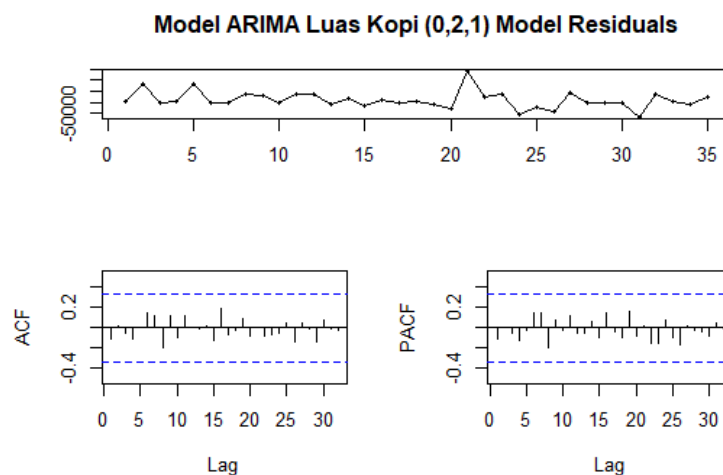
```

---

Hasil uji statistik terhadap model ARIMA (0,1,5) menunjukkan bahwa nilai MAPE data *training* sebesar 2,40% dan untuk data *testing* sebesar 0,40%. Nilai tersebut dinilai telah cukup baik untuk melakukan peramalan dan hasil uji Ljung-Box menunjukkan bahwa tidak terdapat autokorelasi pada residual model (Tabel 6) dan hasil pemeriksaan residual yang sudah cukup stasioner pada rata-rata sebagaimana terlihat pada Gambar 6.

**Tabel 6.** Hasil Uji Ljung-Box test ARIMA (2,1,0) untuk Luas Areal Kopi

lags	statistic	df	p-value
5	1.174840	5	0.9472682
10	5.816989	10	0.8303975
15	7.603219	15	0.9387039
20	11.676747	20	0.9267529
25	13.470901	25	0.9701908
30	22.102148	30	0.8503024



**Gambar 6.** Grafik Pemeriksaan Model *Residual* ARIMA pada Data *Training*

Setelah diperoleh model ARIMA yang optimal, maka model ARIMA tersebut digunakan untuk melakukan peramalan luas areal kopi 5 tahun ke depan (2021-2025). Hasil peramalan luas areal kopi menggunakan ARIMA (0,1,5) ditampilkan pada Tabel 7.

**Tabel 7.** Hasil Peramalan Luas areal Kopi Indonesia menggunakan ARIMA (0,1,5)

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	1250203	1198954	1301451	1171825	1328580
2022	1252694	1172958	1332431	1130747	1374642
2023	1251091	1146940	1355243	1091806	1410377
2024	1251689	1133223	1370155	1070511	1432867
2025	1251850	1111839	1391861	1037721	1465978

### c. Fungsi Transfer

Peubah yang diasumsikan dapat mempengaruhi gejala produksi kopi adalah volume ekspor kopi, dan peubah ini akan digunakan untuk melakukan analisis menggunakan model Fungsi Transfer. Dari pola data volume ekspor dan uji *Augmented Dickey-Fuller* (Tabel 8) terlihat data volume ekspor belum stasioner pada taraf uji 5%. Oleh karena itu, series data volume ekspor dilakukan *differencing* terlebih dahulu sebelum melakukan pemodelan (Tabel 9).

**Tabel 8. Hasil Uji Augmented Dickey-Fuller Series Data Volume Ekspor Kopi**

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-108011  -43173   -2845   41078  108739

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.463e+05  6.258e+04   3.936 0.000476 ***
z.lag.1      -9.689e-01  2.460e-01  -3.939 0.000472 ***
tt           4.918e+03  1.677e+03   2.933 0.006496 **
z.diff.lag   1.121e-01  2.001e-01   0.560 0.579598

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56370 on 29 degrees of freedom
Multiple R-squared:  0.4462,    Adjusted R-squared:  0.3889
F-statistic: 7.788 on 3 and 29 DF,  p-value: 0.0005811

Value of test-statistic is: -3.9391 5.4722 7.7956

Critical values for test statistics:
            1pct  5pct 10pct
tau3      -4.15  -3.50  -3.18
phi2       7.02   5.13   4.31
phi3       9.31   6.73   5.61
```

**Tabel 9. Hasil Uji Augmented Dickey-Fuller Series Data Differencing Volume Ekspor Kopi**

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-103909  -7267   16774   56291  112134

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1      -1.9942    0.2859  -6.976 9.48e-08 ***
z.diff.lag   0.4561    0.1781   2.562 0.0157 *

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 62810 on 30 degrees of freedom
Multiple R-squared:  0.7227,    Adjusted R-squared:  0.7042
F-statistic: 39.1 on 2 and 30 DF,  p-value: 4.404e-09

Value of test-statistic is: -6.9756

Critical values for test statistics:
            1pct  5pct 10pct
tau1      -2.62  -1.95  -1.61
```

Pencarian ordo model ARIMA input untuk volume ekspor dilakukan dengan fungsi autoarima. Dari hasil fungsi autoarima, diperoleh ordo ARIMA terbaik untuk volume ekspor adalah ARIMA (2,1,0) sebagaimana ditampilkan pada Tabel 10. Untuk memperoleh beberapa

alternatif model, dilakukan *overfitting* menggunakan fungsi *armaselect* untuk mencoba ordo lain agar dapat menemukan model ARIMA terbaik (Tabel 11).

**Tabel 10.** Hasil Fungsi Auto ARIMA Volume Ekspor Kopi

---

```

Series: train.h[, "Ekspor"]
ARIMA(2,1,0)

Coefficients:
      ar1      ar2
    -0.5244  -0.4293
s.e.    0.1621   0.1605

sigma^2 estimated as 3.72e+09:  log likelihood=-422.12
    AIC=850.24   AICC=851.04   BIC=854.82

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 11419.84 58322.28 45646.03  1.397478 13.35381 0.8131972 -0.09298098
Test set    -39527.16 85920.87 77795.81 -13.937071 21.75022 1.3859550      NA

```

---

**Tabel 11.** Ordo ARIMA Optimum dengan Fungsi *Armaselect (diff)* Volume Ekspor Kopi

---

```

      p q      sbc
[1,] 5 5 708.5966
[2,] 4 5 745.3868
[3,] 0 2 745.6801
[4,] 0 1 748.1174
[5,] 1 2 748.4482
[6,] 2 1 748.5899
[7,] 0 3 749.2683
[8,] 1 1 749.2974
[9,] 2 2 751.6944
[10,] 3 1 752.0362

arima(x = train.h[, "Ekspor"], order = c(1, 1, 2))

Coefficients:
      ar1      ma1      ma2
    -0.5749  0.0453  -0.4888
s.e.    0.4182  0.3725  0.2133

sigma^2 estimated as 3.513e+09:  log likelihood = -422.17,  aic = 852.34

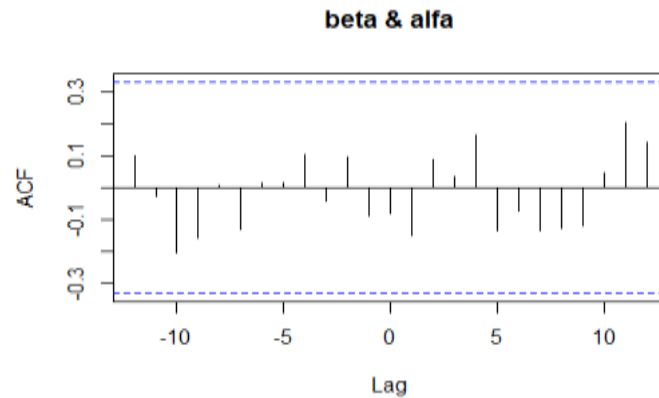
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 15547.94 58418.03 47688.10  2.516299 13.83838 0.8495773 -0.1452724
Test set    -25878.08 79976.70 68394.12 -10.510413 19.19100 1.2184611      NA

```

---

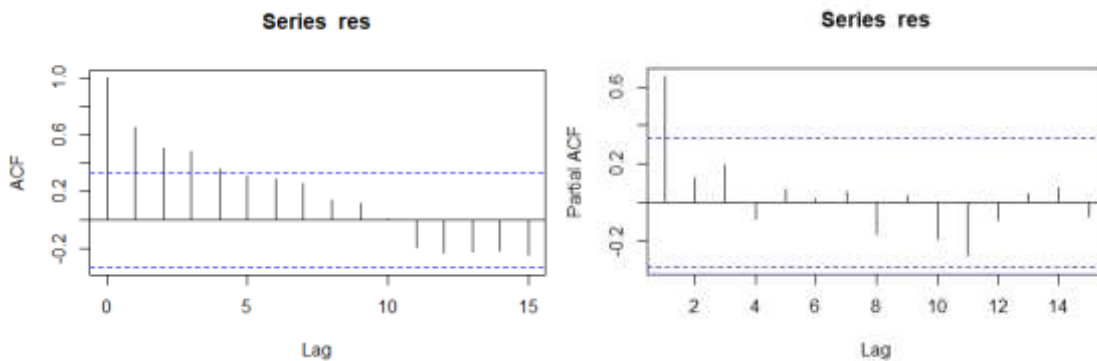
Dari hasil pencarian ordo arima optimum volume ekspor kopi menggunakan fungsi *autoarima* dan fungsi *armaselect*, diputuskan model arima input volume ekspor kopi adalah ARIMA (1,1,2) dengan nilai MAPE data *training* sebesar 13,83% dan MAPE data *testing* sebesar 19,19%. Tahap selanjutnya adalah melakukan *prewhitening* dan analisis korelasi silang antara residual model arima ekspor kopi dan luas areal kopi menggunakan ARIMA (1,1,2). Dari hasil plot *ccf (cross correlation function)* antara alfa (residual ekspor kopi) dan beta (residual luas kopi), diperoleh hasil bahwa tidak terdapat *lag* dari pengaruh yang diberikan oleh ekspor kopi terhadap luas areal kopi atau *lag* = 0. Hal tersebut dapat disimpulkan dari tidak adanya *bar* atau garis yang melewati ambang batas pada nilai *lag* positif (Gambar 7).



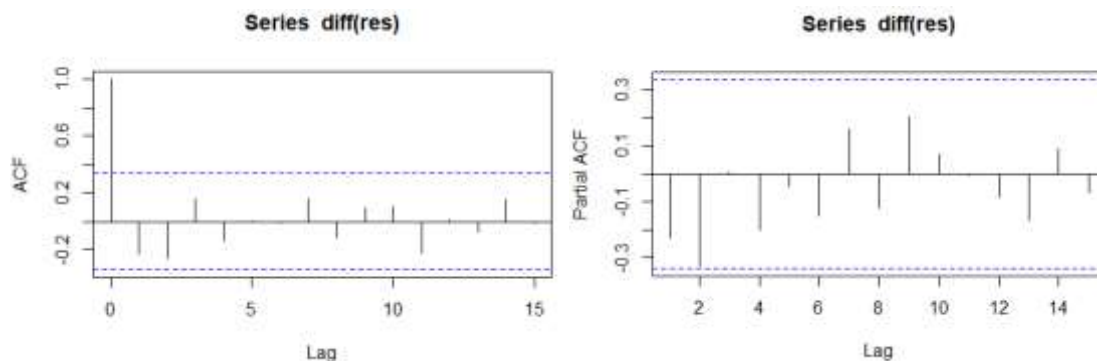


**Gambar 7.** Plot ACF Alfa (Residual Ekspor Kopi) dan Beta (Residual Luas Areal Kopi)

Karena tidak adanya *bar* atau garis yang melewati ambang batas pada nilai *lag* positif, maka model ARIMA untuk pengepasan model fungsi transfer adalah ARIMA (0,0,0). Kemudian residual model ARIMA (0,0,0) tersebut diteliti stasioneritasnya. Dari plot ACF dan PCF residual, diperoleh hasil bahwa residual model tidak stasioner (Gambar 8) sehingga perlu dilakukan *differencing*. Plot ACF dan PACF residual yang telah dilakukan *differencing* dapat dilihat pada Gambar 9.



**Gambar 8.** Plot ACF dan PACF Residual model ARIMA (0,0,0)



**Gambar 9.** Plot ACF dan PACF Residual Model ARIMA (0,0,0) setelah *Differencing*

Setelah data residual stasioner, maka dilakukan pencarian model ARIMA terbaik untuk residual tersebut atau dapat disebut dengan *noise*. Hasil dari fungsi autoarima menunjukkan bahwa model ARIMA (0,1,1) merupakan model terbaik (Tabel 12). Untuk memperoleh referensi ordo ARIMA lain, dilakukan fungsi *armaselect* sebagaimana pada Tabel 13.

**Tabel 12.** Hasil Fungsi Auto ARIMA *Noise*

---

```

Series: res
ARIMA(0,1,1)

Coefficients:
      ma1
      -0.4632
s.e.    0.1791

sigma^2 estimated as 1.008e+10:  log likelihood=-439.43
      AIC=882.85   AICC=883.24   BIC=885.9

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 13717.88 97471.42 78345.71 427.4522 581.3974 0.8854459 0.06600465

```

---

**Tabel 13.** Ordo ARIMA Optimum *Noise* menggunakan Fungsi Armselect (diff)

---

```

      p q      sbc
[1,] 5 5 785.4539
[2,] 0 0 787.2756
[3,] 1 0 788.4917
[4,] 2 0 788.6585
[5,] 4 5 790.9834
[6,] 3 0 793.2125
[7,] 4 0 796.3863
[8,] 0 2 796.7806
[9,] 1 2 797.2912
[10,] 0 1 798.0952

```

---

Setelah dilakukan uji signifikansi model ARIMA dan ARIMA-xreg pada pilihan ordo arima di atas, diperoleh ordo ARIMA *noise* yang signifikan dan menghasilkan MAPE terkecil adalah ARIMA (4,1,5) sebagaimana pada Tabel 14 berikut.

**Tabel 14.** Hasil Uji Signifikansi ARIMA-xreg (4,1,5)

---

```

Series: res
ARIMA(4,1,5)

Coefficients:
      ar1      ar2      ar3      ar4      ma1      ma2      ma3      ma4      ma5
0.6294 -0.4972 0.3840 0.0099 -1.0533 0.5464 -0.1545 -0.6156 0.7524
s.e.    0.2601 0.2985 0.3156 0.2953 0.2792 0.3150 0.3295 0.3034 0.2548

sigma^2 estimated as 8.974e+09:  log likelihood=-435.46
      AIC=890.92   AICC=900.48   BIC=906.18

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 2129.094 80061.24 61187.28 253.907 552.3165 0.6915251 -0.006390868

```

z test of coefficients:

```

      Estimate Std. Error z value Pr(>|z|)
ar1 0.6293565 0.2601079 2.4196 0.0155377 *
ar2 -0.4972021 0.2984767 -1.6658 0.0957535 .
ar3 0.3839853 0.3156165 1.2166 0.2237489
ar4 0.0098789 0.2952924 0.0335 0.9733120
ma1 -1.0532867 0.2791808 -3.7728 0.0001614 ***
ma2 0.5464287 0.3149920 1.7347 0.0827871 .
ma3 -0.1545285 0.3295355 -0.4689 0.6391210
ma4 -0.6155596 0.3033578 -2.0292 0.0424426 *
ma5 0.7523739 0.2547974 2.9528 0.0031487 **

```

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

---

Oleh karena telah didapatkan model input serta *noise* yang terbaik untuk peramalan, maka langkah selanjutnya adalah melakukan peramalan. Peramalan dilakukan menggunakan dua jenis data input, yakni nilai aktual dan data ramalan. Dari hasil peramalan menggunakan nilai aktual sebagai input, diperoleh nilai MAPE sebesar 0,39% (Tabel 15) sementara peramalan menggunakan data ramalan sebagai input memiliki MAPE sebesar 0,59% (Tabel 16).

**Tabel 15.** Pengukuran Tingkat Kesalahan Model Fungsi Transfer dengan Nilai Aktual sebagai Input

---

Series: test.h[, "Luas"]  
Regression with ARIMA(4,1,5) errors

AIC=107.26    AICc=108.59    BIC=106.87

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	618.7518	6471.5	4889.756	0.05047314	<b>0.3914401</b>	0.4728523	-0.09297311

---

**Tabel 16.** Pengukuran Tingkat Kesalahan Model Fungsi Transfer dengan Data Ramalan sebagai Input

---

Series: test.h[, "Luas"]  
Regression with ARIMA(4,1,5) errors

AIC=109.9    AICc=111.24    BIC=109.51

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	762.8562	8417.874	7336.402	0.06011316	<b>0.5888754</b>	0.7094494	-0.5568461

---

Setelah diperoleh model fungsi transfer yang optimal, maka model tersebut digunakan untuk melakukan peramalan luas areal kopi 5 tahun ke depan (2021-2025). Hasil peramalan luas areal kopi Indonesia menggunakan fungsi transfer dengan volume ekspor kopi sebagai peubah ditampilkan pada Tabel 17.

**Tabel 17.** Hasil Peramalan Luas Areal Kopi menggunakan Fungsi Transfer

---

Time Series:  
Start = 2021  
End = 2025  
Frequency = 1

**[1] 1263206 1251595 1252379 1256985 1256287**

---

#### **d. Model Vector Auto Regression (VAR)**

Pemodelan dengan menggunakan model VAR dilakukan dengan memasukkan peubah luas areal kopi, produksi kopi, volume ekspor kopi, dan harga kopi dunia. Hal ini diasumsikan bahwa apabila terjadi gejolak pada variabel-variabel tersebut, akan mempengaruhi gejolak luas areal kopi ataupun sebaliknya.

Pada tahap pertama, perlu dilakukan pemeriksaan apakah terdapat pengaruh trend dalam model. Hasil pengujian signifikansi trend dilakukan pada  $lag=1$  dan  $lag=2$ . Hasil pengujian trend pada  $lag=1$  menunjukkan bahwa *trend* berpengaruh signifikan pada persamaan produksi kopi (signifikan 10%), sementara hasil pengujian *trend* pada  $lag=2$  menunjukkan bahwa *trend* tidak berpengaruh signifikan. Dari hasil *running* model menggunakan  $lag$  1 dan 2, diperoleh informasi komponen *trend* tidak terlalu berpengaruh signifikan sehingga *trend* dikeluarkan dalam model untuk *running* model VAR (p) selanjutnya.

Langkah selanjutnya adalah menentukan  $lag$  (p) yang akan digunakan dalam model VAR. Penentuan  $lag$  (p) dilakukan dengan melakukan pengujian VAR (1) hingga VAR (4) dengan

pertimbangan kecukupan series data. Dari hasil *running* menggunakan p=1 hingga p=4, diperoleh informasi bahwa variabel yang signifikan cukup banyak pada p=1, sehingga model VAR yang digunakan selanjutnya adalah VAR (1). Hal tersebut diperkuat juga dengan nilai *Adjusted R-squared*: 0,9396 atau dapat dikatakan bahwa 93.96% keragaman luas areal kopi dapat dijelaskan oleh ketiga variabel/peubah penjelas termasuk konstanta. Hasil pengujian model VAR(1) dapat dilihat dalam Tabel 18.

**Tabel 18.** Hasil Estimasi VAR Luas Areal Kopi dengan Peubah Produksi, Volume Ekspor, dan Harga Kopi Dunia, pada Lag=1 tanpa *Trend*

---

```

VAR Estimation Results:
=====
Endogenous variables: Luas, Produksi, Ekspor, world.Price
Deterministic variables: const
Sample size: 34
Log Likelihood: -1236.202
Roots of the characteristic polynomial:
0.8731 0.8106 0.6079 0.2299
Call:
VAR(y = serieskopi[1:35, c(2, 3, 4, 5)], p = 1, type = "const")

Estimation results for equation Luas:
=====
Luas = Luas.l1 + Produksi.l1 + Ekspor.l1 + world.Price.l1 + const

      Estimate Std. Error t value Pr(>|t|)
Luas.l1      8.181e-01 1.107e-01  7.389 3.84e-08 ***
Produksi.l1  3.381e-02 1.282e-01  0.264  0.7938
Ekspor.l1    2.330e-02 1.214e-01  0.192  0.8492
world.Price.l1 -1.331e+04 1.264e+04 -1.053  0.3009
const        2.194e+05 9.789e+04  2.241  0.0329 *

Estimation results for equation Produksi:
=====
Produksi = Luas.l1 + Produksi.l1 + Ekspor.l1 + world.Price.l1 + const

      Estimate Std. Error t value Pr(>|t|)
Luas.l1      1.658e-01 8.529e-02  1.944  0.0616 .
Produksi.l1  8.243e-01 9.878e-02  8.345 3.38e-09 ***
Ekspor.l1    -1.139e-01 9.356e-02 -1.217  0.2335
world.Price.l1 3.106e+03 9.734e+03  0.319  0.7519
const        -5.472e+04 7.542e+04 -0.726  0.4739

Estimation results for equation Ekspor:
=====
Ekspor = Luas.l1 + Produksi.l1 + Ekspor.l1 + world.Price.l1 + const

      Estimate Std. Error t value Pr(>|t|)
Luas.l1      1.098e-01 1.644e-01  0.668  0.509
Produksi.l1  2.028e-01 1.903e-01  1.065  0.296
Ekspor.l1    2.172e-01 1.803e-01  1.205  0.238
world.Price.l1 2.104e+04 1.876e+04  1.122  0.271
const        4.725e+03 1.453e+05  0.033  0.974

Estimation results for equation world.Price:
=====
world.Price = Luas.l1 + Produksi.l1 + Ekspor.l1 + world.Price.l1 + const

      Estimate Std. Error t value Pr(>|t|)
Luas.l1      -8.190e-07 1.381e-06 -0.593 0.557692
Produksi.l1  5.376e-07 1.599e-06  0.336 0.739128
Ekspor.l1    2.955e-07 1.515e-06  0.195 0.846646
world.Price.l1 6.619e-01 1.576e-01  4.200 0.000232 ***
const        1.144e+00 1.221e+00  0.937 0.356324

---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 39250 on 29 degrees of freedom
Multiple R-Squared: 0.9469, Adjusted R-squared: 0.9396
F-statistic: 129.4 on 4 and 29 DF, p-value: < 2.2e-16

Covariance matrix of residuals:
      Luas  Produksi  Ekspor world.Price
Luas    1540596889 384196575 -173835923 -5525.0528
Produksi 384196575 914413578 407526034 -622.0452
Ekspor -173835923 407526034 3395050383 -2750.9331
world.Price -5525 -622 -2751 0.2396

```

---

---

Correlation matrix of residuals:				
	Luas	Produksi	Ekspor	world.Price
Luas	1.00000	0.32370	-0.07601	-0.28756
Produksi	0.32370	1.00000	0.23129	-0.04202
Ekspor	-0.07601	0.23129	1.00000	-0.09645
world.Price	-0.28756	-0.04202	-0.09645	1.00000

---

Setelah terpilih *lag* model VAR yang optimal, maka perlu dilakukan pengujian asumsi non-autokorelasi, normalitas, dan homoskedas pada sisaan model VAR (1). Dari hasil pengujian non autokorelasi, diperoleh informasi bahwa model tersebut terbebas dari autokorelasi dengan tingkat kesalahan 5% (Tabel 19). Residual memenuhi asumsi kenormalan dengan tingkat kesalahan 5% melalui *normality test* (*Jarque Bera Test*, pengujian skewness, dan kurtosis) sebagaimana pada Tabel 20, sedangkan asumsi homoskedastisitas atau kehomogenan ragam signifikan terpenuhi dengan *p-value* =1 (Tabel 21).

**Tabel 19.** Hasil Uji Asumsi Non-Autokorelasi model VAR (1) menggunakan *Serial Test* Portmanteau Test (asymptotic)

---

data: Residuals of VAR object varkopi.c1  
Chi-squared = 197.17, df = 240, **p-value = 0.98**

---

**Tabel 20.** Hasil Uji Asumsi Normalitas Regresi Linier Produksi Kopi menggunakan *Normality Test*

---

\$JB  
JB-Test (multivariate)  
data: Residuals of VAR object varkopi.c1  
Chi-squared = 8.6567, df = 8, **p-value = 0.3721**

\$Skewness  
Skewness only (multivariate)  
data: Residuals of VAR object varkopi.c1  
Chi-squared = 6.2226, df = 4, **p-value = 0.1831**

\$Kurtosis  
Kurtosis only (multivariate)  
data: Residuals of VAR object varkopi.c1  
Chi-squared = 2.4341, df = 4, **p-value = 0.6565**

---

**Tabel 21.** Hasil Uji Asumsi Homoskedas Regresi Linier Produksi Kopi menggunakan *ARCH Test*

---

ARCH (multivariate)  
data: Residuals of VAR object varkopi.c1  
Chi-squared = 290, df = 500, **p-value = 1**

---

Sebelum melakukan peramalan, perlu dilakukan evaluasi nilai kesalahan model VAR yang terbentuk dengan menghitung nilai MAPE atau nilai rata-rata persentase kesalahan peramalan. Nilai MAPE VAR (1) untuk dataset *training* dan *testing* masing-masing sebesar 2,39% dan 1,18% seperti tersaji pada Tabel 22.

**Tabel 22.** Nilai MAPE Model VAR (1) Luas Areal Kopi Indonesia

---

	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
MAPE Data <i>Training</i>	0.05193	0.70364	1.71595	<b>2.39440</b>	3.41742	8.93998
MAPE Data <i>Testing</i>	0.2053	0.9334	1.2847	<b>1.1805</b>	1.5362	1.8720

---

Setelah diperoleh model VAR yang optimal yakni VAR (1), maka model tersebut digunakan untuk melakukan peramalan luas areal kopi 5 tahun ke depan (2021-2025). Hasil peramalan luas areal kopi Indonesia menggunakan metode VAR dengan peubah produksi kopi, volume ekspor kopi, dan harga kopi dunia, ditampilkan pada Tabel 23.

**Tabel 23.** Hasil Peramalan Luas Areal Kopi menggunakan VAR (1)

[1] 1256935 1262085 1265886 1268673 1270733

## KESIMPULAN DAN SARAN

### Kesimpulan

- Hasil analisis yang didapat dari tiga model adalah sebagai berikut:
  - Model ARIMA terbaik adalah ARIMA (0,1,5) dengan MAPE data *training* sebesar 2,40% dan MAPE data *testing* sebesar 0,40%
  - Model Fungsi Transfer terbaik dengan peubah input volume ekspor kopi adalah ARIMA (0,0,0)(4,1,5) dengan model input ARIMA (1,1,2). Model Fungsi Transfer tersebut menghasilkan MAPE data *testing* menggunakan nilai aktual sebagai input sebesar 0,39% dan MAPE data *testing* menggunakan data ramalan sebagai input sebesar 0,59%.
  - Model VAR terbaik dengan memasukkan peubah produksi kopi, volume ekspor kopi, dan harga kopi dunia adalah VAR (1) tanpa *trend* (hanya konstanta), menghasilkan MAPE data *training* dan *testing* masing-masing sebesar 2,39% dan 1,18%
- Dari ketiga model yang dicobakan, model terbaik adalah model ARIMA dengan melihat MAPE terkecil yaitu 0,40%
- Hasil ramalan luas areal kopi model ARIMA (0,1,5) untuk Tahun 2021 sampai 2025 adalah 1.250.203 Ha, 1.252.694 Ha, 1.251.091 Ha, 1.251.689 Ha, dan 1.251.850 Ha.

### Saran

- Perlu dilakukan kajian mendalam dengan metode peramalan lainnya.
- Untuk model fungsi transfer dan VAR perlu diujicobakan dengan menggunakan variabel lain yang sekiranya lebih berpengaruh, baik secara teoritis maupun praktis.
- Perlu dikaji metode peramalan untuk data non parametrik, sehingga tidak terhalang oleh tidak terpenuhinya berbagai asumsi.

## DAFTAR PUSTAKA

- Draper, N. R, dan Smith, H. 1992. Analisis Regresi Terapan, Jakarta : PT Gramedia Pustaka Utama
- Direktorat Jenderal Perkebunan. 2019. Statistik Perkebunan Indonesia 2018-2020 Jakarta: Direktorat Jenderal Perkebunan, Kementrian Pertanian.Statistik Perkebunan Indonesia 2018-2020.
- Gujarati, Damodar. N dan Porter, Dawn. C. 2009. Basic Econometrics. Boston: Douglas Reiner.
- Malian, Husni & Sudi Mardianto dan Mewa Ariani. 2004. Faktor-faktor yang Mempengaruhi Produksi, Konsumsi dan Harga Beras serta Inflasi Bahan Makanan. Jurnal Agro Ekonomi, Volume 22 No. 2: 119-146

- Siagian, Viktor dan Muchamad Yusron. 2015. Analisis Faktor-Faktor yang Mempengaruhi Luas Lahan Garapan Usahatani Padi Sawah di Provinsi Banten. Lampung: Politeknik Negeri Lampung
- Thamrin, Syahrini. 2014. Faktor-faktor yang Mempengaruhi Produksi Usahatani Kopi Arabika di Kabupaten Enrekang Sulawesi Selatan. Pangkajene Kepulauan: Politeknik Pertanian Negeri Pangkajene Kepulauan

## SCRIPT PROGRAM RSTUDIO

### MODEL ARIMA

```
#Persiapan Data#
library(readxl)
serieskopi <- read_excel("D:/BPS PERKEBUNAN/ESTIMASI ARIMA/Serie_Kopi.xlsx")
view(serieskopi)

#Plot Deret waktu#
serieskopi <- ts(serieskopi, start=c(1980),end=c(2020),frequency=1)
ts.plot(serieskopi[, "Luas"], type="l", ylab="Luas Kopi")

#Pembagian data training testing#
train=serieskopi[1:35,]
test=serieskopi[36:41,]
total=serieskopi[1:41,]

#Pemeriksaan Kestasioneran#
install.packages("urca")
library(urca)
stasioner <- ur.df(serieskopi[, "Luas"], type="trend")
summary(stasioner)

#Proses Differencing 1#
ts.plot(diff(serieskopi[, "Luas"]), type="l", ylab="Diff. Luas Kopi")
stasioner1 <- ur.df(diff(serieskopi[, "Luas"]), type="none")
summary(stasioner1)

#Pemeriksaan Perilaku Autokorelasi#
acf(train[, "Luas"])
pacf(train[, "Luas"])

acf(diff(train[, "Luas"]))
pacf(diff(train[, "Luas"]))

#Pendugaan Model ARIMA#
install.packages("forecast")
library(forecast)
tmodel <- auto.arima(train[, "Luas"], seasonal = FALSE)
summary(tmodel)

install.packages("caschrono")
library(caschrono)
minic1 <- armaselect(diff(train[, "Luas"]), max.p = 5, max.q = 5, nbmod = 10)
minic1

model2 <- arima(train[, "Luas"], order=c(0,1,5))
summary(model2)

library(lmtest)
coeftest(model2)

install.packages("remotes")
library(remotes)
install_github('cran/portes', upgrade = 'always')
library(portes)
ljbtest <-
LjungBox(residuals(model2), lags=seq(5, 30, 5), order=0, season=1, squared.residuals=FALSE)
ljbtest

#Pemeriksaan Sisaan#
tsdisplay(residuals(model2), lag.max=32, main='Model ARIMA Luas Kopi (0,2,1) Model Residuals')

#Peramalan Data Testing#
ramalan_arima = forecast(model2, 6)
forecast_arima <- ramalan_arima$mean
forecast_arima

accuracy(ramalan_arima, test[, "Luas"])
```



```

#Plot Ramalan#
plot(ramalan_arima)

#Pengepasan Model Untuk Seluruh Data#
model.arima <- Arima(serieskopi[, "Luas"], order=c(0,1,5))
summary(model.arima)

ramalan_arima2 = forecast(model.arima, 5)
ramalan_arima2
plot(ramalan_arima2)

```

## MODEL FUNGSI TRANSFER

```

##### Penyiapan data #####
library(readxl)
serieskopi <- read_excel("D:/BPS PERKEBUNAN/ESTIMASI ARIMA/Serie_Kopi.xlsx")
View(serieskopi)
serieskopi <- ts(serieskopi, start=c(1980), end=c(2020), frequency=1)
ts.plot(serieskopi[, "Luas"], type="l", ylab="Luas Kopi")
ts.plot(serieskopi[, "Ekspor"], type="l", ylab="Ekspor kopi dunia")

##### Pembagian data: training-testing #####
train.h=serieskopi[1:35,]
test.h=serieskopi[36:41,]

##### Pemeriksaan kestasioneran data input #####
library(urca)
stasioner <- ur.df(train.h[, "Ekspor"], type="trend")
summary(stasioner)
stasioner <- ur.df(diff(train.h[, "Ekspor"]), type="none")
summary(stasioner)
acf(train.h[, "Ekspor"])
pacf(train.h[, "Ekspor"])
acf(diff(train.h[, "Ekspor"]))
pacf(diff(train.h[, "Ekspor"]))

##### Pencarian model tentatif model input #####
library(forecast)
tmodel <- auto.arima(train.h[, "Ekspor"], seasonal = FALSE)
summary(tmodel)
library(caschnono)
minic1 <- armaselect(diff(train.h[, "Ekspor"]), max.p = 5, max.q = 5, nbmod = 10)
minic1

model1 <- arima(train.h[, "Ekspor"], order=c(1,1,2))
model1
summary(model1)
library(lmtest)
coefstest(model1)

#Peramalan Data Testing#
ramalan_arima = forecast(model1, 6)
forecast_arima <- ramalan_arima$mean
forecast_arima

accuracy(ramalan_arima, test.h[, "Ekspor"])

#Pengepasan Model Untuk Seluruh Data#
model.arima <- Arima(serieskopi[, "Ekspor"], order=c(1,1,2))
summary(model.arima)

ramalan_arima2 = forecast(model.arima, 5)
ramalan_arima2
plot(ramalan_arima2)

##### Prewhitening dan korelasi silang #####
alfa <- residuals(Arima(train.h[, "Ekspor"], model=model1))
beta <- residuals(Arima(train.h[, "Luas"], model=model1))
ccf(beta, alfa)
ccf(train.h[, "Luas"], train.h[, "Ekspor"])

##### Pengepasan model (r,s,b) = (0, 0, 0) #####
tf.arima0 <- Arima(train.h[, "Luas"], order=c(0,0,0), xreg=train.h[, "Ekspor"])

```

```

#summary(tf.arima0)

##### Identifikasi model noise #####
res <- residuals(tf.arima0)
ts.plot(res, type="l", ylab="Residual")
acf(res)
pacf(res)
stasioner <- ur.df(res,type="drift")
summary(stasioner)
stasioner <- ur.df(diff(res),type="none")
summary(stasioner)
tmodel <- auto.arima(res,seasonal = FALSE)
summary(tmodel)
minic1 <- armaselect(diff(res), max.p = 5, max.q = 5, nbmod = 10)
minic1

modelres <- Arima(res, order=c(4,1,5))
modelres
summary(modelres)
library(lmtest)
coeftest(modelres)

##### Pengepasan model (r,s, b) = (0, 0, 0), Noise(x, y, z) #####
tf.arima1 <- Arima(train.h[,"Luas"], order=c(4,1,5), xreg=train.h[,"Ekspor"])
summary(tf.arima1)
coeftest(tf.arima1)

##### Peramalan berbasis fungsi transfer, data input nilai AKTUAL #####
ramalan.arima.tf1 <- Arima(test.h[,"Luas"], xreg=test.h[,"Ekspor"],
model=tf.arima1)
summary(ramalan.arima.tf1)
ramalan.arima.tf1x <- forecast(tf.arima1,h=6, xreg=test.h[,"Ekspor"])
ramalan.arima.tf1x$mean
plot(ramalan.arima.tf1x)

##### Peramalan berbasis fungsi transfer, data input nilai RAMALAN #####
model1 <- Arima(train.h[,"Ekspor"], order=c(1,1,2))
ramalan.input <- forecast(model1,6)
ramalan.arima.tf2 <- Arima(test.h[,"Luas"], xreg=ramalan.input$mean,
model=tf.arima1)
summary(ramalan.arima.tf2)
ramalan.arima.tf2x <- forecast(tf.arima1,h=6, xreg=ramalan.input$mean)
ramalan.arima.tf2x$mean
plot(ramalan.arima.tf2x)

##### Grafik Data Asli dan Ramalan pada Data testing #####
tmp1 <- as.matrix(ramalan.arima.tf1x$mean)
tmp2 <- as.matrix(ramalan.arima.tf2x$mean)
gabung0 <- as.data.frame(cbind(serieskopi[36:41,1],serieskopi[36:41,2],tmp1,tmp2))
colnames(gabung0) <- c("Tahun", "Luas", "FT1","FT2")
str(gabung0)
plot(gabung0$Tahun,gabung0$Luas, xlab='indeks Tahun', ylab='Luas Kopi', type='l',
col='black', lwd=1, ylim=c(120000,1260000))
lines(gabung0$Tahun,gabung0$FT1, col='red', lwd=1)
lines(gabung0$Tahun,gabung0$FT2, col='blue', lwd=1)
legend("topleft", c("Luas", "FT1","FT2"), col=c("black","red","blue"),
lty=1,lwd=1,cex=0.6, inset=0.02, box.lty=0)

##### Forecast di periode 2021 - 2025 #####
### Menduga ulang model input ###
model.input <- Arima(serieskopi[,"Ekspor"], order=c(1,1,2))
ramalan.input2 <- forecast(model.input,5)

### Menduga ulang fungsi transfer###
tf.arima.final <- Arima(serieskopi[,"Luas"], order=c(4,1,5),
xreg=serieskopi[,"Ekspor"])
summary(tf.arima.final)
coeftest(tf.arima.final)
boxtest<-Box.test(residuals(tf.arima.final),lag=5,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tf.arima.final),lag=10,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tf.arima.final),lag=15,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tf.arima.final),lag=20,type=c("Ljung-Box"),fitdf=0)
boxtest

```

```

boxtest<-Box.test(residuals(tf.arima.final),lag=25,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tf.arima.final),lag=30,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tf.arima.final),lag=35,type=c("Ljung-Box"),fitdf=0)
boxtest
ramalan.tf.final <- forecast(tf.arima.final,h=5, xreg=ramalan.input2$mean)
ramalan.tf.final$mean
plot(ramalan.tf.final)

```

## MODEL VAR

```

###PENYIAPAN DATA##

library(readxl)
serieskopi <- read_excel("D:/BPS PERKEBUNAN/ESTIMASI ARIMA/Serie_Kopi.xlsx")
view(serieskopi)
str(serieskopi)

## PENENTUAN PANJANG LAG OPTIMAL##

library(vars)

##p=1 type=both##
varkopi.b1 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=1, type="both")
summary(varkopi.b1)

##p=2 type=both##
varkopi.b2 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=2, type="both")
summary(varkopi.b2)

##p=1 type=trend##
varkopi.d1 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=1, type="trend")
summary(varkopi.d1)

##p=2 type=trend##
varkopi.d2 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=2, type="trend")
summary(varkopi.d2)

##p=1 type=constant##
varkopi.c1 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=1, type="const")
summary(varkopi.c1)

##p=2 type=constant##
varkopi.c2 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=2, type="const")
summary(varkopi.c2)

##p=3 type=constant##
varkopi.c3 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=3, type="const")
summary(varkopi.c3)

##p=4 type=constant##
varkopi.c4 <- VAR(serieskopi[1:35,c(2,3,4,5)], p=4, type="const")
summary(varkopi.c4)

####UJI ASUMSI MODEL TERPILIH####

##Uji Sisaan##
ser11 <- serial.test(varkopi.c1, lags.pt = 16, type = "PT.asymptotic")
ser11

## Uji Kenormalan Data##
norm1 <- normality.test(varkopi.c1)
norm1

##Uji homogenitas sisaan##
arch1 <- arch.test(varkopi.c1, lags.multi = 5)
arch1

####PLOT DATA AKTUAL VS HASIL MODEL VAR####

##MAPE data training##
fit <- varkopi.c1[["varresult"]][["Luas"]][["fitted.values"]]

```

```

mapet <- abs(serieskopi[2:35,2] - fit)/serieskopi [2:35,2]*100
summary(mapet)

##Plot data training vs hasil model VAR##
gabung2 <- cbind(serieskopi[2:35,1:2],fit)
plot(gabung2$Tahun,gabung2$Luas, xlab='indeks t', ylab='Luas Kopi', type='l',
col='black', lwd=1)
lines(gabung2$Tahun,gabung2$fit, col='red', lwd=1)
legend("topleft", c("Actual","VAR"), col=c("black","red"), lty=10,lwd=10,cex=1,
inset=0.02, box.lty=0)

##MAPE data testing##
forecast <- predict(varkopi.c1, n.ahead = 6, ci = 0.95)
#plot(forecast)

ramalanvar <- forecast[["fcst"]][["Luas"]]
ramalanvar <- ramalanvar[,1]
mape <- abs(serieskopi[36:41,2] - ramalanvar)/serieskopi[36:41,2]*100
summary(mape)

##Plot data testing vs hasil model VAR##
gabung <- cbind(serieskopi[36:41,1:2],ramalanvar)
plot(gabung$Tahun,gabung$Luas, xlab='indeks t', ylab='Luas Kopi', type='l',
col='black', lwd=1, ylim=c(1130000,1350000))
lines(gabung$Tahun,gabung$ramalanvar, col='red', lwd=1)
legend("topleft", c("Actual","VAR"), col=c("black","red"), lty=1,lwd=1,cex=0.6,
inset=0.02, box.lty=0)

##Plot keseluruhan data vs hasil model VAR##
colnames(gabung) <- c("Tahun","Luas","VAR")
colnames(gabung2) <- c("Tahun","Luas","VAR")
gabung3 <- rbind(gabung2,gabung)
plot(gabung3$Tahun,gabung3$Luas, xlab='indeks t', ylab='Luas Kopi', type='l',
col='black', lwd=1, ylim=c(770000,1400000))
lines(gabung3$Tahun,gabung3$VAR, col='red', lwd=1)
legend("topleft", c("Actual","VAR"), col=c("black","red"), lty=1,lwd=1,cex=0.6,
inset=0.02, box.lty=0)

#####RAMALAN PRODUKSI 5 TAHUN KEDEPAN#####
varkopi.c1.all <- VAR(serieskopi[1:41,c(2,3,4,5)], p=1, type="const")
fit1 <- varkopi.c1.all[["varresult"]][["Luas"]][["fitted.values"]]
forecast <- predict(varkopi.c1.all, n.ahead = 5, ci = 0.95)
#plot(forecast)
ramalanvar <- forecast[["fcst"]][["Luas"]]
ramalanvar <- ramalanvar[,1]
ramalanvar

##### PLOT IMPULS RESPONSE FUNCTION DAN FEVD#####

impresp <- irf(varkopi.c1)
plot(impresp)
par(mar=c(0.5,0.5,0.5,0.5))
plot(fevd(varkopi.c1))

```

# ESTIMASI LUAS AREAL TEBU DI INDONESIA

## *Estimation of Sugar Cane Areal in Indonesia*

Afifah Siti Muslikhah<sup>1\*</sup>

<sup>1</sup>Direktorat Statistik Tanaman Pangan, Hortikultura dan Perkebunan-Badan Pusat Statistik  
Jalan Dr. Sutomo no. 6-8 Gd 4 Lt 4, Pasar Baru-Jakarta Pusat, Indonesia  
E-mail: afifah.siti@bps.go.id

### ABSTRACT

Sugar cane is one of the leading commodities in the plantation sub-sector with high economic value. The area of sugar cane in Indonesia has tended to decline over the last five years. The estimation of the area of sugar cane in the next few years is very much needed as a material for determining policies and the Early Warning System (EWS) considering that the Plantation Fixed Rate (ATAP) data has a considerable lag compared to the current year's data. To estimate the area of sugar cane in the next few years, this study uses the ARIMA model, transfer function, and VAR. The estimation results of the three models were compared and the transfer function was selected as the best estimation model. The selected transfer function model is the ARIMA transfer function (2,1,5) with the input variable being the price of sugar (US). The estimation results with the transfer function produce MAPE training data of 4.15% and MAPE data testing of 1.47%. The area of sugar cane in the next year (2021-2025) is predicted to have a positive trend. In 2021, the sugar cane area is forecasted to be 419,852 Ha, then it will increase in 2025 to 430,430 Ha.

**Keywords:** estimation, ARIMA, transfer function, VAR

### ABSTRAK

Tebu merupakan salah satu komoditas unggulan sub sektor perkebunan dengan nilai ekonomis tinggi. Luas areal tebu di Indonesia cenderung menurun selama lima tahun terakhir. Estimasi luas areal tebu beberapa tahun ke depan sangat diperlukan sebagai bahan penentu kebijakan maupun *Early Warning System (EWS)* mengingat data Angka Tetap (ATAP) Perkebunan memiliki lag yang cukup jauh dibandingkan data tahun berjalan. Untuk mengestimasi luas areal tebu beberapa tahun ke depan, penelitian ini menggunakan model ARIMA, fungsi transfer, dan VAR. Hasil estimasi dari ketiga model tersebut dibandingkan dan terpilih fungsi transfer sebagai model estimasi terbaik. Model fungsi transfer yang terpilih adalah fungsi transfer ARIMA (2,1,5) dengan variabel input harga gula (US). Hasil estimasi dengan fungsi transfer menghasilkan MAPE data training sebesar 4,15% dan MAPE data testing 1,47%. Luas areal tebu tahun ke depan (2021-2025) diramalkan memiliki *trend* yang positif. Tahun 2021 luas areal tebu diramalkan sebesar 419.852 Ha kemudian meningkat di tahun 2025 menjadi 430.430 Ha.

**Kata kunci:** estimasi, ARIMA, fungsi transfer, VAR

### PENDAHULUAN

Tanaman tebu merupakan salah satu subsektor perkebunan yang tersebar di berbagai propinsi di Indonesia dan dikelola oleh pemerintah, swasta maupun rakyat (perkebunan rakyat). Tebu merupakan tanaman yang tumbuh dengan baik di Indonesia. Indonesia merupakan daerah yang cocok untuk tanaman tebu, karena iklim yang hadir di Indonesia sangat cocok untuk kebutuhan pertumbuhan tebu, karena tebu membutuhkan musim hujan pada saat penanaman dan sedikit hujan saat proses pemanenan. Luasnya areal pertanian komoditas tebu memberikan peluang tersedianya bahan baku pengelolaan tebu menjadi produk yang memiliki nilai tambah sehingga kesempatan

untuk meningkatkan pendapatan semakin besar (Sholikhah, 2014).

Luas Areal Tebu di Indonesia selama lima tahun terakhir menunjukkan adanya trend penurunan. Tahun 2016 luas areal tebu sebesar 44.5075 Ha dan menurun menjadi 42.0505 Ha pada tahun 2020 atau rata-rata terjadi penurunan 1,39% per tahun (Direktorat Jendral Perkebunan, 2019). Estimasi luas areal tebu beberapa tahun ke depan sangat diperlukan sebagai bahan penentu kebijakan maupun *Early Warning System (EWS)* mengingat data Angka Tetap (ATAP) Perkebunan memiliki lag yang cukup jauh dibandingkan data tahun berjalan.

Estimasi areal pada tahun berjalan maupun beberapa periode ke depan sangat *urgent* untuk

dilakukan. Hal ini karena informasi tersebut menjadi bahan untuk penentuan kebijakan di subsektor perkebunan. Estimasi luas areal komoditas perkebunan untuk lima tahun ke depan masih belum tersedia. Estimasi Ditjen Perkebunan hanya dilakukan untuk satu tahun ke depan menggunakan model *univariate* seperti *Double Exponential Smoothing (DES)*. Salah satu kelemahan dari model *univariate* yaitu variabel yang digunakan hanya satu misalnya produksi. Akibatnya hasil analisis hanya mampu memberikan gambaran terhadap satu variabel saja tanpa adanya intervensi dari variabel lain. Padahal produksi komoditas perkebunan tidak terlepas dari pengaruh variabel-variabel lain seperti harga, luas areal, ekspor-impor serta variabel lainnya. Selain itu, untuk menentukan kebijakan subsektor perkebunan, diperlukan informasi variabel input lain yang diduga turut berpengaruh terhadap produksi komoditas perkebunan sehingga intervensi kebijakan dapat dilakukan pada variabel input tersebut. Oleh karena itu diperlukan model yang mampu menyajikan analisis mendalam dalam mengestimasi produksi dengan melibatkan variabel input lain, misalnya model *multivariate*.

Penelitian ini akan menganalisis hasil estimasi luas areal tebu di Indonesia dengan model *univariate* maupun *multivariate*. Terdapat tiga model yang digunakan dalam mengestimasi luas areal tebu antara lain *Autoregressive Integrated Moving Average (ARIMA)*, fungsi transfer dan *Vector Autoregression (VAR)*. Model ARIMA menghasilkan estimasi luas areal tebu tanpa ada pengaruh dari variabel lain. Model fungsi transfer menghasilkan angka estimasi produksi dengan memasukkan intervensi dari satu variabel pendukung yang dianggap paling berpengaruh terhadap produksi. Model VAR mengestimasi produksi dengan dengan mempertimbangkan pengaruh dari beberapa variabel lain atau terdapat lebih dari satu variabel pendukung yang diduga berpengaruh terhadap produksinya. Hasil estimasi dari keempat model tersebut akan dibandingkan untuk selanjutnya ditentukan model terbaik untuk meramalkan luas areal tebu di Indonesia beberapa tahun ke depan.

## METODE PENELITIAN

### Kerangka Pemikiran

Kerangka pemikiran dalam penelitian menggambarkan hasil estimasi luas areal tebu dengan tiga model yaitu ARIMA, fungsi transfer dan VAR. Model ARIMA umumnya digunakan untuk peramalan jangka pendek, sedangkan

untuk peramalan jangka panjang kurang baik ketepatannya. Estimasi dengan model ARIMA hanya menggunakan variabel dependen dan mengabaikan variabel independennya. Dengan kata lain, untuk mengestimasi luas areal tebu beberapa tahun ke depan maka variabel yang digunakan hanya areal itu sendiri.

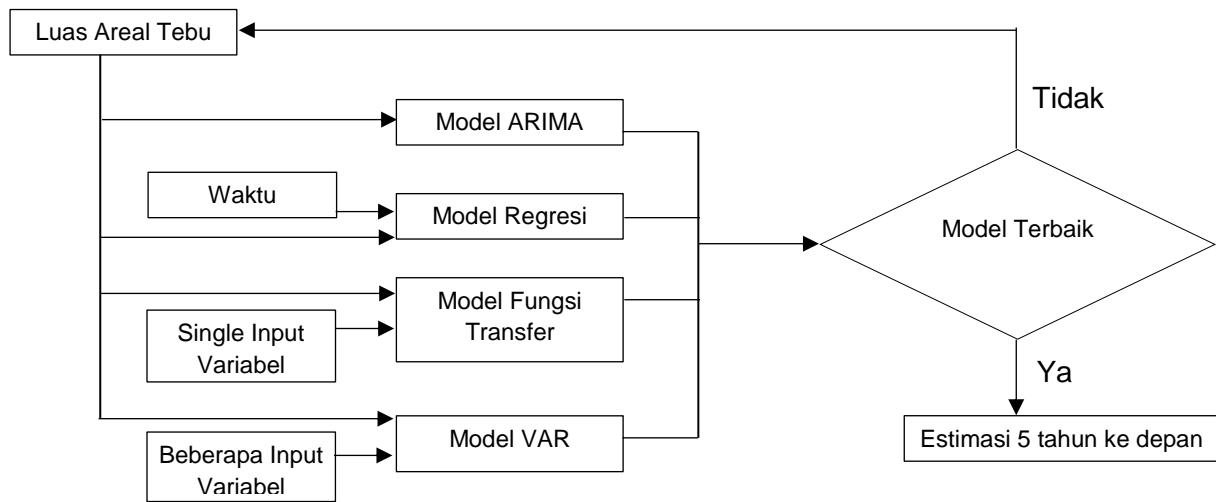
Model fungsi transfer menggambarkan nilai ramalan masa depan dari suatu deret berkala (deret output) yang didasarkan pada nilai-nilai masa lalu dari deret itu sendiri serta didasarkan pula pada suatu deret berkala yang berhubungan (deret input). Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linear antara waktu ke- $t$  dengan deret/variabel input, tetapi juga terdapat hubungan antara variabel input dengan variabel output pada waktu ke- $t$ ,  $t+1$ , ...,  $t+k$ . Pada fungsi transfer terdapat rangkaian output yang mungkin dipengaruhi oleh rangkaian *multiple* input. Untuk kasus *single input* variabel pada fungsi transfer, dapat menggunakan metode korelasi silang. Penelitian ini menggunakan *single input* variabel yaitu volume ekspor untuk meramalkan luas areal tebu sebagai variabel outputnya.

Model VAR menggunakan pendekatan non-struktural atau tidak mendasarkan pada teori ekonomi tertentu dalam melakukan peramalan. Model ini memperlakukan seluruh peubah secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model VAR memperlakukan seluruh variabel sebagai variabel dependen/endogen, karena pada kenyataannya suatu variabel dapat bertindak sebagai endogen atau dependen dari suatu peubah lain (Gujarati & Porter, 2010). Untuk meramalkan luas areal tebu beberapa tahun ke depan, penelitian ini menggunakan beberapa variabel antara lain Luas areal tebu, Produksi gula, Harga Gula EU, Harga Gula US, Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung.

Pembentukan model estimasi luas areal tebu dilakukan dengan membagi series data aktual menjadi data training dan data testing. Data training digunakan untuk menentukan model estimasi dan meramalkan data testing yang sebenarnya sudah tersedia data aktualnya. Hasil ramalan data testing tersebut kemudian dibandingkan dengan data aktual untuk dihitung tingkat kesalahan (*error*) hasil ramalan. Model terbaik untuk estimasi adalah model dengan tingkat *error* yang paling kecil, dalam hal ini ditunjukkan oleh nilai *Mean Absolute Percentage Error (MAPE)* terkecil. Selain MAPE, pemilihan model terbaik juga mempertimbangkan kelogisan hasil ramalan dengan historis data sebelumnya. Berdasarkan hasil identifikasi model ARIMA, regresi, fungsi transfer dan VAR, dipilih model

terbaik untuk meramalkan luas areal tebu di Indonesia selama lima tahun ke depan. Secara

umum tahapan penelitian ini disajikan melalui kerangka konseptual penelitian pada Gambar 1.



Gambar 1. Kerangka Konseptual Penelitian

### Pengumpulan Data

Data yang digunakan pada penelitian ini bersumber dari Direktorat Jenderal Perkebunan, Kementerian Pertanian. Variabel yang digunakan antara lain Luas areal tebu, Produksi gula, Harga Gula EU, Harga Gula US, Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung. Luas areal tebu merupakan total luas panen areal tebu yang diusahakan oleh Perkebunan Rakyat (PR), Perkebunan Besar Negara (PBN) dan Perkebunan Besar Swasta (PBS). Harga Gula EU, Harga Gula US, Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung didapatkan dari web World Bank dan Pusat Informasi Harga Pangan Strategis Nasional (PIHPS). Series data yang digunakan pada penelitian ini yaitu data tahun 1972-2020. Berdasarkan series data tersebut, selanjutnya dilakukan pengelompokan data training untuk periode 1972-2014 dan data testing untuk periode 2015-2020, sehingga diperoleh total observasi sebanyak 49. Berdasarkan model terbaik yang terpilih, estimasi luas areal tebu di Indonesia dilakukan untuk lima tahun ke depan yaitu 2020-2024.

### Analisis Data

Secara empiris, penelitian ini membandingkan hasil estimasi luas areal tebu dengan tiga model yaitu ARIMA, fungsi transfer dan VAR. Tahapan penelitian dimulai dengan mencari model estimasi berdasarkan historis data training untuk

meramalkan data testing. Selanjutnya hasil estimasi data testing dibandingkan dengan nilai aktual produksinya untuk mengetahui tingkat kesalahan berdasarkan nilai MAPE. Berdasarkan nilai MAPE yang dihasilkan oleh ketiga model estimasi tersebut dipilih model ramalan dengan MAPE terkecil. Model dengan MAPE terkecil dipilih sebagai model terbaik untuk mengestimasi luas areal tebu lima tahun ke depan. Model terbaik yang terpilih juga harus memenuhi asumsi statistik yang ditetapkan di masing-masing model. Pengolahan data untuk estimasi luas areal tebu baik dengan model ARIMA, fungsi transfer maupun VAR dilakukan dengan *software* RStudio.

### Estimasi dengan Model ARIMA

Model ARIMA dibagi ke dalam tiga kelompok model yaitu *Autoregressive Model (AR)*, *Moving Average Model (MA)* dan *Autoregressive Integrated Moving Average Model (ARIMA)*. Model AR menjelaskan pergerakan suatu peubah itu sendiri di masa lalu. Model AR ordo ke-*p* untuk mengestimasi luas areal tebu atau dapat ditulis ARIMA (*p*, 0, 0) sebagai berikut:

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t \dots \dots \dots (1)$$

dimana:

- $Y_t$  = luas areal tebu pada waktu ke-*t*
- $Y_{t-p}$  = luas areal tebu pada kurun waktu ke (*t-p*)
- $\mu$  = suatu konstanta
- $\theta_1 \dots \theta_p$  = parameter autoregressive ke-*p*

$\varepsilon_t$  = nilai kesalahan pada waktu ke- $t$

Model MA menjelaskan pergerakan peubahnya melalui sisaannya di masa lalu. Model MA dengan ordo  $q$  untuk mengestimasi luas areal tebu atau ARIMA (0,0, $q$ ) ditulis sebagai berikut:

$$Y_t = \mu - \phi_1\varepsilon_{t-1} - \phi_2\varepsilon_{t-2} - \dots - \phi_q\varepsilon_{t-q} + \varepsilon_t \dots\dots\dots(2)$$

dimana:

$Y_t$  = luas areal tebu pada waktu ke- $t$

$\phi_1 \dots \phi_q$  = parameter-parameter moving average

$\varepsilon_{t-q}$  = nilai kesalahan pada waktu ke ( $t-q$ )

Model ARIMA merupakan model dari fungsi linear nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk model ARIMA ( $p,d,q$ ) untuk mengestimasi luas areal tebu ditulis sebagai berikut:

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1\varepsilon_{t-1} - \phi_2\varepsilon_{t-2} - \dots - \phi_q\varepsilon_{t-q} + \varepsilon_t \dots\dots\dots(3)$$

dimana:

$Y_t$  = luas areal tebu pada waktu ke- $t$

$Y_{t-p}$  = luas areal tebu pada kurun waktu ke ( $t-p$ )

$\mu$  = suatu konstanta

$\theta_1 \theta_2 \phi_1 \phi_n$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke- $(t-q)$

Penggunaan model ARIMA mensyaratkan series data yang stasioner. Suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melakukan *differencing*. *Differencing* yaitu menghitung perubahan atau selisih nilai observasi. Data yang telah dilakukan *differencing* perlu dicek kembali apakah telah stasioner atau belum. Pengecekan stasioneritas data dapat dilihat dari beberapa cara antara lain melihat sebaran data, menggunakan uji *Augmented Dickey-Fuller Test (ADF) Test Unit Root Test* dan melihat dari perilaku autokorelasi berdasarkan plot *Autocorrelation Function (ACF)* dan *Partial Autocorrelation Function (PACF)*.

Berdasarkan sebaran datanya, data yang telah stasioner menyebar secara acak dan tidak memiliki pola-pola tertentu baik pola musiman maupun *trend*. Pengecekan stasioneritas dengan uji ADF memiliki hipotesis sebagai berikut:

Hipotesis: .....(4)

$H_0$  : Data tidak stasioner

$H_1$  : Data stasioner

Jika nilai *test-statistic* pada uji ADF lebih kecil dari *critical value for test-statistic* baik pada taraf ( $\alpha$ )

5% maka  $H_0$  ditolak yang berarti data telah stasioner. Pengecekan stasioneritas dari perilaku *autokorelasi* dilihat dari plot ACF dan PACF. Jika pada kedua plot tersebut tidak banyak lag yang keluar dari *confidence interval* maka data telah stasioner.

Data yang telah stasioner selanjutnya dilakukan tahapan pendugaan model ARIMA menggunakan fungsi *auto.arima* atau *armaselect* yang tersedia pada software RStudio. Software tersebut akan memberikan rekomendasi model terbaik untuk mengestimasi luas areal tebu. Berdasarkan model terbaik yang terpilih, kemudian dilakukan pemeriksaan sisaan menggunakan pengujian LjungBox. Jika autokorelasi sisaan tidak signifikan yang ditandai dengan nilai p-value yang lebih besar dari 5%, maka model ARIMA tersebut sudah cukup baik untuk mengemas data luas areal tebu.

Model ARIMA yang terpilih digunakan untuk mengestimasi data testing. Hasil ramalan data testing selanjutnya dibandingkan dengan data aktualnya untuk mengecek akurasi hasil ramalan. Akurasi hasil ramalan model ARIMA ditunjukkan oleh MAPE data training dan data testing. Jika model terpilih dirasa telah menghasilkan MAPE yang kecil, maka model tersebut dapat digunakan untuk mengestimasi luas areal tebu untuk beberapa periode ke depan.

**Estimasi dengan Model Fungsi Transfer**

Luas areal tebu di Indonesia tidak terlepas dari pengaruh variabel lain. Menurut Abdul et al. (2017) salah satu variabel yang berpengaruh terhadap luas areal tebu di Indonesia adalah harga gula. Penelitian ini menggunakan harga gula US sebagai variabel input dalam mengestimasi areal (variabel output) menggunakan model fungsi transfer. Model fungsi transfer pada penelitian ini menggambarkan ramalan luas areal tebu yang didasarkan pada nilai-nilai masa lalu dari areal itu sendiri, serta didasarkan pada harga gula US (variabel input) dan gangguan/noise. Model fungsi transfer untuk mengestimasi luas areal tebu dituliskan sebagai berikut:

$$y_t = \nu(B)x_t + N_t \quad \rightarrow \quad y_t = \frac{\omega_s(B)}{\delta_s(B)} x_{t-s} + \frac{\theta_s(B)}{\varphi_s(B)} \varepsilon_t \dots\dots(5)$$

dimana:

$y_t$  = luas areal tebu tahun ke- $t$

$x_t$  = harga gula US tahun ke- $t$

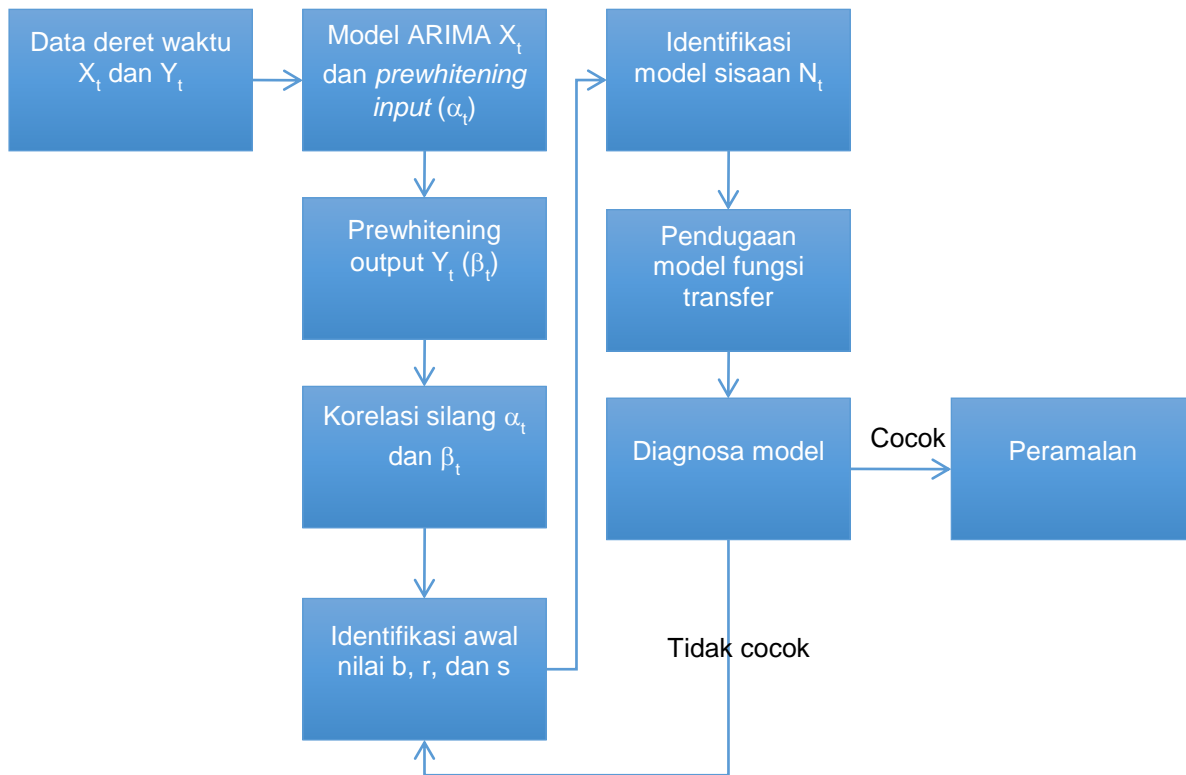
$b$  = panjang jeda pengaruh harga gula US terhadap luas areal tebu

$r$  = panjang lag luas areal tebu periode sebelumnya yang masih



mempengaruhi luas areal tebu tahun- $t$   
 $s$  = panjang jeda harga gula US  
periode sebelumnya yang masih  
mempengaruhi luas areal tebu tahun- $t$   
 $p$  = ordo AR bagi noise  $N_t$   
 $q$  = ordo MA bagi noise  $N_t$

Pemodelan untuk estimasi luas areal tebu dengan fungsi transfer dilakukan melalui beberapa langkah sebagaimana disajikan pada Gambar 2 berikut:



Gambar 2. Langkah-langkah Estimasi Luas areal tebu dengan Fungsi Transfer

### Estimasi dengan Model VAR

Pemodelan dengan *Vector Autoregression (VAR)* memperlakukan seluruh peubah secara simetris tanpa mempermasalahkan apakah variabel tersebut merupakan variabel dependen maupun independen. Penelitian ini menggunakan luas areal tebu sebagai variabel dependen. Luas areal tebu tersebut diduga dipengaruhi oleh beberapa variabel, tidak seperti model fungsi transfer yang hanya menggunakan harga gula US saja sebagai variabel inputnya. Selain harga gula US, variabel input lain yang diduga juga berpengaruh terhadap luas areal tebu di Indonesia antara lain Produksi gula, Harga Gula EU, Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung (Abdul et al., 2017)

Model VAR termasuk kategori model sistem, dimana ketika tidak ada kepastian untuk menentukan bahwa suatu peubah adalah eksogen (independen) maka suatu perluasan analisis fungsi perpindahan alami akan

memperlakukan masing-masing peubah secara simetris (Enders, 2004). Sebagai contoh, pada kasus-kasus peubah yang membiarkan alur waktu atau *time path*  $\{s_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{y_t\}$  dan membiarkan *time path*  $\{y_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{s_t\}$ . Di dalam sistem *bivariate*, hubungan tersebut dapat digambarkan seperti pada persamaan berikut:

$$\begin{aligned} s_t &= b_{10} - b_{12}y_t + \gamma_{11}s_{t-1} + \gamma_{12}y_{t-1} + \varepsilon_{s_t} \\ y_t &= b_{20} - b_{21}s_t + \gamma_{21}s_{t-1} + \gamma_{22}y_{t-1} + \varepsilon_{y_t} \end{aligned} \quad \dots(6)$$

Dengan mengasumsikan bahwa kedua peubah  $s_t$  dan  $y_t$  adalah stasioner:  $\varepsilon_{s_t}$  dan  $\varepsilon_{y_t}$  adalah *disturbances* yang memiliki rata-rata nol dan matriks kovarians terbatas atau bersifat *white noise* dengan standar deviasi yang berurutan  $\sigma_s$

dan  $\sigma_y$ : serta  $\{\varepsilon_{s_t}\}$  dan  $\{\varepsilon_{y_t}\}$  adalah *disturbances* yang independen dengan rata-rata nol dan kovarian terbatas (*uncorrelated white-noise disturbances*). Kedua persamaan di atas merupakan orde pertama VAR, karena panjang lag nya hanya satu. Agar Persamaan (11) lebih mudah dipahami dan digunakan sebagai alat analisis maka ditransformasikan dengan menggunakan matriks aljabar, dan hasilnya dapat dituliskan secara bersama seperti pada persamaan di bawah ini:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

atau dengan bentuk lain:

$$Bx_t = \Gamma_0 + \Gamma_1 x_{t-1} + \varepsilon_t \dots \dots \dots (7)$$

dimana:

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \quad x_t = \begin{bmatrix} s_t \\ y_t \end{bmatrix} \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Dengan melakukan pengalihan antara persamaan (12) dengan  $B^{-1}$  atau invers matriks B, maka akan dapat ditentukan model VAR dalam bentuk standar, seperti dituliskan pada persamaan di bawah ini:

$$x_t = A_0 + A_1 x_{t-1} + \ell_t \dots \dots \dots (8)$$

dimana:

$$A_0 = B^{-1} \Gamma_0$$

$$A_1 = B^{-1} \Gamma_1$$

$$\ell_t = B^{-1} \varepsilon_t$$

Pada penelitian ini  $X_t$  merupakan matriks yang dibentuk dari Luas areal Tebu, Produksi gula, Harga Gula EU, Harga Gula US, Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung. Sebagaimana model regresi, model VAR juga didasarkan pada beberapa asumsi antara lain:

**- Sisaan mengikuti fungsi distribusi normal**  
 Uji normalitas pada model VAR didasarkan pada nilai *Jarque-Bera (JB) test (multivariate)*, *Skewness only (multivariate)*, dan *Kurtosis only (multivariate)*. Hipotesis yang mendasari yaitu:

Hipotesis: .....(9)

$H_0$  : Sisaan terdistribusi normal  
 $H_1$  : *Otherwise*

Jika nilai p-value dari *JB test*, *Skewness only* dan *Kurtosis only* lebih besar dari nilai  $\alpha$  maka  $H_0$  diterima, dengan kata lain asumsi normalitas terpenuhi.

**- Varians sisaan konstan untuk setiap data pengamatan (homoskedastisitas)**

Asumsi homoskedastisitas pada model VAR didasarkan pada nilai *ARCH (multivariate)* dengan hipotesis sebagai berikut:

Hipotesis: .....(10)

$H_0$  : Homoskedastisitas  
 $H_1$  : Heterokedastisitas

Jika nilai p-value pada *ARCH (multivariate)* lebih besar dari nilai  $\alpha$  maka  $H_0$  diterima, dengan kata lain asumsi homoskedastisitas terpenuhi.

**- Tidak terdapat autokorelasi antar sisaan untuk setiap data pengamatan**

*Pormanteau Test (asymptotic)* digunakan untuk menguji ada tidaknya autokorelasi antar sisaan pada data amatan. Hipotesis yang digunakan sebagai berikut:

Hipotesis: .....(11)

$H_0$  : Tidak ada autokorelasi antar sisaan  
 $H_1$  : *Otherwise*

Jika nilai p-value hasil uji *Pormanteau Test* lebih besar dari nilai  $\alpha$  maka  $H_0$  diterima, dengan kata lain asumsi tidak adanya autokorelasi antar sisaan telah terpenuhi.

## HASIL DAN PEMBAHASAN

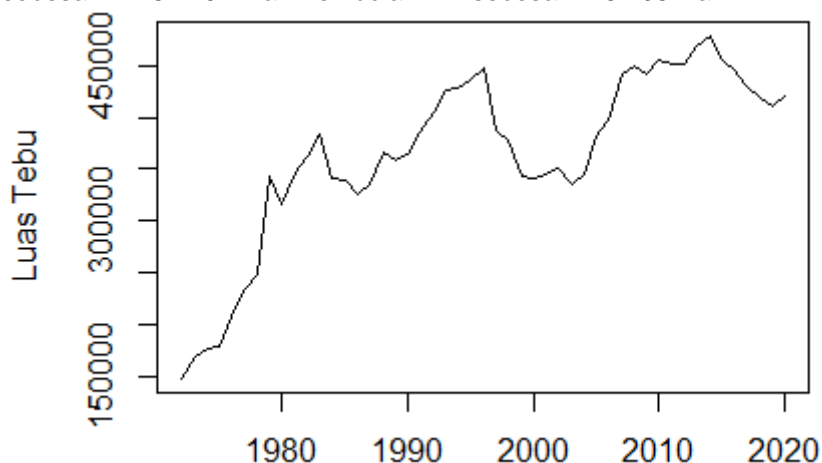
### Keragaman Tebu di Indonesia

Tebu merupakan salah satu komoditas unggulan sub sektor perkebunan dengan nilai ekonomis tinggi. Berdasarkan Statistik Tebu Indonesia 2020 (BPS, 2021) Perkebunan Besar (PB) dan Perkebunan Rakyat (PR) tebu tersebar di sepuluh provinsi di Indonesia, yaitu Sumatera Utara, Sumatera Selatan, Lampung, Jawa Barat, Jawa Tengah, DI Yogyakarta, Jawa Timur, Nusa Tenggara Barat, Sulawesi Selatan, dan Gorontalo. Dilihat dari luas areal, lima provinsi dengan luas areal tebu terluas pada tahun 2020 yaitu Provinsi Jawa Timur, Lampung, Jawa

Tengah, Sumatera Selatan, dan Sulawesi Selatan.

Perkembangan luas areal tebu dari tahun 1972 hingga tahun 2020 jika dilihat pada Gambar 3 cenderung meningkat. Pada tahun 1972 luas areal tebu sebesar 148.710 Ha kemudian

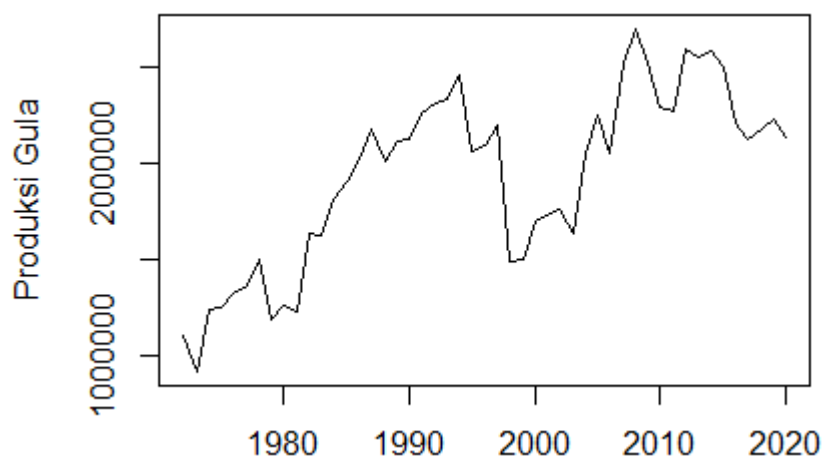
meningkat menjadi 420.505 Ha pada tahun 2020. Namun, jika dilihat sepuluh tahun terakhir (2011-2020) luas areal tebu tercatat mengalami rata-rata penurunan sebesar 0,47% per tahun, dengan areal tertinggi dicapai pada tahun 2014 sebesar 478.108 Ha.



Gambar 3. Perkembangan Luas areal tebu di Indonesia Tahun 1972-2020 (Ha)

Penurunan luas areal tebu dalam sepuluh tahun diikuti dengan produksi gula yang juga memiliki trend negatif di tahun 2011 hingga 2020 (Gambar 4). Pada tahun 2011, produksi gula di Indonesia mencapai 2.267.887 ton, lalu mengalami penurunan menjadi 2.130.719 ton pada 2020. Penurunan hingga mencapai titik terendah dalam sepuluh tahun terakhir terjadi

pada tahun 2017 dimana produksi gula hanya menyentuh angka 2.121.671 ton. Menurut Direktur Jenderal Perkebunan, Kementerian Pertanian, Kasdi Subagyono, adanya penurunan luasan areal tebu, dan daya saing yang juga berkurang karena tebu menggunakan lahan irigasi sehingga berebut dengan komoditas padi dan jagung (Republika, 2020).

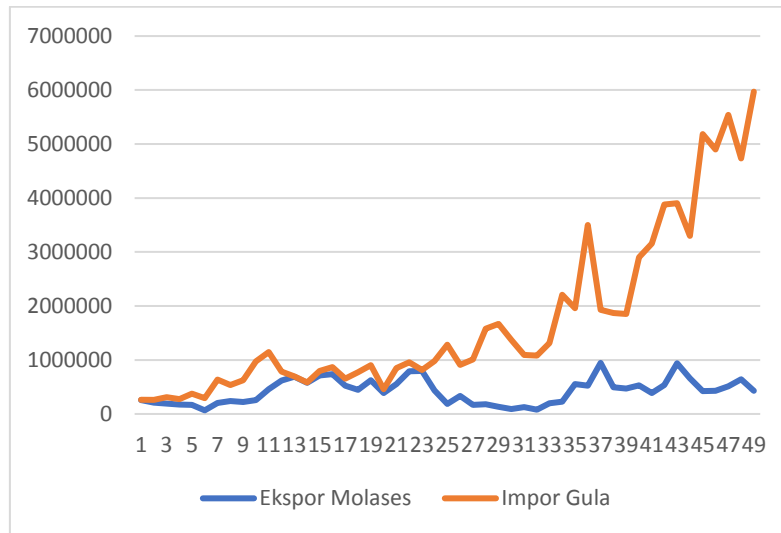


Gambar 4. Perkembangan Produksi Gula di Indonesia Tahun 1976-2020 (Ton)

Sebagai salah satu konsumen gula terbesar di dunia, Indonesia juga merupakan negara yang paling banyak mengimpor gula. Pada tahun 1972 volume ekspor molases sebesar 255.915 ton dengan volume impor gula awalnya hanya sebesar 6,123 ton. Di tahun 2020 jika dilihat pada Gambar 5, volume ekspor molases meningkat menjadi 430.707 ton sedangkan volume impor gula meningkat tajam menjadi sebesar

5.539.679 ton. Dilansir dari worldtoexports, nilai impor gula Indonesia pada tahun 2020 mencapai USD 1,9 miliar atau 7,7% dari total impor gula dunia. Angka tersebut hanya lebih kecil dari Amerika Serikat di peringkat pertama yang mengimpor gula sebesar USD 2,1 miliar. Namun, impor gula Indonesia lebih besar dari China, padahal negeri yang dipimpin Xi Jinping itu

memiliki penduduk 5x lipat lebih dari Indonesia (CNBC Indonesia, 2021).

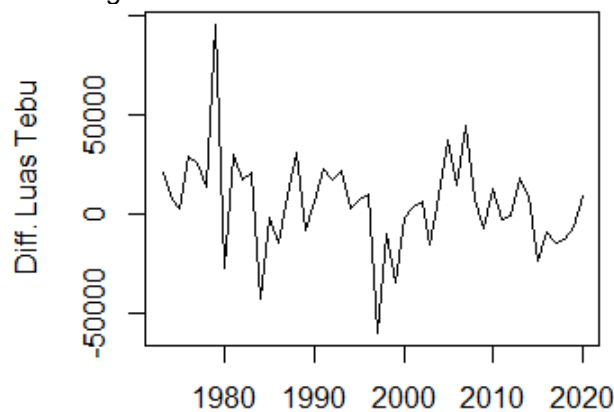


Gambar 5. Perkembangan Volume Ekspor Molases dan Volume Impor Gula di Indonesia Tahun 1972-2020 (Ton)

### Hasil Estimasi dengan Model ARIMA

Pemodelan ARIMA dilakukan untuk data yang telah stasioner. Karena berdasarkan plot data serta hasil uji ADF terindikasi bahwa data luas areal tebu belum stasioner, dilakukan *differencing* satu kali. Setelah *differencing* data tersebut

menjadi stasioner (Gambar 6) yang ditunjukkan dengan nilai *test-statistic* -3,37, dimana nilai tersebut lebih kecil dari *critical value* baik untuk  $\alpha$  5%, sehingga data areal tebu sudah dinyatakan stasioner pada *differencing* pertama.



Gambar 6. Plot Luas Areal Tebu Setelah Differencing

Setelah mendapatkan data luas areal tebu yang stasioner, maka langkah selanjutnya adalah menentukan ordo AR dan MA menggunakan fungsi auto arima yang terdapat pada RStudio. Dari hasil pendugaan auto arima, maka model ARIMA yang terpilih adalah ARIMA (0,1,0) yang berarti model ARIMA yang optimal untuk mengestimasi data yang diolah adalah ARIMA dengan ordo Auto Regressive (AR) = 0, ordo differencing = 1, dan ordo Moving Average (MA) = 0.

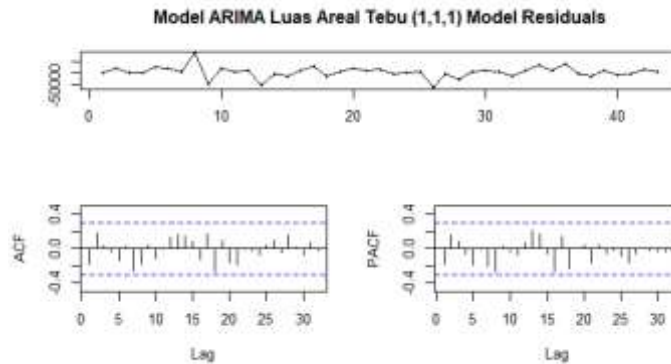
Namun model ARIMA yang dipilih oleh fungsi autoarima dirasa kurang tepat karena pada AR dan MA bernilai nol artinya estimasi akan menghasilkan angka yang konstan, artinya estimasi 5 tahun kedepan akan sama nilainya dari tahun ke tahun. Oleh karena itu, dilakukan overfitting menggunakan fungsi armaselect untuk mencoba ordo lain agar dapat menemukan model ARIMA terbaik yang dapat digunakan untuk melakukan peramalan luas areal tebu Indonesia.

Dari hasil pencarian ordo ARIMA optimum menggunakan fungsi armaselect (diff), maka

diperoleh beberapa kombinasi AR dan MA. Dari beberapa alternatif tersebut diperoleh kandidat model terbaik yaitu ARIMA (1,1,1)

Hasil uji statistik terhadap model ARIMA (1,1,1) menunjukkan bahwa nilai MAPE data training sebesar 5,23% dan untuk data testing sebesar 13,31%. Nilai tersebut dinilai telah cukup

baik untuk melakukan peramalan dan hasil uji Ljung-Box keduanya menunjukkan bahwa tidak terdapat autokorelasi pada residual model dan hasil pemeriksaan residual yang sudah cukup stasioner pada rata-rata sebagaimana terlihat pada Gambar 7.



Gambar 7. Grafik Pemeriksaan Model Residual ARIMA pada Data Training

Setelah diperoleh model ARIMA yang optimal, maka model ARIMA tersebut digunakan untuk melakukan peramalan luas areal tebu 5

tahun ke depan (2021-2025). Hasil peramalan luas areal tebu menggunakan ARIMA (1,1,1) ditampilkan pada Tabel 1.

**Tabel 1.** Hasil Peramalan Luas areal Tebu Indonesia menggunakan ARIMA (1,1,1)

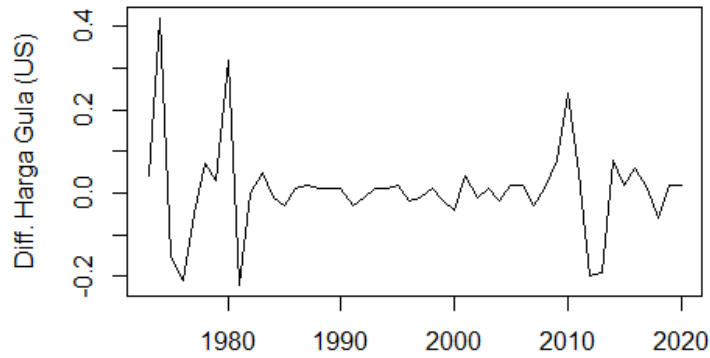
	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	419970	388582.9	451356.5	371967.7	467971.6
2022	419578	371794.3	467361.9	346499.0	492657.2
2023	419292	357342.0	481241.3	324547.8	514035.5
2024	419082	344286.8	493877.4	304692.5	533471.6
2025	418928	332288.3	505569.3	286423.6	551434.0

**Hasil Estimasi dengan Model Fungsi Transfer**

Estimasi dengan model fungsi transfer diawali dengan meramalkan nilai variabel input yaitu Harga Gula (US) menggunakan model ARIMA. Dikarenakan data Harga Gula (US) belum stasioner maka dilakukan *differencing* satu kali sehingga menghasilkan data yang telah stasioner (Gambar 8) yang ditunjukkan dengan nilai *test-statistic* -7,49, dimana nilai tersebut lebih kecil dari *critical value* untuk  $\alpha$  sebesar 5%. Dari hasil fungsi autoarima diperoleh ordo ARIMA terbaik untuk Harga Gula (US) adalah ARIMA (0,1,0). Hasil autoarima dirasa kurang tepat karena akan menghasilkan estimasi yang konstan, sehingga hasil tersebut tidak dipilih untuk forecasting dalam penelitian ini. Untuk memperoleh beberapa alternatif model, dilakukan overfitting

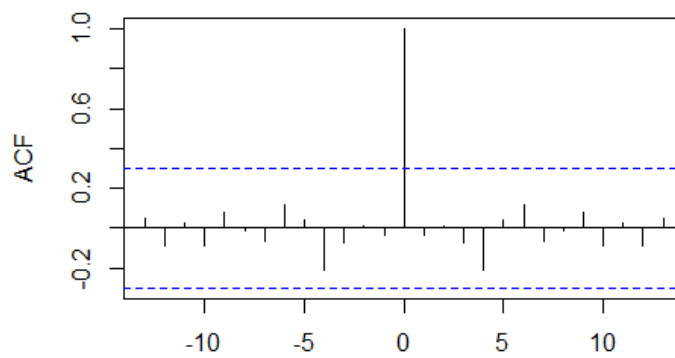
menggunakan fungsi *armaselect* untuk mencoba ordo lain agar dapat menemukan model ARIMA terbaik.

Dari hasil pencarian ordo arima optimum Harga Gula (US) fungsi *armaselect*, diputuskan model arima input Harga Gula (US) adalah ARIMA (2,1,1) dengan nilai MAPE data training sebesar 10,10% dan MAPE data testing sebesar 5,16%. Tahap selanjutnya adalah melakukan *prewhitening* dan analisis korelasi silang antara residual model arima Harga Gula (US) dan luas areal tebu menggunakan ARIMA (1,1,2). Dari hasil plot *ccf* (*cross correlation function*) antara alfa (residual Harga Gula (US)) dan beta (residual luas areal tebu), diperoleh hasil bahwa tidak terdapat lag dari pengaruh yang diberikan oleh Harga Gula (US) terhadap luas areal tebu atau lag = 0. Hal tersebut dapat disimpulkan dari tidak adanya bar atau garis yang melewati ambang batas pada nilai lag positif kecuali pada lag ke-nol (Gambar 9).



Gambar 8. Plot Harga Gula (US) Setelah *Differencing*

**beta & alfa**

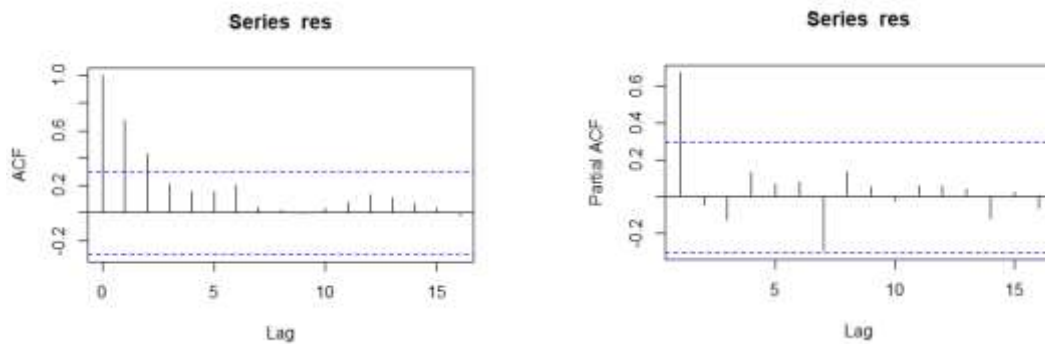


**Lag**

Gambar 9. Plot Harga Gula (US) Setelah *Differencing*

Karena tidak adanya bar atau garis yang melewati ambang batas pada nilai lag positif (kecuali lag ke nol), maka model ARIMA untuk pengepasan model fungsi transfer adalah ARIMA (0,0,0). Kemudian residual model ARIMA (0,0,0) tersebut diteliti stasioneritasnya. Dari plot ACF dan PCF residual, diperoleh hasil bahwa residual model tidak stasioner (Gambar 10), sehingga perlu dilakukan differencing pada. Setelah dilakukan differencing sehingga data residual stasioner, maka dilakukan pencarian model

ARIMA terbaik untuk residual tersebut atau dapat disebut dengan noise. Hasil dari fungsi autoarima menunjukkan bahwa model ARIMA (0,1,0) merupakan model terbaik. Untuk memperoleh referensi ordo ARIMA lain, dilakukan fungsi armaselect. Setelah dilakukan uji signifikansi model ARIMA dan ARIMA-xreg, diperoleh ordo ARIMA noise yang signifikan dan menghasilkan MAPE terkecil adalah ARIMA (2,1,5) yaitu MAPE Training dan Testing masing masing sebesar 4,15% dan 1,47%.



Gambar 10. Plot ACF dan PACF Residual model ARIMA (0,0,0)

Oleh karena telah didapatkan model input serta noise yang terbaik untuk peramalan, maka langkah selanjutnya adalah melakukan peramalan. Peramalan dilakukan menggunakan dua jenis data input, yakni nilai aktual dan data ramalan. Dari hasil peramalan menggunakan nilai aktual sebagai input, diperoleh nilai MAPE sebesar 4,41%, sementara peramalan

menggunakan data ramalan sebagai input memiliki MAPE sebesar 1,46%.

Setelah diperoleh model fungsi transfer yang optimal, maka model tersebut digunakan untuk melakukan peramalan luas areal tebu 5 tahun ke depan (2021-2025). Hasil peramalan luas areal tebu Indonesia menggunakan fungsi transfer dengan harga gula (US) sebagai peubah ditampilkan pada Tabel 2.

**Tabel 2.** Hasil Peramalan Luas Areal Tebu menggunakan Fungsi Transfer

---

Time Series:  
 Start = 2021  
 End = 2025  
 Frequency = 1  
**[1] 419852 424608 429695 431837 430430**

---

**Hasil Estimasi dengan Model VAR**

Pemodelan dengan menggunakan model VAR dilakukan dengan memasukkan peubah luas areal tebu luas areal tebu di Indonesia antara lain Produksi gula, Harga Gula (EU), Harga Gula (US), Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung. Hal ini diasumsikan bahwa apabila terjadi gejala pada variabel-variabel tersebut, akan mempengaruhi gejala luas areal tebu ataupun sebaliknya.

Awalnya perlu menentukan lag (p) yang akan digunakan dalam model VAR. Penentuan

lag (p) dilakukan dengan melakukan pengujian VAR (1) hingga VAR (4) dengan pertimbangan kecukupan series data. Dari hasil running menggunakan p=1 hingga p=5, diperoleh informasi bahwa variabel yang signifikan cukup banyak pada p=2, sehingga model VAR yang digunakan selanjutnya adalah VAR (2). Hal tersebut diperkuat juga dengan nilai Adjusted R-squared: 0,9095 atau dapat dikatakan bahwa 90,95% keragaman luas areal tebu dapat dijelaskan oleh sembilan variabel/peubah penjelas termasuk konstanta. Hasil pengujian model VAR(1) dapat dilihat dalam Tabel 3.

**Tabel 3.** Hasil Estimasi VAR Luas Areal Tebu

---

VAR Estimation Results:  
 Endogenous variables: Luas, Produksi, Harga\_EU, Harga\_US, Ekspor\_Molases, Impor\_Gula, Hrg\_Kons\_Gula, Hprod\_Jagung  
 Deterministic variables: both  
 Sample size: 41  
 Log Likelihood: -2451.845  
 Roots of the characteristic polynomial:  
 0.9844 0.8209 0.8209 0.8131 0.8131 0.804 0.804 0.7667 0.7667 0.6601 0.6601 0.5543 0.5543 0.4119 0.4119 0.2834  
 Call:  
 VAR(y = data\_gula[1:43, c(2, 3, 4, 5, 6, 7, 8, 10)], p = 2, type = "both")

Estimation results for equation Luas:  
 =====  
 Luas = Luas.l1 + Produksi.l1 + Harga\_EU.l1 + Harga\_US.l1 + Ekspor\_Molases.l1 + Impor\_Gula.l1 + Hrg\_Kons\_Gula.l1 + Hprod\_Jagung.l1 + Luas.l2 + Produksi.l2 + Harga\_EU.l2 + Harga\_US.l2 + Ekspor\_Molases.l2 + Impor\_Gula.l2 + Hrg\_Kons\_Gula.l2 + Hprod\_Jagung.l2 + const + trend

	Estimate	Std. Error	t value	Pr(> t )
Luas.l1	4.521e-01	2.104e-01	2.148	0.0424 *
Produksi.l1	3.169e-02	2.253e-02	1.407	0.1728
Harga_EU.l1	2.330e+05	1.279e+05	1.822	0.0816 .
Harga_US.l1	-8.731e+04	6.250e+04	-1.397	0.1757
Ekspor_Molases.l1	5.597e-02	3.603e-02	1.554	0.1339
Impor_Gula.l1	-2.986e-03	1.375e-02	-0.217	0.8300
Hrg_Kons_Gula.l1	2.839e+01	1.255e+01	2.261	0.0335 *
Hprod_Jagung.l1	-2.376e+01	3.424e+01	-0.694	0.4946
Luas.l2	2.709e-01	2.110e-01	1.284	0.2118
Produksi.l2	-4.018e-02	3.043e-02	-1.320	0.1997
Harga_EU.l2	5.386e+04	1.331e+05	0.405	0.6896
Harga_US.l2	-8.298e+03	4.384e+04	-0.189	0.8515

---

Ekspor_Molases.12	-2.707e-02	4.160e-02	-0.651	0.5216
Impor_Gula.12	-2.214e-02	1.527e-02	-1.451	0.1604
Hrg_Kons_Gula.12	1.819e+00	1.307e+01	0.139	0.8906
Hprod_Jagung.12	-2.817e+00	3.978e+01	-0.071	0.9442
const	7.536e+04	3.665e+04	2.056	0.0513 .
trend	-4.594e+03	1.931e+03	-2.380	0.0260 *

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 22490 on 23 degrees of freedom  
Multiple R-Squared: 0.948, Adjusted R-squared: 0.9095  
F-statistic: 24.64 on 17 and 23 DF, p-value: 7.595e-11

Setelah terpilih lag model VAR yang optimal, maka perlu dilakukan pengujian asumsi non-autokorelasi, normalitas, dan homoskedas pada sisaan model VAR (2). Dari hasil pengujian non autokorelasi, diperoleh informasi bahwa model tersebut terbebas dari autokorelasi dengan p-value = 0,7595. Residual memenuhi asumsi kenormalan dengan tingkat kesalahan 5% melalui normality test menggunakan pengujian skewness, dan kurtosis dengan nilai p-value masing masing sebesar 0,08359 dan 0,1041, sedangkan asumsi homoskedastisitas atau kehomogenan ragam signifikan terpenuhi dengan p-value =1.

Sebelum melakukan peramalan, perlu dilakukan evaluasi nilai kesalahan model VAR

yang terbentuk dengan menghitung nilai MAPE atau nilai rata-rata persentase kesalahan peramalan. Nilai MAPE VAR (2) untuk dataset training dan testing masing-masing sebesar 5,82% dan 2,90%.

Setelah diperoleh model VAR yang optimal yakni VAR (2), maka model tersebut digunakan untuk melakukan peramalan luas areal tebu 5 tahun ke depan (2021-2025). Hasil peramalan luas areal tebu Indonesia menggunakan metode VAR dengan peubah Produksi gula, Harga Gula (EU), Harga Gula (US), Ekspor Molases, Impor Gula, Harga Konsumen Gula, Harga produsen Jagung ditampilkan pada Tabel 4.

Tabel 4. Hasil Peramalan Luas Areal Tebu menggunakan VAR (2)

[1]	445075	430363	417576	411435	420505
-----	--------	--------	--------	--------	--------

### Pemilihan Model Estimasi Terbaik

Kriteria pemilihan model estimasi terbaik pada penelitian ini didasarkan pada nilai terkecil dari MAPE data training dan data testing serta kelogisan hasil ramalan jika dibandingkan dengan perkembangan luas areal tebu beberapa tahun terakhir. Dari ketiga model estimasi yang dibandingkan, diperoleh informasi bahwa model fungsi transfer memberikan MAPE data training dan data testing paling kecil dibandingkan model

lain. Dengan luas areal tebu pada tahun 2020 sebesar 420.505 Ha, dan rata-rata pertumbuhan luas areal selama lima tahun terakhir (2016-2020) sebesar -1,39%, maka luas areal tebu pada periode 2021-2025 diramalkan naik 0,47%. Perbandingan hasil estimasi luas areal tebu selama lima tahun ke depan dengan model ARIMA, fungsi transfer dan VAR disajikan pada Tabel 5.



	Pengujian MAPE	Model ARIMA		Fungsi Transfer		Model VAR		
		ARIMA (1,1,1)	(%)	Arima (2,1,5) Xreg=Harga_US	(%)	VAR (2) type=both	(%)	
		MAPE Training	5,028		4,151		5,820	
		MAPE Testing	13,308		1,473		2,901	
ATAP	2016	445.075		445.075		445.075		
	2017	430.363	-3,31	430.363	-3,31	430.363	-3,31	
	2018	417.576	-2,97	417.576	-2,97	417.576	-2,97	
	2019	411.435	-1,47	411.435	-1,47	411.435	-1,47	
	2020	420.505	2,20	420.505	2,20	420.505	2,20	
Angka Estimasi (AESTI)	2021	419.970	-0,13	419.852	-0,16	428.470	1,89	
	2022	419.578	-0,09	424.608	1,13	414.032	-3,37	
	2023	419.292	-0,07	429.695	1,20	419.893	1,42	
	2024	419.082	-0,05	431.837	0,50	419.014	-0,21	
	2025	418.929	-0,04	430.430	-0,33	413.336	-1,36	
Rata-rata Pertumbuhan	ATAP 2016 - 2020		-1,39		-1,39		-1,39	
	AESTI 2021 - 2025		-0,08		0,47		-0,32	

Tabel 5. Perbandingan Hasil Estimasi Luas areal tebu dengan Model ARIMA, Fungsi Transfer dan VAR

## KESIMPULAN DAN SARAN

### Kesimpulan

Model terbaik yang terpilih untuk mengestimasi luas areal tebu adalah fungsi transfer ARIMA (2,1,5) dengan variabel inputnya harga gula (US) dengan nilai MAPE training dan MAPE testing masing masing sebesar 4,15% dan 1,47%. Pemilihan fungsi transfer sebagai model estimasi dengan pertimbangan MAPE data training dan MAPE data testing yang dihasilkan merupakan yang terkecil dibandingkan model lain. Selain itu, hasil estimasi luas areal tebu selama lima tahun ke depan memiliki rata rata kenaikan sebesar 0,47%. Meskipun historis luas areal tebu beberapa tahun ke belakang mengalami trend yang negatif, namun angka estimasi luas areal tebu periode 2021-2025 dianggap masih logis, mengingat dengan beroperasinya beberapa PG baru dan program intensifikasi lahan tebu dari pemerintah akan berdampak pada penambahan luas areal. Hasil estimasi dengan fungsi transfer menunjukkan pada tahun 2021 luas areal tebu di Indonesia sebesar 419.852 Ha kemudian meningkat di tahun 2025 menjadi 430.430 Ha.

### Saran

Penelitian ini membatasi estimasi luas areal tebu menggunakan tiga model yaitu ARIMA, fungsi transfer dan VAR. Meskipun model fungsi transfer dipilih menjadi model terbaik untuk mengestimasi luas areal tebu, akan lebih baik jika pada penelitian berikutnya mencoba model lain untuk mengestimasi luas areal tebu baik itu *univariate* atau *multivariate*. Selain itu, variabel lain selain yang telah digunakan pada penelitian ini dapat juga dicobakan untuk mengestimasi luas areal tebu.

## DAFTAR PUSTAKA

- Abdul, Sayyidah, and Eka. 2017. Factors That Influence Of Supply And Demand Of Crystal Sugar In Indonesia. <https://doi.org/10.18551/rjoas.2017-12.47>
- Badan Pusat Statistik. 2021. Publikasi Statistik Tebu Indonesia 2020.
- CNBC Indonesia. 2021. Alamak! RI Diam-Diam Jadi Negara Importir Terbesar Gula Dunia. Diakses melalui <https://www.cnbcindonesia.com/news/20211029120527-4-287519/alamak-ri-diam-diam-jadi-negara-importir-terbesar-gula-dunia>
- Direktorat Jenderal Perkebunan. 2019. Statistik Perkebunan Indonesia 2018-2020 Jakarta: Direktorat Jenderal Perkebunan, Kementerian Pertanian. Statistik Perkebunan Indonesia 2018-2020.
- Enders, W. 2004. Applied Econometric Time Series. USA: University of Alabama.
- Gujarati, D.N., & Porter, D.C. 2010. Dasar-dasar Ekonometrika. Jakarta: Salemba Empat, Buku 2, Edisi 5.
- Republika. 2020. Kementan: Lahan Tebu Terus Berkurang, Daya Saing Menurun diakses melalui <https://republika.co.id/berita/qka10w370/kementan-lahan-tebu-terus-berkurang-daya-saing-menurun>
- Sholikhah, Umni. 2014. (ibM) Kelompok Usaha Gula Merah di Kecamatan Semboro Kabupaten JEMBER. Inovasi dan Kewirausahaan, Vol. 3, No. 2, Mei 2014 hal 97-104

# **KAJIAN MODEL PERAMALAN LUAS AREAL LADA DI INDONESIA: PENDEKATAN MODEL ARIMA, FUNGSI TRANSFER DAN VAR**

Lasmiyati

<sup>1</sup>*Direktorat Statistik Tanaman Pangan, Hortikultura dan Perkebunan-Badan Pusat Statistik*

*Jalan Dr. Sutomo no. 6-8 Gd 4 Lt 4, Pasar Baru-Jakarta Pusat, Indonesia.*

## **ABSTRAK**

Indonesia dahulu pernah dikenal sebagai negeri rempah-rempah yang menjadi alasan utama banyak bangsa Eropa jauh berlayar meninggalkan negerinya untuk datang ke Indonesia. Setidaknya ada tujuh jenis rempah-rempah yang menjadi kekayaan Indonesia, satu diantaranya adalah lada yang dianggap sebagai salah satu 'emas hitam', yang justru kini malah harus diimpor dari negara lain. Guna merumuskan kebijakan terkait luas areal, sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Penelitian ini akan mengkaji tiga metode yaitu metode *Auto-Regressive Integrated Moving Average* (ARIMA), fungsi transfer dan *Vector Auto Regression* (VAR) dengan menggunakan R Studio guna melakukan pemodelan luas areal lada di Indonesia. Metode yang dipilih guna melakukan peramalan data luas areal Lada adalah metode yang terbaik yang dikaji dari nilai *Mean Absolute Percentage Error* (MAPE) terkecil. Berdasarkan dari hasil nilai MAPE disimpulkan bahwa fungsi transfer adalah metode yang terbaik dengan MAPE 2,02% untuk luas areal lada Indonesia.

## **ABSTRACT**

*Indonesia was once known as the land of spices, which is the main reason many European nations sailed away from their country to come to Indonesia. There are at least seven types of spices that have become Indonesia's wealth, one of which is pepper, which is considered one of 'black gold', which now has to be imported from other countries. In order to formulate policies related to area size, the availability of up-to-date data, and even forecasts for several periods in the future, is highly demanded. This study will examine three methods, namely the Auto-Regressive Integrated Moving Average (ARIMA) method, transfer function and Vector Auto Regression (VAR) using R Studio to model the area of oil palm in Indonesia. The method chosen for forecasting pepper area data is the best method which is assessed from the smallest MAPE (Mean Absolute Percentage Error) value. Based on the results of the MAPE value, it is concluded that the transfer function is the best method with a MAPE of 2,02% for the total area of Indonesian pepper.*

## PENDAHULUAN

Indonesia dahulu pernah dikenal sebagai negeri rempah-rempah yang menjadi alasan utama banyak bangsa Eropa jauh berlayar meninggalkan negerinya untuk datang ke Indonesia. Setidaknya ada tujuh jenis rempah-rempah yang menjadi kekayaan Indonesia, satu diantaranya adalah lada yang dianggap sebagai salah satu 'emas hitam', yang justru kini malah harus diimpor dari negara lain.

Lada atau merica merupakan salah satu bumbu dapur yang paling banyak digunakan dalam masakan, khususnya masakan Indonesia. Selain itu, lada mengandung berbagai nutrisi yang baik untuk kesehatan. Oleh karena itu, menjadi sangat penting untuk mempelajari estimasi data luas areal lada sebagai dasar penentu berbagai kebijakan terkait.

Guna merumuskan kebijakan terkait luas areal lada, sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Makalah ini akan mencoba membandingkan beberapa metode statistik yang telah dikembangkan oleh para ahli, diantaranya metode *Auto-Regressive Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Auto Regression* (VAR) dalam melakukan pemodelan dan peramalan luas areal lada di Indonesia.

Oleh karena itu, tujuan dari disusunnya kegiatan ini adalah:

- f. Melakukan analisis dan peramalan data luas areal lada menggunakan model ARIMA, Fungsi Transfer dan VAR di Indonesia.
- g. Membandingkan metode tersebut dalam memperoleh ramalan data luas areal lada di Indonesia
- h. Menentukan metode terbaik dalam meramal data luas areal lada di Indonesia.

## BAHAN DAN METODE

Data yang digunakan dalam melakukan pemodelan luas areal lada adalah data *series* luas areal lada dalam satuan hektar di Indonesia tahun 1973 s.d. 2020, produksi lada dalam satuan ton di Indonesia tahun 1973 s.d. 2020, volume ekspor dan volume impor lada dalam satuan ton di Indonesia tahun 1977-2020. Peubah yang diasumsikan mempengaruhi besaran luas areal Lada untuk Fungsi Transfer adalah volume ekspor lada. Sedangkan untuk pemodelan VAR luas areal lada, peubah yang diasumsikan mempengaruhi adalah produksi lada Indonesia, volume ekspor lada, dan volume impor lada. Peningkatan volume ekspor dan penurunan impor diperkirakan akan membangkitkan minat petani untuk mengusahakan Lada.

Peramalan data luas areal lada dilakukan melalui pengujian dengan beberapa metode, yakni ARIMA, Fungsi Transfer, dan VAR menggunakan R Studio.

Penjelasan metode yang digunakan dalam melakukan peramalan luas areal Lada di Indonesia adalah sebagai berikut:

### i. ARIMA

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \theta_2 \dots \theta_p$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

#### j. Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut.

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\varphi_p(B)} \varepsilon_t$$

Dimana:

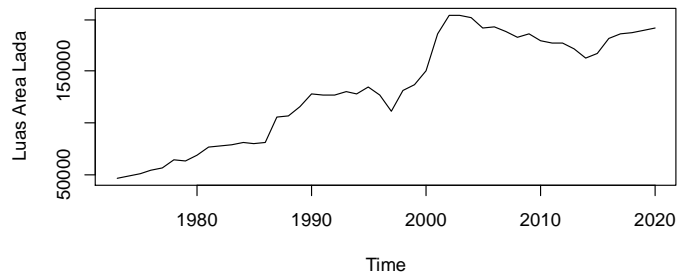
- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$

#### k. VAR

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen, karena pada kenyataannya suatu variabel dapat bertindak sebagai endogen atau dependen dari suatu variabel yang lainnya (Gujarati, Damodar. N dan Porter, Dawn. C. 2009).

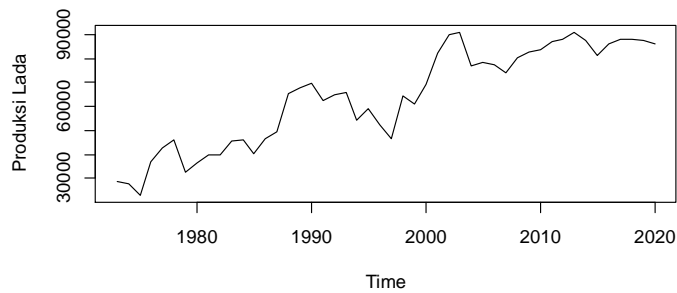
## HASIL DAN PEMBAHASAN

### e. Eksplorasi Data



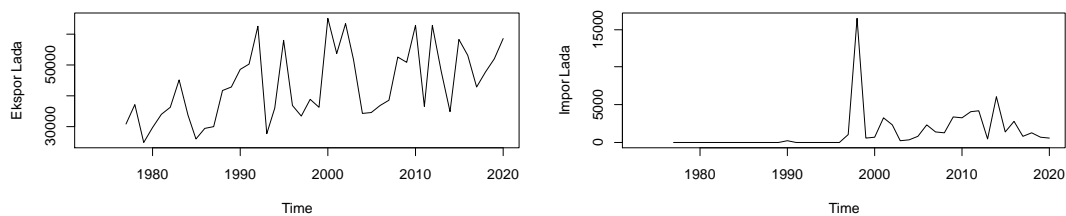
**Gambar 1.** Perkembangan Luas Areal Lada (Ha), 1973-2020

Luas areal lada di Indonesia selama periode 1973 – 2020 juga cenderung mengalami trend peningkatan dari tahun ke tahun (Gambar 1).



**Gambar 2.** Perkembangan Produksi Lada (Ton), 1973- 2020

Produksi lada di Indonesia selama periode 1973 – 2020 cenderung mengalami trend peningkatan dari tahun ke tahun (Gambar 2).



**Gambar 3.** Perkembangan Volume Eskpor Lada (Ton) 1977-2020 dan Volume Impor Lada (Ton) 1977-2020

Pergerakan volume ekspor lada selama periode 1977 – 2020 cenderung menunjukkan trend peningkatan, sedangkan pergerakan volume impor lada selama periode 1977 – 2020 menunjukkan pergerakan yang cenderung tidak stabil (Gambar.3).

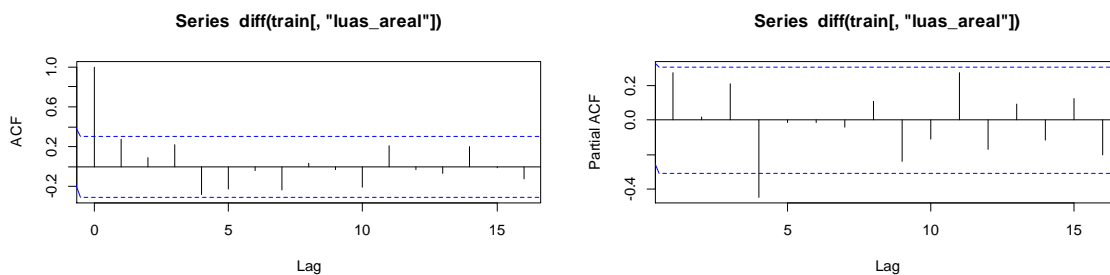
**f. ARIMA**

Dalam melakukan pemodelan luas areal lada menggunakan ARIMA, data yang digunakan adalah periode 1973-2020. Periode data tersebut kemudian dipisahkan menjadi set data *training* dan *testing*. Panjang *series* data pada data *training* adalah tahun 1973-2014, sementara data *testing* adalah periode tahun 2015-2020. Data *training* digunakan untuk melakukan penyusunan model, sementara data *testing* digunakan untuk validasi model.

Uji kestasioneran data dilakukan menggunakan uji formal statistik *Augmented Dickey-Fuller* maupun secara visual dengan melihat plot ACF dan PACF. Tabel 1 menunjukkan bahwa data luas areal lada stasioner pada *differencing* 1. Berdasarkan plot ACF dan PACF pada *differencing* 1 terlihat data sudah stasioner atau dengan kata lain rata-rata dan variannya konstan (Kirchgassner, Gebhard dan Wolters, Jurgen. 2007) (Gambar 1).

**Tabel 1.** Hasil Uji ADF

Kondisi Data	Hasil Uji ADF Luas Areal Lada
Asli	Value of test-statistic is: -1.9962 2.4249 2.2217 Critical values for test statistics: 1pct 5pct 10pct <b>Blm stasioner</b> tau3 -4.15 -3.50 -3.18 phi2 7.02 5.13 4.31 phi3 9.31 6.73 5.61
Diff 1	Value of test-statistic is: -3.6606 Critical values for test statistics: 1pct 5pct 10pct <b>stasioner</b> tau1 -2.62 -1.95 -1.61
Diff 2	Value of test-statistic is: -8.2821 Critical values for test statistics: 1pct 5pct 10pct <b>stasioner</b> tau1 -2.62 -1.95 -1.61



**Gambar 1.** Plot ACF dan Partial ACF data *differencing* 1 luas areal Lada

Selanjutnya dilakukan pendugaan model ARIMA dengan hasil seperti tertera dalam tabel 2.

**Tabel 2.** Pendugaan Model ARIMA dengan Auto Arima

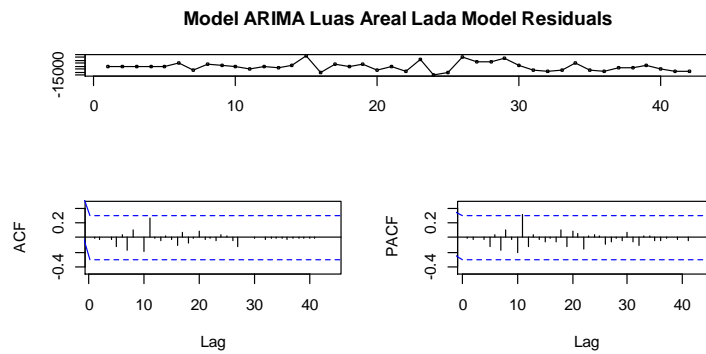
---

```
Series: train[, "luas_areal"]
ARIMA(0,1,1)
sigma^2 estimated as 86428112: log likelihood=-432.4
AIC=868.8 AICC=869.12 BIC=872.23
Training set error measures:
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	1907.116	9072.623	6446.351	1.909527	<b>5.359029</b>	0.9930543	-0.09363891

---

Dari hasil *running* program model auto ARIMA menggunakan R Studio pada luas areal lada diperoleh model ARIMA (0,1,1) dengan nilai MAPE 5,35%. Selain itu, dengan menggunakan bantuan ARMASELECT pada differencing 1 dan 2 diperoleh beberapa kombinasi AR dan MA. Dari beberapa alternatif tersebut diperoleh model ARIMA (3,2,4) dengan nilai MAPE yang cukup kecil yaitu 4,31% dengan didukung hasil pemeriksaan *residual* yang sudah cukup stasioner pada rata-rata sebagaimana terlihat pada Gambar 2. Selain itu, hasil pemeriksaan kenormalan *residual* menggunakan Ljung Box juga menunjukkan bahwa *residual* sudah secara signifikan mendekati distribusi normal pada lag 5,10,15,20,25,30,35, dan 40. Hal tersebut terlihat dari nilai p yang signifikan atau di atas taraf 5 persen (Tabel 3).



**Gambar 2.** Grafik Pemeriksaan Model *Residual* ARIMA pada Data *Training*

**Tabel 3.** Hasil Uji L-Jung test *Residual* ARIMA pada Data *Training*

Lag	Box-Ljung test Luas Areal Lada	Hasil
5	X-squared=2.1904, df= 5, p-value = 0.9751	Sign normal
10	X-squared=8.9751, df= 10, p-value = 0.8673	Sign normal
15	X-squared=9.3678, df= 15, p-value = 0.8208	Sign normal
20	X-squared=10.095, df= 20, p-value = 0.9081	Sign normal
25	X-squared=10.204, df= 25, p-value = 0.9799	Sign normal
30	X-squared=10.414, df= 30, p-value = 0.9909	Sign normal
35	X-squared=10.454, df = 35, p-value = 0.9988	Sign normal
40	X-squared=10.806, df = 40, p-value = 0.9999	Sign normal

Selanjutnya dilakukan peramalan data *testing* diikuti penghitungan nilai MAPE. Tabel 4 menunjukkan perbandingan nilai MAPE data *training* dan data *testing*, terlihat bahwa nilai MAPE luas areal lada cukup bagus yaitu 4,31% dan 15,06%.



**Tabel 4.** Perbandingan MAPE Data *Training* dan Data *Testing*

Luas Area1	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-571.1859	7221.629	5524.588	-0.2754542	<b>4.312452</b>	0.8510574	-0.00552611
Test set	27942.7759	28917.852	27942.776	15.0553181	<b>15.055318</b>	4.3045580	NA

Hasil pengepasan model ARIMA (3,2,4) untuk luas areal lada pada keseluruhan data *training* dan *testing* diperoleh model sebagai tertera pada Tabel 5 dengan nilai MAPE 4,3% untuk luas areal Lada.

**Tabel 5.** Hasil Pengepasan Model ARIMA untuk Keseluruhan Data

```

arima(x = total[, "luas_areal"], order = c(3, 2, 4))
Coefficients:
      ar1      ar2      ar3      ma1      ma2      ma3      ma4
-0.2495 -0.3317 -0.2249 -0.3336 -0.2011  0.2871 -0.7523
s.e.    0.2900  0.4164  0.2959  0.2923  0.4023  0.3530  0.3716
sigma^2 estimated as 59181715: log likelihood = -479.49, aic = 974.98
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -294.8986 7530.995 5712.601 -0.1276427 4.303317 0.9101649 -0.006740618

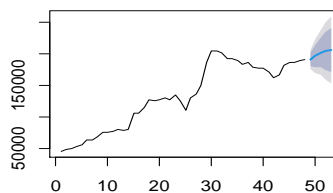
```

Hasil peramalan 5 tahun ke depan dengan menggunakan model ARIMA (3,2,4) untuk luas areal Lada tertera pada Tabel 6. dengan plot yang terlihat pada Gambar 3. Dari plot terlihat bahwa nilai ramalan cenderung mengalami kenaikan pada 5 tahun mendatang.

**Tabel 6.** Peramalan Luas Areal ARIMA (3,2,4) 2021-2025

Point Forecast	Luas areal
49	191259.2
50	196681.0
51	202413.5
52	204758.6
53	206551.1

Forecasts from ARIMA(3,2,4)



**Gambar 3.** Plot peramalan Luas Areal ARIMA (3,2,4)

### g. Fungsi Transfer

Peubah yang diasumsikan dapat mempengaruhi gejala luas areal lada adalah volume ekspor lada sehingga akan digunakan untuk melakukan analisis dengan model fungsi transfer. Panjang *series* data pada data *training* adalah tahun 1977-2014, sementara data *testing* adalah periode tahun 2014-2020. Dari pola data dan uji *Augmented Dickey-Fuller* (Tabel 7 dan Gambar 4) volume ekspor lada terlihat bahwa data belum stasioner sehingga akan dilakukan *differencing* terlebih dahulu sebelum digunakan dalam permodelan. Hasil pengujian deret *differencing* (Tabel 8) menunjukkan bahwa data *differencing* secara nyata sudah stasioner (nilai uji *mutlak* > nilai kritis *mutlak*), oleh karena itu bisa dilanjutkan dengan identifikasi model ARIMA untuk volume ekspor lada sebagai peubah input.

**Tabel 7.** Uji Augmented Dickey-Fuller Volume Ekspor Lada

---

F-statistic: 3.455 on 2 and 34 DF, p-value: **0.04306**  
 Value of test-statistic is: **-0.5443**  
 Critical values for test statistics:  
     1pct 5pct 10pct  
 tau1 **-2.62** -1.95 -1.61

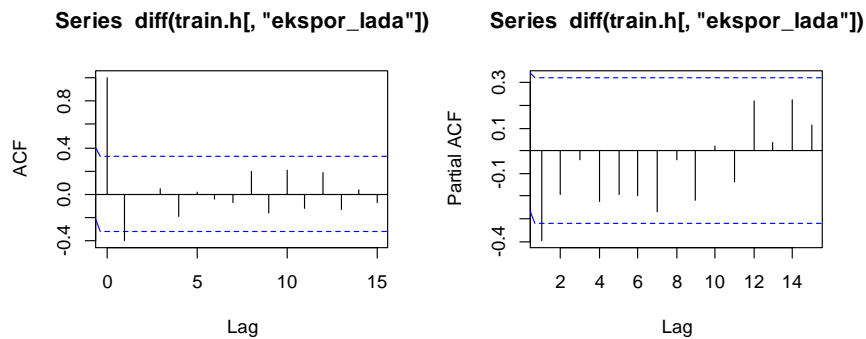
---

**Tabel 8.** Uji Augmented Dickey-Fuller untuk *differencing* Volume Ekspor Lada

---

F-statistic: 40.81 on 2 and 33 DF, p-value: **1.195e-09**  
 Value of test-statistic is: **-5.7811**  
 Critical values for test statistics:  
     1pct 5pct 10pct  
 tau1 **-2.62** -1.95 -1.61

---



**Gambar 4.** ACF dan PACF Data *Differencing* 1 Volume Ekspor Lada

Pendugaan model ARIMA dilakukan dengan memperhatikan hasil uji ADF dan pola dari ACF dan PACF yang signifikan. Model ARIMA deret input volume ekspor lada data *training* ini akan digunakan untuk tahap *prewhitening*. Dengan menggunakan auto ARIMA untuk data training diperoleh model ARIMA (0,1,1) dengan MAPE 21,41 %. ARIMA (0,1,1) akan menghasilkan ramalan yang sama untuk periode ke depan sehingga dilakukan *trial error* dengan menggunakan ordo AR dan MA yang lain hingga diperoleh model ARIMA (4,1,5) dengan MAPE 15,63 %.

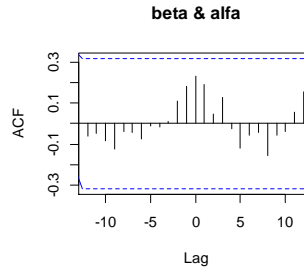
**Tabel 9.** Pendugaan Model ARIMA Deret Input Volume Ekspor Lada Data *Training*

---

```
Call:
arima(x = train.h[, "ekspor_lada"], order = c(4, 1, 5))
Coefficients:
      ar1      ar2      ar3      ar4      ma1      ma2      ma3      ma4      ma5
-0.0818  0.6811  0.0574 -0.8345 -0.6561 -1.1229  0.5989  1.0443 -0.5791
s.e.    0.1505  0.1231  0.1483  0.1460  0.2531  0.2064  0.3820  0.2041  0.2669
sigma^2 estimated as 70246246:  log likelihood = -391.39,  aic = 802.78
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 1500.619 8270.289 6563.741 0.1763659 15.63375 0.6377836 -0.1092505
```

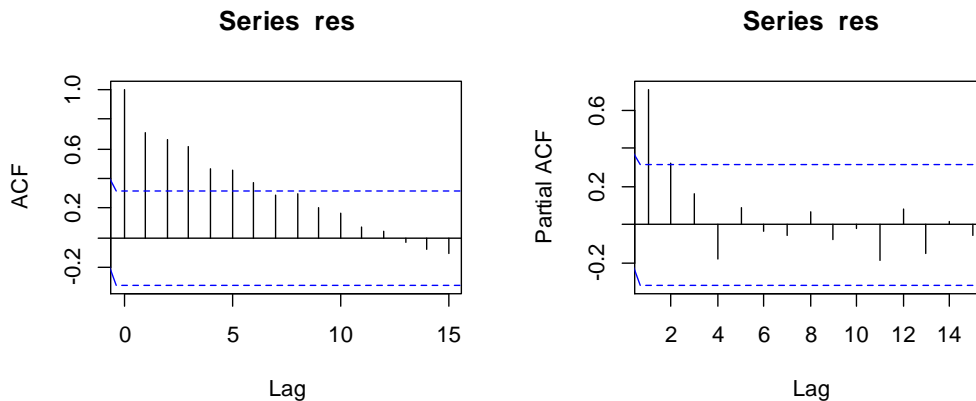
---

Selanjutnya dilakukan *prewhitening* dengan terlebih dahulu menentukan nilai alfa dan beta. Alfa merupakan nilai *residual* dari model arima variabel input yang dalam hal ini adalah volume ekspor lada, sedangkan beta merupakan nilai *residual* dari model arima output yang dalam hal ini adalah luas areal lada. Korelasi nilai beta dan alfa akan menentukan nilai r, b, dan s yang akan digunakan dalam permodelan fungsi transfer.



**Gambar 5.** *Cross Corelation* dari Beta dan Alfa Data Training Luas Areal Lada

Berdasarkan grafik *cross corelation* beta dan alfa pada Gambar 5. dapat dilihat bahwa pengaruh dari input volume ekspor lada kepada luas areal Lada Indonesia terjadi pada saat yang sama tanpa jeda, sehingga nilai koefisien  $b$  dan  $s$  adalah 0, begitu juga halnya dengan nilai  $r$  atau jeda pengaruh variabel output sendiri adalah 0. Selanjutnya dilakukan pengepasan model awal data *training* dengan membuat permodelan fungsi transfer awal menggunakan koefisien  $b$ ,  $s$ , dan  $r$  untuk mendapatkan nilai *residual*. Plot ACF dan PACF *residual* data *training* dapat dilihat pada Gambar 5 sedangkan hasil model hasil pendugaan residual dapat dilihat pada Tabel 10.



**Gambar 5.** Plot ACF dan PACF *Residual* Pendugaan Fungsi Transfer Data *Training*

**Tabel 10.** Hasil Pendugaan Model ARIMA Residual Data *Training* dengan Auto Arima

---

```
Series: res
ARIMA(0,1,1)
Coefficients:
      ma1
      -0.5022
s.e.      0.1406
sigma^2 estimated as 6.3e+08: log likelihood=-426.97
AIC=857.94  AICC=858.3  BIC=861.17
Training set error measures:
```

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	4618.091	24430.31	18880.79	-298.4057	<b>414.9619</b>	0.8751202	-0.09535892

---

Dengan auto arima, diperoleh model ARIMA *residual* (0,1,1) untuk luas areal lada. Nilai MA1 model ARIMA *residual* luas areal akan memberikan estimasi yang tidak jauh berbeda sehingga model ini tidak dapat digunakan dan perlu mencari model lain. Dengan bantuan ARMA SELECT diperoleh model ARIMA *residual* (2,2,0) sehingga model ini dapat digunakan untuk permodelan

fungsi transfer lebih lanjut meskipun nilai MAPE cukup tinggi. Selanjutnya dilakukan pendugaan model fungsi transfer dengan menggunakan model ARIMA *residual* (2,2,0) untuk output luas areal lada dan input volume ekspor lada pada data *training* (Tabel 11).

**Tabel 11.** Pendugaan Fungsi Transfer Arima *Residual* (2,2,0) Output Luas Areal Lada dengan Input Volume Eskpor Lada Data *Training*

---

```
Series: train.h[, "luas_areal"]
Regression with ARIMA(2,2,0) errors
Coefficients:
      ar1      ar2      xreg
      -0.6001  -0.4087  0.1018
s.e.      0.1916   0.1472  0.1295
sigma^2 estimated as 1.07e+08: log likelihood=-382.59
AIC=773.17  AICC=774.46  BIC=779.51
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set  -644.3046  9639  6843.417  -0.468848  5.037419  0.9881805  0.09334488
```

---

Nilai MAPE pada tabel 11 menunjukkan bahwa model ARIMA *residual* (2,2,0) cukup bagus yaitu 5,04%. Berikutnya dilakukan peramalan dengan data *testing* menggunakan input nilai aktual dan nilai ramalan, hasil dapat dilihat pada Tabel 12.

**Tabel 12.** Peramalan Data *Testing* menggunakan Input Nilai Aktual dan Ramalan

---

```
Input Aktual Series: test.h[, "luas_areal"]
Regression with ARIMA(2,2,0) errors
Coefficients:
      ar1      ar2      xreg
      -0.6001  -0.4087  0.1018
s.e.      0.0000   0.0000  0.0000

sigma^2 estimated as 1.07e+08: log likelihood=-40.79
AIC=83.59  AICC=85.59  BIC=82.98
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set  -3757.263  4946.334  3781.365  -1.999065  2.013447  0.7876214  0.3318222

Input Ramalan Series: test.h[, "luas_areal"]
Regression with ARIMA(2,2,0) errors
Coefficients:
      ar1      ar2      xreg
      -0.6001  -0.4087  0.1018
s.e.      0.0000   0.0000  0.0000
sigma^2 estimated as 1.07e+08: log likelihood=-40.87
AIC=83.73  AICC=85.73  BIC=83.12
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set  -3764.227  5037.173  3788.481  -2.004515  2.018988  0.7891038  0.2193194
```

---

Selanjutnya dilakukan pendugaan ulang model input untuk keseluruhan data *training* dan *testing*, hasil dapat dilihat pada Tabel 13. Terlihat nilai MAPE masih cukup bagus yaitu 4,9% untuk model luas area lada. Sedangkan ramalan untuk lima tahun ke depan dapat dilihat pada Tabel 14, terlihat bahwa nilai ramalan cenderung mengalami kenaikan.

**Tabel 13.** Pendugaan Ulang Model Fungsi Transfer Keseluruhan Data

---

```
Model Information:
Series: dataestimasi[, "luas_areal"]
Regression with ARIMA(2,2,0) errors
Coefficients:
      ar1      ar2      xreg
      -0.5231  -0.4005  0.0680
s.e.      0.1680   0.1370  0.1115
sigma^2 estimated as 100501968: log likelihood=-445.23
AIC=898.46  AICC=899.54  BIC=905.41
Error measures:
```

---

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-232.4646	9438.288	6820.652	-0.2171161	<b>4.895155</b>	1.028796	0.07617186

**Tabel 14.** Peramalan Fungsi Transfer (0,0,0) (2,2,0) Luas Areal Lada 2021-2025

Point Forecas Luas Area1	
45	192120.2
46	193854.3
47	195281.0
48	196407.9
49	198449.9

## h. Model VAR

Pemodelan dengan menggunakan model VAR dilakukan dengan menggunakan peubah luas areal, produksi lada, volume ekspor lada, dan volume impor lada. Hal ini diasumsikan bahwa apabila terjadi gejolak produksi lada, volume ekspor lada, dan volume impor lada akan mempengaruhi pergerakan luas areal. Panjang *series* data pada data *training* adalah tahun 1977-2014, sementara data *testing* adalah periode tahun 2015-2020. Panjang *lag* ( $p$ ) dalam model VAR ( $\text{lag}=p$ ) ditentukan dengan melakukan *trial error* menggunakan nilai  $p$  dari 1 s.d 6 lalu memperhatikan hasil signifikansi lag baik dengan maupun tanpa menggunakan konstanta dan *trend*. Dari hasil *trial error* diperoleh nilai  $p=5$  dengan *type* tanpa menggunakan konstanta dan *trend*. Tipe tanpa penggunaan konstanta dan *trend* dipilih karena dari sekian banyak *trial error* diperoleh informasi bahwa kedua hal tersebut seringkali tidak signifikan. Hasil pendugaan model VAR (5) dapat dilihat pada tabel 15. *R-squared* sebesar 99,67% menunjukkan bahwa peubah produksi lada, volume ekspor lada, dan volume impor lada tanpa konstanta dan *trend* mampu menjelaskan 99,67% keragaman luas areal Lada.

**Tabel 15.** Pendugaan Model VAR (5) Data *Training*

```

VAR Estimation Results:
=====
Endogenous variables: luas_areal, prod_lada, ekspor_lada, impor_lada
Deterministic variables: none
Sample size: 33
Log Likelihood: -1270.943
Roots of the characteristic polynomial:
1.036 0.9645 0.9645 0.9451 0.9451 0.9286 0.9286 0.9145 0.8598 0.8598 0.8369 0.8369 0.8
318 0.8318 0.8217 0.8217 0.6821 0.6821 0.6666 0.2578
Call:
VAR(y = lada[5:42, c(2, 3, 5, 6)], p = 5, type = "none")
Estimation results for equation luas_areal:
=====
luas_areal = luas_areal.l1 + prod_lada.l1 + ekspor_lada.l1 + impor_lada.l1 + luas_area
l.l2 + prod_lada.l2 + ekspor_lada.l2 + impor_lada.l2 + luas_areal.l3 + prod_lada.l3 +
ekspor_lada.l3 + impor_lada.l3 + luas_areal.l4 + prod_lada.l4 + ekspor_lada.l4 + impor
_lada.l4 + luas_areal.l5 + prod_lada.l5 + ekspor_lada.l5 + impor_lada.l5
      Estimate Std. Error t value Pr(>|t|)
luas_areal.l1  1.48520    0.31171   4.765 0.00037 ***
prod_lada.l1   0.82605    0.42442   1.946 0.07356 .
ekspor_lada.l1 -0.24217    0.20942  -1.156 0.26833 .
impor_lada.l1  -0.80959    0.97383  -0.831 0.42079
luas_areal.l2  -0.56138    0.41572  -1.350 0.19993
prod_lada.l2   -1.30772    0.54549  -2.397 0.03225 *
ekspor_lada.l2 -0.22494    0.19384  -1.160 0.26673
impor_lada.l2   1.51364    0.94457   1.602 0.13306
luas_areal.l3   0.58066    0.42568   1.364 0.19570
prod_lada.l3    0.41428    0.57285   0.723 0.48237
ekspor_lada.l3  0.71966    0.23472   3.066 0.00902 **
impor_lada.l3   0.14203    0.83910   0.169 0.86820
luas_areal.l4  -1.09265    0.45022  -2.427 0.03051 *

```

prod_lada.l4	0.19101	0.47390	0.403	0.69345
ekspor_lada.l4	-0.29212	0.27657	-1.056	0.31010
impor_lada.l4	0.82570	0.80717	1.023	0.32499
luas_areal.l5	0.50531	0.28502	1.773	0.09966
prod_lada.l5	0.06747	0.42847	0.157	0.87730
ekspor_lada.l5	0.04052	0.25044	0.162	0.87395
impor_lada.l5	-1.45956	0.74553	-1.958	0.07208
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 8644 on 13 degrees of freedom				
Multiple R-Squared: 0.9987, Adjusted R-squared: <b>0.9967</b>				
F-statistic: 496.1 on 20 and 13 DF, p-value: 1.374e-15				

**Tabel 16.** Uji Asumsi Klasik Residual Model VAR (5) Data *Training* dan Keseluruhan Data

<b>Portmanteau Test (asymptotic) Training</b>	
Chi-squared=176.22, df=176, p-value=	<b>0.4812</b> Chi-squared=176.32, df=350, p-value=
<b>JB-Test (multivariate)</b>	
Chi-squared=66.31, df=8, p-value=	<b>0.5769</b> Chi-squared=19.531, df= 8, p-value=0.01226
<b>Skewness only (multivariate)</b>	
Chi-squared=4.2437, df=4, p-value=	<b>0.374</b> Chi-squared=8.6878, df=4, p-value=
<b>Kurtosis only (multivariate)</b>	
Chi-squared=2.3873, df=4, p-value=	<b>0.6649</b> Chi-squared=10.843, df=4, p-value=0.02838
<b>ARCH (multivariate) Training</b>	
Chi-squared=280, df=500, p-value=	<b>1</b> Chi-squared=340, df=500, p-value=

Dari hasil pengujian asumsi non autokorelasi, normalitas, dan homoskedastisitas, diperoleh informasi bahwa *residual* yang diperoleh dari model VAR (5) data *training* memenuhi ketiga asumsi tersebut, dibuktikan dengan nilai *p-value* yang tidak signifikan (di atas 0,05). Selanjutnya dilakukan ramalan pada data *testing* dan pengepasan model pada seluruh data. Hasil pengujian normalitas, non autokorelasi, serta homoskedastisitas dengan model VAR (5) pada data keseluruhan menunjukkan bahwa *residual* yang diperoleh memenuhi ketiga asumsi tersebut yang ditunjukkan dengan nilai *p-value* yang tidak signifikan atau lebih dari 0,05 (Tabel 21).

**Tabel 17.** MAPE Data *Training*, *Testing*, dan Pengepasan Keseluruhan Data Model VAR (5)

Luas Areal Lada	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
Training	0.1718	1.2870	2.9594	<b>3.2785</b>	4.2943	13.4093
Testing	1.209	2.131	5.422	<b>5.363</b>	7.513	10.884
Keseluruhan Data	0.03209	0.78237	2.60134	<b>3.27161</b>	4.22184	13.95998

Dari hasil perhitungan MAPE data *training*, *testing*, dan pengepasan keseluruhan data diperoleh informasi MAPE data *testing* sudah cukup baik, yaitu sebesar 5,36% dan 3,27% untuk data *training*. Sedangkan ramalan untuk lima tahun ke depan dapat dilihat pada Tabel 18, terlihat bahwa nilai ramalan cenderung mengalami kenaikan.

**Tabel 18.** Peramalan VAR (5) Luas Areal Lada 2021-2025

Fcst	luas areal Lada
[1,]	187684.9
[2,]	185230.9
[3,]	184677.9
[4,]	185711.2
[5,]	190024.1

**Tabel 19.** Rekapitulasi MAPE *Training*, *Testing*, dan Pengepasan Keseluruhan Data Model ARIMA, Fungsi Transfer, dan VAR

LUAS AREAL LADA	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
ARIMA							
Training	-571.1859	7221.629	5524.588	-0.2754542	<b>4.312452</b>	0.8510574	-0.00552611
Test	27942.7759	28917.852	27942.776	15.053181	<b>15.055318</b>	4.3045580	NA

Total	-294.8986	7530.995	5712.601	-0.1276427	<b>4.303317</b>	0.9101649	-0.006740618	
Fungsi Transfer		ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training	-644.3046	9639	6843.417	-0.468848	<b>5.037419</b>	0.9881805	0.09334488	
Test (aktual)	-3757.263	4946.334	3781.365	-1.999065	<b>2.013447</b>	0.7876214	0.3318222	
Test (ramalan)	-3764.227	5037.173	3788.481	-2.004515	<b>2.018988</b>	0.7891038	0.2193194	
Total	-232.4646	9438.288	6820.652	-0.2171161	<b>4.895155</b>	1.028796	0.07617186	
VAR		Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	
Training		0.1718	1.2870	2.9594	<b>3.2785</b>	4.2943	13.4093	
Testing		1.209	2.131	5.422	<b>5.363</b>	7.513	10.884	
Total		0.03209	0.78237	2.60134	<b>3.27161</b>	4.22184	13.95998	

Dari hasil rekapitulasi nilai MAPE *training*, *testing*, dan pengepasan keseluruhan data model ARIMA, Fungsi Transfer, dan VAR, terlihat bahwa MAPE data *testing* model fungsi transfer memberikan nilai yang paling kecil yaitu 2,02% untuk peramalan luas areal lada.

## KESIMPULAN DAN SARAN

### Kesimpulan

- Hasil analisis yang didapat dari tiga model adalah sebagai berikut:

Peramalan Luas Areal Lada

- Model ARIMA terbaik adalah ARIMA (3,2,4) dengan MAPE data *training* 4,31% dan MAPE data *testing* 15,06%.
- Model Fungsi Transfer terbaik dengan peubah input volume ekspor lada adalah ARIMA (0,0,0)(2,2,0) dengan model input ARIMA (4,1,5) dengan MAPE data *training* 5,04% dan MAPE data *testing* 2,02%.
- Model VAR terbaik dengan memasukkan peubah produksi lada, volume ekspor lada, dan volume impor lada adalah VAR(5) type “none” dengan MAPE data *training* 3,28% dan MAPE data *testing* 5,36%.
- Dari ketiga model yang diujicobakan, model terbaik adalah fungsi transfer dengan nilai MAPE data *testing* terkecil yaitu 2,02% untuk luas areal lada.
- Hasil ramalan fungsi transfer luas areal lada untuk lima tahun ke depan adalah 192.120 Ha, 193.854 Ha, 195.281 Ha, 196.407 Ha, dan 198.449 Ha.

### Saran

- Perlu dilakukan kajian mendalam dengan metode peramalan lainnya.
- Untuk model fungsi transfer dan VAR perlu diujicobakan dengan menggunakan variabel lain yang sekiranya lebih berpengaruh, baik secara teoritis maupun praktis.
- Perlu dikaji metode peramalan untuk data non parametrik, sehingga tidak terhalang oleh tidak terpenuhinya berbagai asumsi.

## DAFTAR PUSTAKA

- Direktorat Jenderal Perkebunan. 2019. Statistik Perkebunan Indonesia 2018-2020 Jakarta: Direktorat Jenderal Perkebunan, Kementrian Pertanian. Statistik Perkebunan Indonesia 2018-2020.
- Gujarati, Damodar. N dan Porter, Dawn. C. 2009. Basic Econometrics. Boston: Douglas Reiner.
- Kirchgassner, Gebhard dan Wolters, Jurgen. 2007. Introduction to Modern Time Series Analysis. Berlin: Springer Berlin Heidelberg Newyork.
- Wei, William WS. 2006. Time Series Analyis. Phladelphia: Department of Statistics The Fox School of Business and Management Temple University.
- Wooldridge, Jeffrey M. 2015. Introductory Econometrics:A Modern Approach. Boston:Cegage Learning.



## Lampiran

### ARIMA

```
#Luas Areal#
```

```
## Plot Deret Waktu ##
```

```
lada <- ts(Lada, start=c(1973), end=c(2020), frequency=1)  
ts.plot(lada["luas_areal"], type="l", ylab="Luas Area Lada")
```

```
## Pemeriksaan Kestasioneran ##
```

```
#install.packages("urca")  
library(urca)  
stasioner <- ur.df(lada["luas_areal"],type="trend")  
summary(stasioner)
```

```
ts.plot(lada["luas_areal"], type="l", ylab="Luas Areal Lada")  
ts.plot(diff(lada["luas_areal"]), type="l", ylab="Diff. Luas Areal Lada")  
ts.plot(diff(lada["luas_areal"],differences = 2), type="l", ylab="Diff. Luas Areal Lada")
```

```
stasioner2 <- ur.df(diff(lada["luas_areal"]),type="none")  
summary(stasioner2)  
stasioner3 <- ur.df(diff(lada["luas_areal"],differences = 2),type="none")  
summary(stasioner3)
```

```
# Pembagian data training testing
```

```
train=lada[1:42,]  
test=lada[43:48,]  
total=lada[1:48,]
```

```
acf(train["luas_areal"])  
pacf(train["luas_areal"])
```

```
acf(diff(train["luas_areal"]))          # Differences ordo 1  
pacf(diff(train["luas_areal"]))        # Differences ordo 1
```

```
acf(diff(train["luas_areal"],differences = 2))      # Differences ordo 2  
pacf(diff(train["luas_areal"],differences = 2))    # Differences ordo 2
```

```
## Pendugaan Model ARIMA ##
```

```
#install.packages("forecast")  
library(forecast)  
tmodel <- auto.arima(train["luas_areal"],seasonal = FALSE)  
summary(tmodel)
```

```
##Pemilihan kombinasi AR MA yang lain##
```

```
library(caschrono)
```

```

minic1 <- armaselect(diff(train[,"luas_areal"]), max.p = 5, max.q = 5, nbmod = 10)
minic1
minic2 <- armaselect(diff(train[,"luas_areal"],differences = 2), max.p = 5, max.q = 5, nbmod =
10)
minic2
tmodel <- arima(train[,"luas_areal"], order=c(3,2,4))
summary(tmodel)

## Peramalan data testing (terpilih ARIMA (3,2,4))##
ramalan_arima = forecast(tmodel, 6)
forecast_arima <- ramalan_arima$mean
forecast_arima
accuracy(ramalan_arima,test[,"luas_areal"])

## Pemeriksaan Sisaan ##
tsdisplay(residuals(tmodel), lag.max=44, main='Model ARIMA Luas Areal Lada Model
Residuals')
#install.packages('remotes')
library(remotes)
#install_github('cran/portes',upgrade = 'always')
#install.packages('portes')
library(portes)
ljbtest <-
LjungBox(residuals(tmodel),lags=seq(5,30,5),order=0,season=1,squared.residuals=FALSE)
ljbtest
boxtest<-Box.test(residuals(tmodel),lag=5,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=10,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=15,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=20,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=25,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=30,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=35,type=c("Ljung-Box"),fitdf=0)
boxtest
boxtest<-Box.test(residuals(tmodel),lag=40,type=c("Ljung-Box"),fitdf=0)
boxtest

## Pengepasan model untuk seluruh data ##
model.arima <- arima(total[,"luas_areal"],order=c(3,2,4))
summary(model.arima)

```

```
ramalan_arima2 = forecast(model.arima, 5)
ramalan_arima2
plot(ramalan_arima2)
```

## **FUNGSI TRANSFER**

```
#Luas Areal#
```

```
lada <- ts(Lada, start=c(1973),end=c(2020),frequency=1)
ts.plot(lada["luas_areal"], type="l", ylab="Luas Areal")
ts.plot(lada["ekspor_lada"], type="l", ylab="Ekspor Lada")
```

```
##### Pembagian data: training-testing #####
```

```
train.h=lada[5:42,]
test.h=lada[43:48,]
```

```
##### Pemeriksaan kestasioneran #####
```

```
library(urca)
stasioner <- ur.df(train.h["ekspor_lada"],type="none")
summary(stasioner)
stasioner <- ur.df(diff(train.h["ekspor_lada"]),type="none")
summary(stasioner)
#acf(train.h["ekspor_lada"])
#pacf(train.h["ekspor_lada"])
acf(diff(train.h["ekspor_lada"]))
pacf(diff(train.h["ekspor_lada"]))
```

```
##### Pencarian model tentatif model input #####
```

```
library(forecast)
tmodel <- auto.arima(train.h["ekspor_lada"],seasonal = FALSE)
summary(tmodel)
library(caschnono)
#minic0 <- armaselect(train.h["ekspor_lada"], max.p = 5, max.q = 5, nbmod = 10)
#minic0
minic1 <- armaselect(diff(train.h["ekspor_lada"]), max.p = 5, max.q = 5, nbmod = 10)
minic1
model1 <- arima(train.h["ekspor_lada"], order=c(4,1,5))
summary(model1)
library(lmtest)
coeftest(model1)
```

```
##### Prewhitening dan korelasi silang #####
```

```
alfa <- residuals(Arima(train.h["ekspor_lada"],model=model1))
beta <- residuals(Arima(train.h["luas_areal"],model=model1))
ccf(beta,alfa)
#ccf(alfa,beta)
ccf(train.h["luas_areal"],train.h["ekspor_lada"])
```

```

##### Pengepasan model (r,s, b) = (0, 0, 0) #####
tf.arima0 <- Arima(train.h[,"luas_areal"], order=c(0,0,0), xreg=train.h[,"ekspor_lada"])
#summary(tf.arima0)

##### Identifikasi model noise #####
res <- residuals(tf.arima0)
plot(res)
acf(res)
pacf(res)
#stasioner <- ur.df(diff(res,type="drift"))
#summary(stasioner)
tmodel <- auto.arima(res,seasonal = FALSE)
summary(tmodel)
minic1 <- armaselect(diff(res), max.p = 5, max.q = 5, nbmod = 10)
minic1
minic2 <- armaselect(diff((res),differences = 2), max.p = 5, max.q = 5, nbmod = 10)
minic2
modelres <- Arima(res, order=c(2,2,0))
summary(modelres)
coeftest(modelres)

##### Pengepasan model (r,s, b) = (0, 0, 0), Noise(0, 1, 1) #####
tf.arima1 <- Arima(train.h[,"luas_areal"], order=c(2,2,0), xreg=train.h[,"ekspor_lada"])
summary(tf.arima1)
coeftest(tf.arima1)

##### Peramalan berbasis fungsi transfer, data input nilai AKTUAL #####
ramalan.arima.tf1 <- Arima(test.h[,"luas_areal"], xreg=test.h[,"ekspor_lada"], model=tf.arima1)
summary(ramalan.arima.tf1)

##### Peramalan berbasis fungsi transfer, data input nilai RAMALAN #####
modell1 <- Arima(train.h[,"ekspor_lada"], order=c(4,1,5))
ramalan.input <- forecast(modell1,6)
ramalan.arima.tf2 <- Arima(test.h[,"luas_areal"], xreg=ramalan.input$mean, model=tf.arima1)
summary(ramalan.arima.tf2)

### Menduga ulang model input ###
dataestimasi <- lada[5:48,]
model.input <- Arima(dataestimasi[,"ekspor_lada"], order=c(4,1,5))
ramalan.input2 <- forecast(model.input,5)
plot(ramalan.input)

### Menduga ulang fungsi transfer ###
tf.arima.final <- Arima(dataestimasi[,"luas_areal"], order=c(2,2,0),
xreg=dataestimasi[,"ekspor_lada"])
ramalan.tf.final <- forecast(tf.arima.final,h=5, xreg=ramalan.input2$mean)

```

```
summary(ramalan.tf.final)
coefficients(ramalan.tf.final)
ramalan.tf.final$mean
plot(ramalan.tf.final)
```

## **VAR**

```
#Luas Areal#
```

```
lada <- ts(Lada, start=c(1973),end=c(2025),frequency=1)
library(vars)
library(zoo)
library(urca)
library(lmtest)
library(sandwich)
library(strucchange)
library(MASS)
varhsheet1 <- VAR(lada[5:42,c(2,3,5,6)], p=1, type="none")
summary(varhsheet1)
varhsheet2 <- VAR(lada[5:42,c(2,3,5,6)], p=2, type="none")
summary(varhsheet2)
varhsheet3 <- VAR(lada[5:42,c(2,3,5,6)], p=3, type="none")
summary(varhsheet3)
varhsheet4 <- VAR(lada[5:42,c(2,3,5,6)], p=4, type="none")
summary(varhsheet4)
varhsheet5 <- VAR(lada[5:42,c(2,3,5,6)], p=5, type="none")
summary(varhsheet5)
varhsheet6 <- VAR(lada[5:42,c(2,3,5,6)], p=6, type="none")
summary(varhsheet6)
```

```
varhsheet1 <- VAR(lada[5:42,c(2,3,5,6)], p=1, type="trend")
summary(varhsheet1)
varhsheet2 <- VAR(lada[5:42,c(2,3,5,6)], p=2, type="trend")
summary(varhsheet2)
varhsheet3 <- VAR(lada[5:42,c(2,3,5,6)], p=3, type="trend")
summary(varhsheet3)
varhsheet4 <- VAR(lada[5:42,c(2,3,5,6)], p=4, type="trend")
summary(varhsheet4)
varhsheet5 <- VAR(lada[5:42,c(2,3,5,6)], p=5, type="both")
summary(varhsheet5)
varhsheet6 <- VAR(lada[5:42,c(2,3,5,6)], p=6, type="trend")
summary(varhsheet6)
```

```
varhsheet1 <- VAR(lada[5:42,c(2,3,5,6)], p=1, type="const")
summary(varhsheet1)
varhsheet2 <- VAR(lada[5:42,c(2,3,5,6)], p=2, type="const")
summary(varhsheet2)
varhsheet3 <- VAR(lada[5:42,c(2,3,5,6)], p=3, type="const")
```

```

summary(varhsheet3)
varhsheet4 <- VAR(lada[5:42,c(2,3,5,6)], p=4, type="const")
summary(varhsheet4)
varhsheet5 <- VAR(lada[5:42,c(2,3,5,6)], p=5, type="const")
summary(varhsheet5)
varhsheet6 <- VAR(lada[5:42,c(2,3,5,6)], p=6, type="const")
summary(varhsheet6)

varhsheet1 <- VAR(lada[5:42,c(2,3,5,6)], p=1, type="both")
summary(varhsheet1)
varhsheet2 <- VAR(lada[5:42,c(2,3,5,6)], p=2, type="both")
summary(varhsheet2)
varhsheet3 <- VAR(lada[5:42,c(2,3,5,6)], p=3, type="both")
summary(varhsheet3)
varhsheet4 <- VAR(lada[5:42,c(2,3,5,6)], p=4, type="both")
summary(varhsheet4)
varhsheet5 <- VAR(lada[5:42,c(2,3,5,6)], p=5, type="both")
summary(varhsheet5)
varhsheet6 <- VAR(lada[5:42,c(2,3,5,6)], p=6, type="both")
summary(varhsheet6)

ser11 <- serial.test(varhsheet5, lags.pt = 16, type = "PT.asymptotic")
ser11
norm1 <- normality.test(varhsheet5)
norm1
arch1 <- arch.test(varhsheet5, lags.multi = 5)
arch1

impresp <- irf(varhsheet5)
plot(impresp)
plot(fevd(varhsheet5))

fit <- varhsheet5[["varresult"]][["luas_areal"]][["fitted.values"]]
mapet <- abs((lada[10:42,2] - fit)/lada[10:42,2]*100)
summary(mapet)

gabung2 <- as.data.frame(cbind(lada[10:42,1],lada[10:42,2],fit))
colnames(gabung2) <- c("Tahun", "luas_areal", "fit")
str(gabung2)
plot(gabung2$Tahun,gabung2$luas_areal, xlab='indeks t', ylab='Luas Areal Lada', type='l',
col='black', lwd=1)
lines(gabung2$Tahun,gabung2$fit, col='red', lwd=1)
legend("topleft", c("Actual", "VAR"), col=c("black","red"), lty=1,lwd=1,cex=0.6, inset=0.02,
box.lty=0)

forecast <- predict(varhsheet5, n.ahead = 6, ci = 0.95)

```

```

plot(forecast)

ramalanvar <- forecast[["fcst"]][["luas_areal"]]
ramalanvar <- ramalanvar[,1]
mape <- abs((lada[43:48,2] - ramalanvar)/lada[43:48,2]*100)
summary(mape)

gabung <- as.data.frame(cbind(lada[43:48,1],lada[43:48,2],ramalanvar))
colnames(gabung) <- c("Tahun", "luas_areal", "ramalanvar")
str(gabung)
plot(gabung$Tahun,gabung$luas_areal, xlab='indeks t', ylab='Luas Areal Lada', type='l',
col='black', lwd=1, ylim=c(120000,190000))
lines(gabung$Tahun,gabung$ramalanvar, col='red', lwd=1)
legend("topleft", c("Actual", "VAR"), col=c("black", "red"), lty=1,lwd=1,cex=0.6, inset=0.02,
box.lty=0)

colnames(gabung) <- c("Tahun", "luas_areal", "VAR")
colnames(gabung2) <- c("Tahun", "luas_areal", "VAR")
gabung3 <- rbind(gabung2,gabung)
plot(gabung3$Tahun,gabung3$luas_areal, xlab='indeks t', ylab='Luas Areal Lada', type='l',
col='black', lwd=1, ylim=c(77000,250000))
lines(gabung3$Tahun,gabung3$VAR, col='red', lwd=1)
legend("topleft", c("Actual", "VAR"), col=c("black", "red"), lty=1,lwd=1,cex=0.6, inset=0.02,
box.lty=0)

#pengepasan model pada seluruh data#
varhsheet5.all <- VAR(lada[5:48,c(2,3,5,6)], p=5, type="none")
summary(varhsheet5.all)
ser11 <- serial.test(varhsheet5.all, lags.pt = 16, type = "PT.asymptotic")
ser11
norm1 <- normality.test(varhsheet5.all)
norm1
arch1 <- arch.test(varhsheet5.all, lags.multi = 5)
arch1
fit1 <- varhsheet5.all[["varresult"]][["luas_areal"]][["fitted.values"]]
mape1 <- abs(lada[10:48,2] - fit1)/lada[10:48,2]*100
summary(mape1)
forecast <- predict(varhsheet5.all, n.ahead = 5, ci = 0.95)
plot(forecast)
ramalanvar <- forecast[["fcst"]][["luas_areal"]]
ramalanvar <- ramalanvar[,1]
ramalanvar

```





# KAJIAN METODE ESTIMASI LUAS AREAL TEH INDONESIA

**Diah Indarti - Statisticians**

*Center for Agricultural Data and Information System - Ministry of Agriculture  
Jln. Harsono RM 3 Gd D Lt 4, Ragunan-Jakarta Selatan, Indonesia*

## ABSTRAK

Komoditas teh adalah salah satu komoditas perkebunan penghasil devisa setelah minyak dan gas, hal ini ditunjang dengan luasan perkebunan teh di Indonesia tergolong cukup luas. Tujuan dari analisis kajian produksi teh Indonesia ini adalah mencari angka estimasi tahun 2022 dengan menggunakan model *Autoregressive Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Autoregression* (VAR) dengan menggunakan software RStudio. Tujuan dari ketiga model ini mencari keakurasian data agar menjadi lebih baik yang ditandai dengan *Mean Absolut Percentage Error* (MAPE) untuk data training maupun data testing semakin kecil. Berdasarkan dari hasil perbandingan MAPE disimpulkan bahwa model yang digunakan adalah model ARIMA (0,2,1) dengan MAPE training sebesar 2,63% dan data testing sebesar 3,47% dengan hasil estimasi luas areal teh Indonesia tahun 2022 sebesar 110.208 ha.

*Kata kunci: teh, ARIMA, fungsi transfer, var, MAPE*

## ABSTRACT

*Tea commodity is one of the foreign exchange-producing plantation commodities after oil and gas, this is supported by the relatively large area of tea plantations in indonesia. The purpose of this analysis of the study of indonesian tea production is to find the estimated number for 2022 using the autoregressive integrated moving average (arima), transfer function and vector autoregression (var) model using the rstudio software. The purpose of these three models is to look for data accuracy to be better, which is marked by the mean absolute percentage error (mape) for training data and testing data is getting smaller. Based on the results of the mape comparison, it is concluded that the model used is the arima model (0.2.1) with mape training of 2.63% and testing data of 3.47% with the estimated indonesian tea area in 2022 of 110,208 ha.*

*Keywords: tea, ARIMA, transfer function, var, MAPE*

## PENDAHULUAN

Teh adalah salah satu minuman yang paling banyak dikonsumsi dunia. Teh diduga berasal dari Tiongkok. Temperatur dan kelembaban yang konstan adalah keadaan ideal untuk pertumbuhan tanaman teh. Kondisi tersebut dapat ditemukan di wilayah iklim tropis dan subtropis di Asia tempat lebih dari 60% teh dunia diproduksi. Dataran tinggi yang dingin merupakan tempat paling baik untuk memproduksi daun teh berkualitas tinggi. Tanaman teh dapat dipanen untuk pertama kalinya setelah mencapai usia kira-kira empat tahun. Ketika panen, hanya daun-daun muda yang dipilih, mengimpilkasikan bahwa pemetikan manual lebih efisien dibandingkan menggunakan peralatan mesin.

Estimasi produksi teh beberapa tahun ke depan sangat diperlukan sebagai bahan penentu kebijakan maupun *Early Warning System (EWS)* mengingat data Angka Tetap (ATAP) Perkebunan memiliki lag yang cukup jauh dibandingkan data tahun berjalan. Pada kajian metode estimasi produksi teh Indonesia ini menggunakan model *univariate* maupun model *multivariate*. Metode yang digunakan antara lain *Autoregressive Integrated Moving Average (ARIMA)*, Fungsi Transfer dan *Vector Autoregression (VAR)*. Metode *Autoregressive Integrated Moving Average (ARIMA)* menghasilkan angka estimasi produksi teh tanpa ada pengaruh dari variabel lain. Metode Fungsi Transfer menghasilkan angka estimasi produksi teh dengan memasukkan intervensi dari satu variabel pendukung yang dianggap paling berpengaruh terhadap produksi. Metode VAR menghasilkan angka estimasi produksi teh dengan mempertimbangkan pengaruh dari beberapa variabel lain atau lebih dari satu variabel pendukung yang diduga berpengaruh terhadap produksinya.

Saat ini rilis resmi data produksi teh oleh Badan Pusat Statistik (BPS) adalah data tahunan dengan lag  $n-1$  tahun. Guna merumuskan kebijakan produksi, pasokan dan distribusi komoditas teh sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Metode ramalan produksi teh dapat digunakan dengan menggunakan beberapa metode statistik yang telah dikembangkan oleh para ahli.

Pada makalah ini akan dikaji metode ARIMA, Fungsi Transfer dan VAR dalam melakukan pemodelan dan peramalan produksi teh Indonesia.

Tujuan dari disusunnya makalah ini adalah:

- i. Melakukan analisis dan peramalan data produksi teh menggunakan model ARIMA, Regresi, Fungsi Transfer dan VAR di Indonesia

- j. Membandingkan ketiga metode tersebut dalam memperoleh ramalan data produksi teh.
- k. Menentukan metode terbaik dalam meramal data produksi teh di Indonesia.

## **METODOLOGI**

### 1. Sumber Data

Data yang digunakan dalam kajian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik dan Direktorat Jenderal Perkebunan. Data-data tersebut terdiri dari variabel luas areal teh, produksi teh, luas tanaman menghasilkan teh, volume ekspor teh, volume impor teh dari tahun 1980 hingga tahun 2020. Data tahun 1980 sampai tahun 2014 digunakan untuk membangun model sedangkan data tahun 2015 sampai tahun 2020 digunakan untuk validasi model.

### 2. Software

Software yang digunakan dalam menyusun makalah ini menggunakan software aplikasi R-Studio. Keunggulan dari software R-Studio ini merupakan software yang open source sehingga tidak memerlukan biaya untuk pembelian maupun perpanjangan lisensi. Keunggulan lain dari R adalah mudah dalam melakukan transformasi dan pemrosesan data. Karena R adalah program untuk analisis data, maka kemampuan R dalam transformasi data seperti penyiapan data, import dan export data dalam berbagai format, dan lain-lain.

### 3. Tinjauan Literature

Metode statistik yang digunakan dalam peramalan ini menggunakan peubah tunggal maupun peubah ganda. Metode yang digunakan antara lain ARIMA (*Auto Regressive Integrated Moving Average*), Fungsi Transfer dan VAR (*Vector Autoregression*).

#### **a. *Autoregressive Integrated Moving average (ARIMA)***

*Autoregressive Integrated Moving average (ARIMA)* atau biasa disebut juga dengan metode time series Box Jenkins, sangat sesuai digunakan untuk melakukan peramalan jangka pendek, sementara untuk peramalan jangka panjang kurang baik ketepatannya. Metode ARIMA merupakan metode yang hanya menggunakan peubah dependen dan mengabaikan peubah independen sewaktu melakukan peramalan.

Metode ARIMA dibagi kedalam tiga kelompok model, yaitu *autoregressive model* (AR), *moving average model* (MA) dan model campuran yang memiliki karakteristik kedua model di atas yaitu *autoregressive integrated moving average* (ARIMA).

**1) Autoregressive Model (AR)**

AR adalah suatu model yang menjelaskan pergerakan suatu peubah melalui peubah itu sendiri di masa lalu.

Model *autoregressive* orde ke-p dapat ditulis sebagai berikut:

ARIMA (p,0,0)

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t \dots\dots\dots(1)$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \dots \theta_p$  = parameter autoregresive ke-p

$\varepsilon_t$  = nilai kesalahan pada waktu ke t

**2) Moving Average Model (MA)**

MA adalah suatu model yang melihat pergerakan peubahnya melalui sisaannya di masa lalu.

Bentuk model MA dengan ordo q atau MA (q) atau model ARIMA (0,d,g) dapat ditulis sebagai berikut:

$$Y_t = \mu - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t \dots\dots\dots(2)$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$\phi_1 \dots \phi_q$  = parameter-parameter moving average

$\varepsilon_{t-q}$  = nilai kesalahan pada waktu ke (t-q)

**3) Autoregressive Intergrated Moving Everage (ARIMA)**

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t \dots\dots\dots(3)$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \theta_q \phi_1 \phi_n$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

### c. Metode Fungsi Transfer

Dalam bidang pertanian, produksi komoditas pertanian sangat dipengaruhi oleh peubah lainnya seperti serangan OPT, penggunaan saprodi, gejolak harga komoditas tersebut atau komoditas lainnya, dan lainnya. Apabila peramalan produksi hanya didasarkan pada besarnya produksi komoditas yang bersangkutan saja tanpa memperhatikan faktor-faktor yang mempengaruhinya, maka informasi untuk pembuatan perencanaan menjadi tidak lengkap, sehingga tujuan peramalan menjadi tidak tercapai secara utuh. Salah satu upaya menganalisis data deret waktu multivariat agar diperoleh hasil yang dapat memberikan informasi yang lengkap dan simultan, adalah dengan menggunakan model Fungsi Transfer.

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke-t, tetapi juga pada waktu t+1, t+2, ..., t+k. Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara peubah *input* dan peubah *output*.

Tujuan pemodelan fungsi transfer adalah untuk menetapkan model yang sederhana, yang menghubungkan deret output ( $Y_i$ ) dengan deret input ( $X_i$ ) dan gangguan/noise( $n_i$ ). Wei (1994) juga menjelaskan bahwa di dalam fungsi transfer terdapat rangkaian *output* yang mungkin dipengaruhi oleh rangkaian *multiple input*. Pada kasus *single input* peubah, dapat menggunakan metode korelasi silang yang dianjurkan oleh Box and Jenkins (1976). Teknik ini juga dapat digunakan ketika terdapat *single input* peubah yang lebih dari satu selama antar variable *input* tidak berkorelasi silang. Jika beberapa atau semua peubah input berkorelasi silang maka teknik prewhitening atau metode korelasi silang tidak dapat digunakan secara langsung. Alasan utama bagi perlunya suatu perencanaan atau peramalan adalah adanya tenggang waktu

pengambilan keputusan yang dapat berkisar dari beberapa hari atau sampai beberapa tahun. Pada analisis fungsi transfer untuk peramalan deret berkala univariate, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (noise). Deret input mempengaruhi deret output melalui sebuah fungsi transfer yang mendistribusikan pengaruhnya secara dinamis melalui beberapa periode waktu yang akan datang dengan persentase tertentu yang disebut sebagai bobot respons impuls atau bobot fungsi transfer.

Model umum Fungsi Transfer:

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\varphi_p(B)} \varepsilon_t \quad \dots\dots\dots(4)$$

Dimana:

- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$
- $q \rightarrow$  ordo MA bagi noise  $N_t$

**d. Vector Autoregression (VAR)**

*Vector Autoregression* (VAR) dikemukakan pertama kali oleh Christopher Sims (1980). Sims mengembangkan model ekonometri dengan mengabaikan pengujian asumsi secara apriori. VAR dikembangkan oleh Sims sebagai kritik atas metode simultan. Jumlah peubah yang besar dan klasifikasi endogen dan eksogen pada metode simultan merupakan dasar dari kritik tersebut. Penggunaan pendekatan struktural atas pemodelan persamaan simultan biasanya menerapkan teori ekonomi di dalam usahanya untuk mendeskripsikan hubungan antar peubah yang ingin diuji. Akan tetapi sering ditemukan bahwa teori ekonomi saja ternyata tidak cukup kaya di dalam menyediakan spesifikasi yang ketat dan tepat atas hubungan dinamis antar peubah (Gujarati, 2010). Model VAR merupakan jalan keluar atas permasalahan ini melalui pendekatan non-struktural atau tidak berdasarkan pada teori ekonomi tertentu (*atheoretical*). Metode VAR memperlakukan seluruh peubah secara simetris tanpa memperlakukan peubah dependen dan independen atau dengan kata lain model ini memperlakukan seluruh peubah sebagai peubah endogen., karena pada kenyataannya suatu peubah dapat bertindak sebagai endogen atau dependen dari suatu peubah yang lainnya (Gujarati, 2010).

Kelebihan dalam penggunaan metode VAR (Gujarati, 2010):

- e. Kemudahan dalam penggunaan, tidak perlu mengkhawatirkan tentang penentuan peubah endogen dan peubah eksogen.
- f. Kemudahan dalam estimasi, metode *Ordinary Least Square* (OLS) dapat diaplikasikan pada tiap persamaan secara terpisah.
- g. *Forecast* atau peramalan yang dihasilkan pada beberapa kasus ditemukan lebih baik daripada yang dihasilkan oleh model persamaan simultan yang kompleks.
- h. Perangkat estimasi yang digunakan adalah *Impulse Respon Function* (IRF) untuk melacak respon dari peubah dependen dalam sistem VAR terhadap *shock* dari *error term* dan *Variance Decomposition* yang memberikan informasi mengenai pentingnya masing-masing *error term* dalam mempengaruhi peubah-peubah dalam VAR.

Di sisi lain, terdapat beberapa kritik terhadap model VAR menyangkut permasalahan berikut (Gujarati, 2010) :

- 6) Model VAR merupakan model yang *atheoretic* atau tidak berdasarkan teori, hal ini tidak seperti pada persamaan simultan. Pada persamaan simultan, pemilihan peubah yang akan dimasukkan dalam persamaan memegang peranan penting dalam mengidentifikasi model.
- 7) Pada model VAR penekanannya terletak pada *forecasting* atau peramalan sehingga model ini kurang cocok digunakan dalam menganalisis kebijakan.
- 8) Permasalahan yang besar dalam model VAR adalah pada pemilihan *lag length* atau panjang lag yang tepat. Karena semakin panjang lag, maka akan menambah jumlah parameter yang akan bermasalah pada *degrees of freedom*.
- 9) Peubah yang tergabung pada model VAR harus stasioner. Apabila tidak stasioner, perlu dilakukan transformasi bentuk data, misalnya melalui *first difference*.
- 10) Sering ditemui kesulitan dalam menginterpretasi tiap koefisien pada estimasi model VAR, sehingga sebagian besar peneliti melakukan interpretasi pada estimasi fungsi *impulse respon*.

**Estimasi Model VAR**

Enders (2004) menjelaskan, ketika peneliti tidak memiliki kepastian untuk menentukan bahwa suatu peubah adalah eksogen, maka suatu perluasan analisis fungsi perpindahan alami akan memperlakukan masing-masing peubah secara simetris. Sebagai contoh, pada kasus-kasus peubah yang membiarkan alur waktu atau *time path*  $\{s_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{y_t\}$  dan membiarkan *time path*  $\{y_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{s_t\}$ .

Di dalam sistem *bivariate*, hubungan tersebut dapat digambarkan seperti pada persamaan (1.4) di bawah ini:

$$\begin{aligned} s_t &= b_{10} - b_{12}y_t + \gamma_{11}s_{t-1} + \gamma_{12}y_{t-1} + \varepsilon_{s_t} \\ y_t &= b_{20} - b_{21}s_t + \gamma_{21}s_{t-1} + \gamma_{22}y_{t-1} + \varepsilon_{y_t} \end{aligned} \dots\dots\dots(5)$$

Dengan mengasumsikan bahwa kedua peubah  $s_t$  dan  $y_t$  adalah stasioner:  $\varepsilon_{s_t}$  dan  $\varepsilon_{y_t}$  adalah *disturbances* yang memiliki rata-rata nol dan matriks kovarians terbatas atau bersifat *white noise* dengan standar deviasi yang berurutan  $\sigma_s$  dan  $\sigma_y$ : serta  $\{\varepsilon_{s_t}\}$  dan  $\{\varepsilon_{y_t}\}$  adalah *disturbances* yang independen dengan rata-rata nol dan kovarian terbatas (*uncorrelated white-noise disturbances*). Kedua persamaan di atas merupakan orde pertama VAR, karena panjang *lag* nya hanya satu. Agar persamaan (4.1) lebih mudah dipahami dan digunakan sebagai alat analisis maka ditransformasikan dengan menggunakan matriks aljabar, dan hasilnya dapat dituliskan secara bersama seperti pada persamaan di bawah ini:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Atau dengan bentuk lain:

$$Bx_t = \Gamma_0 + \Gamma_1 X_{t-1} + \varepsilon_t \dots\dots\dots(6)$$

Dimana:

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \quad x_t = \begin{bmatrix} s_t \\ y_t \end{bmatrix} \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Dengan melakukan pengalihan antara persamaan (4.2) dengan  $B^{-1}$  atau invers matriks B, maka akan dapat ditentukan model VAR dalam bentuk standar, seperti dituliskan pada persamaan di bawah ini:



$$x_t = A_0 + A_1 x_{t-1} + \ell_t \dots \dots \dots (7)$$

$$A_0 = B^{-1} \Gamma_0$$

dimana  $A_1 = B^{-1} \Gamma_1$

$$\ell_t = B^{-1} \varepsilon_t$$

Untuk tujuan notasi, maka  $\{a_{i0}\}$  dapat didefinisikan sebagai elemen ke-i dari vektor  $A_0$ ;  $\{a_{ij}\}$  sebagai elemen dalam baris ke-i dan baris ke-j dari matriks  $A_1$ ; dan  $\{e_{it}\}$  sebagai elemen ke-i dari vektor  $e_t$ . Dengan menggunakan notasi baru yang telah dijelaskan sebelumnya, maka persamaan (4.3) dapat ditulis menjadi:

$$\begin{aligned} s_t &= a_{10} + a_{11}s_{t-1} + a_{12}y_{t-1} + e_{1t} \\ y_t &= a_{20} + a_{21}s_{t-1} + a_{22}y_{t-1} + e_{2t} \dots \dots \dots (8) \end{aligned}$$

**Fungsi Impulse Response**

Fungsi *impulse response* menggambarkan tingkat laju dari *shock* peubah yang satu terhadap peubah yang lainnya pada suatu rentang periode tertentu. Sehingga dapat dilihat lamanya pengaruh dari *shock* suatu peubah terhadap peubah lain sampai pengaruhnya hilang atau kembali ke titik keseimbangan. Analisis fungsi *impulse respon* dapat dituliskan dalam bentuk *Vector Moving Avarage (VMA)* dari bentuk standar VAR pada persamaan (4.4).

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix} \dots \dots \dots (9)$$

dimana  $s_t$  dan  $y_t$  memiliki hubungan dengan  $e_{1t}$  dan  $e_{2t}$  secara berurutan. Selanjutnya dengan menggunakan operasi aljabar matriks maka *vector error* dapat ditentukan sebagai berikut :

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{st-i} \\ \varepsilon_{yt-i} \end{bmatrix} \dots \dots \dots (10)$$

dengan menggabungkan persamaan (4.5) dan (4.6) akan didapat:

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{st-i} \\ \varepsilon_{yt-i} \end{bmatrix} \dots \dots \dots (11)$$

Persamaan (4.7) dapat disederhanakan dengan mendefinisikan matriks 2x2  $\Phi_i$  dengan elemen  $\Phi_{jk}$  (i) seperti persamaan berikut :

$$\Phi_i = A_1^i / (1 - b_{12}b_{21}) \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \dots \dots \dots (12)$$

sehingga diperoleh bentuk matriks persamaan fungsi *impulse respon*:

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \sum_{i=0}^n \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{st-1} \\ \varepsilon_{yt-1} \end{bmatrix} \dots \dots \dots (13)$$

dimana :

$\Phi_{ij}(i)$  = efek dari *structural shock* pada s dan y

$\Phi_{ij}(0)$  = *impact multipliers*

$\sum \Phi_{ij}(i)$  = *cumulative multipliers*

$\sum \Phi_{ij}(i)$  pada saat  $n \rightarrow \infty$  = *long run multipliers*

### Variance Decomposition

*Variance decomposition* atau disebut juga *forecast error variance decomposition* merupakan perangkat pada model VAR yang akan memisahkan variasi dari sejumlah peubah yang diestimasi menjadi komponen-komponen *shock* atau menjadi peubah *innovation*, dengan asumsi bahwa peubah-peubah *innovation* tidak saling berkorelasi. Kemudian, *variance decomposition* akan memberikan informasi mengenai proporsi dari pergerakan pengaruh *shock* pada sebuah peubah terhadap *shock* peubah yang lain pada periode saat ini dan periode yang akan datang.

Bentuk VMA dari peubah x pada satu periode ke depan dapat dituliskan sbb.:

$$x_{t+1} = \bar{x} + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t+1-i} \dots \dots \dots (14)$$

*Forecast error* pada satu periode kedepan adalah:

$$E_t x_{t+1} - \bar{x} = \sum_{i=1}^{\infty} \phi_i \varepsilon_{t+1-i} \dots \dots \dots (15)$$

*Forecast* satu periode ke depan dilambangkan dengan  $\Phi_0 \varepsilon_{t+1}$ . *Forecast error* pada periode n ke depan adalah:

$$x_{t+n} - E_t x_{t+n} = \bar{x} + \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i} \dots \dots \dots (16)$$

*Forecast error* pada n periode ke depan untuk peubah s adalah:

$$s_{t+n} - E_t y_{t+n} = \phi_{11}(0) \varepsilon_{st+n} + \phi_{11}(1) \varepsilon_{st+n-1} + \dots + \phi_{11}(n-1) \varepsilon_{yt+1}$$

$$+ \phi_{12}(0)\varepsilon_{y_{t+n}} + \phi_{12}(1)\varepsilon_{y_{t+n-1}} + \dots + \phi_{12}(n-1)\varepsilon_{y_{t+1}} \dots\dots\dots(17)$$

Variance dari forecast error  $s_{t+n}$  periode  $n$  ke depan adalah  $\sigma_s(n)^2$  dimana:

$$\sigma_s(n)^2 = \sigma_s^2[\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2] + \sigma_y^2[\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2] \dots\dots\dots(18)$$

Forecast error variance decomposition adalah proporsi dari  $\sigma_s(n)^2$  terhadap shock  $s$  dan shock  $y$ .

Sehingga forecast error variance decomposition pada shock  $s$  adalah:

$$\sigma_s^2[\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2] / \sigma_s(n)^2 \dots\dots\dots(19)$$

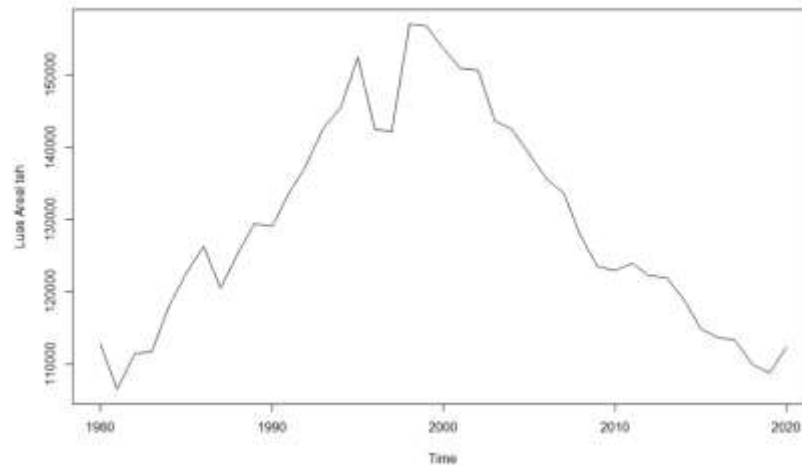
sedangkan forecast error variance decomposition pada shock  $y$  adalah:

$$\sigma_y^2[\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2] / \sigma_s(n)^2 \dots\dots\dots(20)$$

## HASIL DAN PEMBAHASAN

### A. Model ARIMA

ARIMA sering disebut juga dengan metode Box-Jenkins, maka baik digunakan untuk peramalan jangka pendek sedangkan untuk peramalan jangka panjang untuk ketepatan peramalannya masih kurang baik, biasanya akan cenderung *flat* (mendatar/konstan). Eksplorasi data luas areal teh Indonesia dimulai dari tahun 1980 hingga tahun 2020 dalam satuan hektar. Gambar 1 menunjukkan perkembangan luas areal teh tahun 1980 hingga 2020, dimana perkembangan luas areal teh di Indonesia berfluktuasi namun cenderung sedikit meningkat dengan rata-rata pertumbuhan sebesar 0,04% per tahun. Pada tahun 1980 total luas areal teh di Indonesia sebesar 112,70 ribu ton dan mengalami sedikit penurunan di tahun 2020 menjadi 112,05 ribu ton.



Gambar 1. Perkembangan Luas Areal Teh Tahun 1980-2020

Pada pemodelan *Autoregressive Integrated Average* (ARIMA), data yang digunakan dapat dilakukan pemisahan antara data testing dan data training, pemisahan data ini berfungsi untuk menguji tingkat akurasi dalam melakukan peramalan. Panjang series data luas areal teh pada data set training adalah tahun 1980 sampai tahun 2014, sementara dataset testing adalah periode tahun 2015 sampai tahun 2020. Dataset training digunakan untuk melakukan penyusunan model, sementara dataset testing digunakan untuk validasi model.

Uji kestasioneran data pada pemodelan ARIMA dapat dilihat dari hasil plot data dan uji formal statistik. Dari hasil uji formal statistik dapat dilihat dengan uji *Augmented Dickey-Fuller*, yang

menunjukkan bahwa data luas areal teh belum stasioner. Hal ini terlihat dari hasil *value of test statistic* -1.6817, nilai kritis pada tingkat kepercayaan 5% sebesar -3,50 artinya nilai ini lebih kecil dari nilai uji statistik maka  $H_0$  tidak ditolak. Untuk mendapat data produksi teh yang stasioner maka dilakukan differencing satu kali, seperti gambar 2.

Tabel 1. Hasil Uji Augmented Dickey Fuller

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

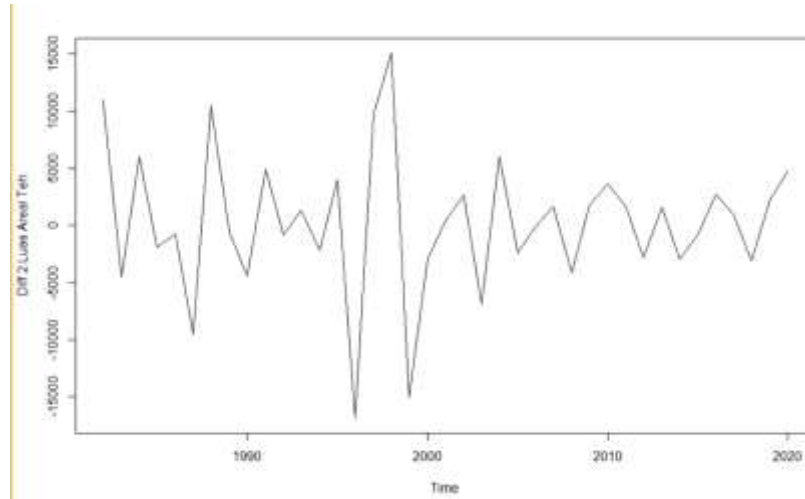
Residuals:
  Min   1Q Median   3Q   Max
-9111 -1855  226  1569 15079

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  1.438e+04  6.468e+03  2.224 0.03271 *
z.lag.1      -8.018e-02  4.768e-02 -1.682 0.10154
tt           -1.804e+02  6.505e+01 -2.773 0.00884 **
z.diff.lag   -1.641e-02  1.595e-01 -0.103 0.91860
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4231 on 35 degrees of freedom
Multiple R-squared:  0.2313, Adjusted R-squared:  0.1654
F-statistic: 3.511 on 3 and 35 DF, p-value: 0.02515

Value of test-statistic is: -1.6817 3.3709 5.0292

Critical values for test statistics:
  1pct  5pct 10pct
tau3  -4.15 -3.50 -3.18
phi2   7.02  5.13  4.31
phi3   9.31  6.73  5.61
```



Gambar 2. Perkembangan Luas Areal Teh Tahun 1980-2020, Setelah Differencing Dua Kali

Dari hasil plot luas areal teh yang telah dilakukan differencing dua kali, maka plot tersebut sudah menunjukkan stasioner. Selain itu didukung juga dari hasil uji *augmented Dickey Fuller* yang menunjukkan Value of test-statistic is: -8.3617 dengan nilai kritis pada tingkat kepercayaan 95% sebesar -1,95 dan tingkat kepercayaan 99% sebesar -2,62 artinya lebih besar dari nilai uji statistik sehingga data produksi teh setelah differencing dua kali sudah stasioner. Stasioneritas data tersebut juga didukung dengan plot ACF dan PACF, dimana pada plot tersebut tidak banyak lag yang keluar dari *confidence interval* (Gambar 3)

Tabel 2. Hasil Uji Augmented Dickey Fuller Setelah Differencing Dua Kali

```
#####
Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min     1Q   Median     3Q    Max
-15508.6 -1767.0  692.8  2176.1 12871.4

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1   -2.0662    0.2471  -8.362 7.34e-10 ***
z.diff.lag  0.4723    0.1438   3.284 0.00233 **
---

```

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5001 on 35 degrees of freedom

Multiple R-squared: 0.7662,

Adjusted R-squared: 0.7529

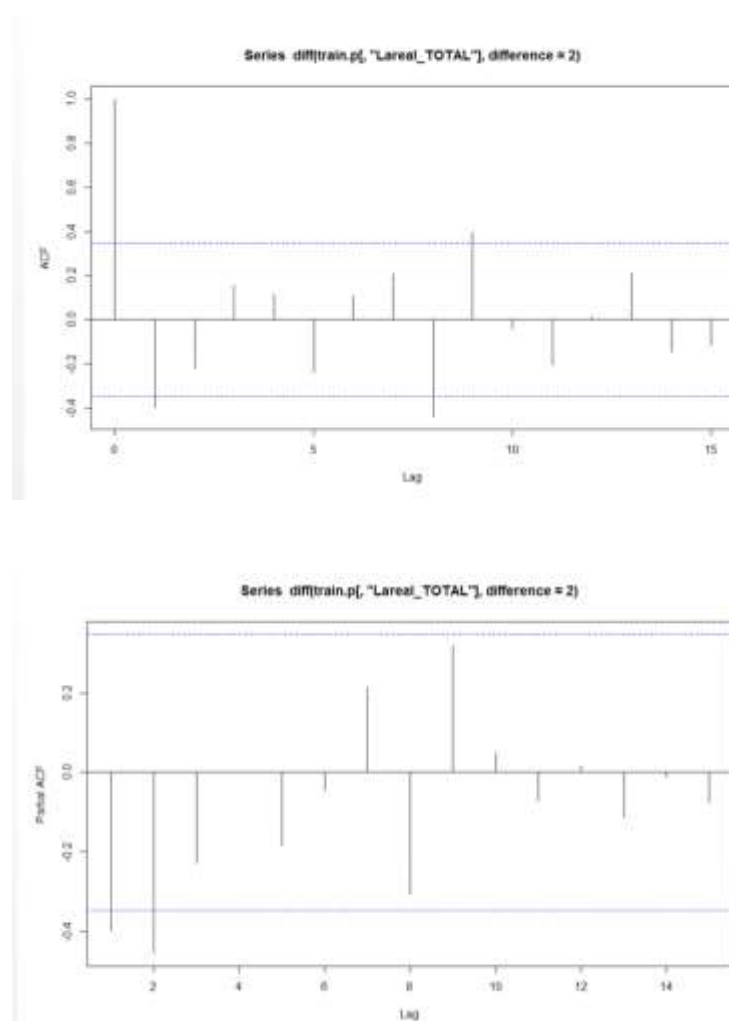
F-statistic: 57.36 on 2 and 35 DF, p-value: 9.001e-12

Value of test-statistic is: -8.3617

Critical values for test statistics:

1pct 5pct 10pct

tau1 -2.62 -1.95 -1.61



Gambar 3. Plot ACF dan PACF Produksi Teh Differencing Dua Kali

Setelah dilakukan penelusuran dengan menggunakan dari hasil plot ACF dan PACF, maka sulit untuk ditentukan orde ARIMA. Untuk itu dilakukan run model dengan menggunakan auto arima. Dari hasil auto arima disarankan untuk menggunakan ARIMA (0,2,1). Dengan menggunakan ARIMA (0,2,1) menghasilkan MAPE yang cukup kecil yaitu 2,63% artinya model arima akan menyimpang sekitar -2,63% sampai +2,63% dari data aktual.

Tabel 3. ARIMA dengan Model Auto Arima

Series: train.p[, "Lareal_TOTAL"]							
ARIMA(0,2,1)							
Coefficients:							
ma1							
-0.8565							
s.e. 0.0969							
sigma^2 estimated as 25963394: log likelihood=-318.71							
AIC=641.43 AICc=641.84 BIC=644.36							
Training set error measures:							
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-189.8523	4865.44	3550.839	-0.0691502	2.631734	0.927692	-0.00731326

Hasil model Auto Arima (0,2,1) dapat menghasilkan nilai estimasi, maka untuk mendapatkan orde ARIMA yang lebih baik lagi dapat menggunakan *Arima selection*. Pada metode ini akan dikeluarkan beberapa model terbaik. Model terbaik pertama menurut metode *Arima selection* adalah ARIMA (5,1,5) menghasilkan nilai sbc paling kecil sebesar 396.3627. Dari penelusuran model dengan menggunakan *Arima selection* belum diperoleh hasil yang terbaik, maka diputuskan untuk menggunakan model dari hasil auto arima.

Tabel 4. Arima Selection

	p	q	sbc
[1,]	5	5	396.3627
[2,]	3	5	447.9011
[3,]	4	5	451.0534
[4,]	4	3	452.0562
[5,]	5	3	455.1378
[6,]	4	4	456.8146
[7,]	5	4	460.2140
[8,]	3	4	460.8208
[9,]	3	3	463.9945
[10,]	2	4	485.8555



Model ARIMA terbaik telah di dapat maka dilakukan pengujian berdasarkan data training dan data testing. Model terbaik adalah model yang menghasilkan MAPE terkecil. Pada analisis ini sudah dilakukan pembagian data training dan data testing. Data training merupakan data produksi teh dari tahun 1980 – 2020, sementara data testing diambil 6 data terakhir, yaitu produksi teh tahun 2015 – 2020. Untuk menguji performa model ARIMA terbaik, dilakukan pengujian dengan data testing. Hasil pengujian data training dan testing terlihat pada tabel 5.

Selanjutnya dilakukan pengujian model ARIMA (0,2,1) apakah koefisien sudah signifikan. Untuk model ARIMA (0,2,1) koefisien  $ma_1$  sebesar 0,02 dan koefisien ini signifikan pada taraf 1%. Sehingga model ARIMA (0,2,1) layak digunakan.

Tabel 5. Uji Koefisien Model ARIMA (0,2,1)

```
Call:
arima(x = train.p[, "Lareal_TOTAL"], order = c(0, 2, 1))

Coefficients:
    ma1
-0.8565
s.e.  0.0969

sigma^2 estimated as 25151108: log likelihood = -318.71, aic = 641.43
> library(lmtest)
> coeftest(model1)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ma1 -0.856493   0.096906 -8.8384 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan Ljung-Box. Hasil pengujian nilai p-value lag 5 sampai dengan lag 30 tidak ada yang signifikan, sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (0,2,1) layak digunakan.

Tabel 6. Uji Sisaan Ljung Box ARIMA (0,2,1)

lags	statistic	df	p-value
5	3.287517	5	0.6557534
10	11.234632	10	0.3395327
15	17.271868	15	0.3028736
20	21.515550	20	0.3673657
25	23.806620	25	0.5305687
30	24.788329	30	0.7352520

Tahap selanjutnya adalah pengepasan model untuk seluruh data. Pengepasan data luas areal dari tahun 1980 hingga 2020 dengan model ARIMA (0,2,1) akan menghasilkan koefisien  $\alpha$  sebesar -0.8629, dengan MAPE 2,49%. Dari hasil MAPE tersebut maka akan diartikan bahwa antara data estimasi dengan data aktual akan terdapat perbedaan dengan rata-rata berkisar antara -2,49% sampai +2,49%.

Tabel 6. Pengepasan Model ARIMA (0,2,1) Untuk Seluruh Data

Series: teh[, "Lareal_TOTAL"]							
ARIMA(0,2,1)							
Coefficients:							
$\alpha_1$							
-0.8629							
s.e. 0.0873							
sigma <sup>2</sup> estimated as 22304349: log likelihood=-385.46							
AIC=774.92 AICc=775.25 BIC=778.25							
Training set error measures:							
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-119.7972	4546.683	3298.813	-0.01909835	2.493562	0.9217586	-0.000206116

Hasil estimasi luas areal teh Indonesia dengan menggunakan model ARIMA (0,2,1) menghasilkan angka estimasi untuk 5 tahun ke depan untuk tahun 2021-2025. Hasil estimasi luas areal teh tahun 2021 sebesar 111.257 ha. Tahun 2022 luas areal teh turun menjadi 110.208 ha kemudian mengalami penurunan luas areal di tahun 2023 menjadi 109.158 ha. Kembali terjadi penurunan

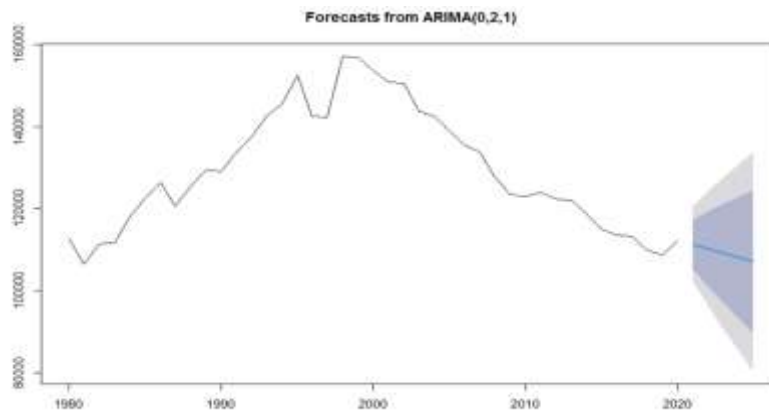
luas areal di tahun 2024 dan tahun 2025 masing-masing sebesar 108.108 ha dan 107.059 ha. Selama tahun 2021 hingga tahun 2025 rata-rata pertumbuhan turun sebesar 0,96%.

Tabel 7A. Output Estimasi Luas Areal Teh ARIMA (0,2,1)

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	111257.3	105204.88	117309.8	102000.91	120513.8
2022	110207.7	101042.72	119372.6	96191.10	124224.2
2023	109158.0	97180.16	121135.8	90839.48	127476.5
2024	108108.3	93396.89	122819.7	85609.14	130607.5
2025	107058.7	89616.01	124501.3	80382.44	133734.9

Tabel 7B. Hasil Estimasi Luas Areal Teh ARIMA

Tahun	Est. Luas Areal Ha	Pertumbuhan %
2021	111.257	
2022	110.208	-0,94
2023	109.158	-0,95
2024	108.108	-0,96
2025	107.059	-0,97
<b>Rata-rata Pertumbuhan (%)</b>		<b>-0,96</b>



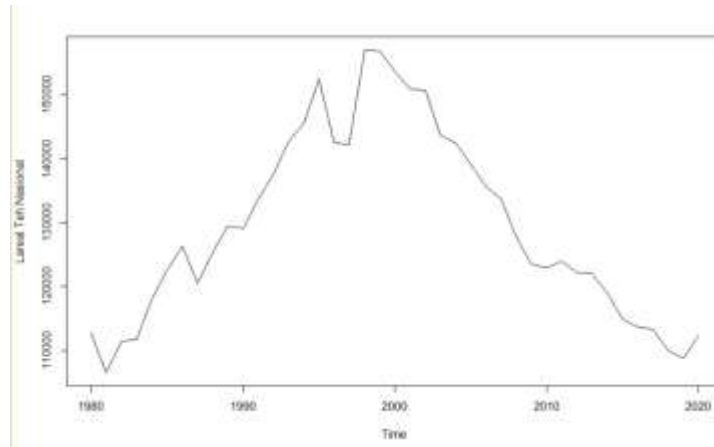
Gambar 4. Hasil Estimasi Luas Areal Teh Tahun 2021-2025

## B. Model Fungsi Transfer

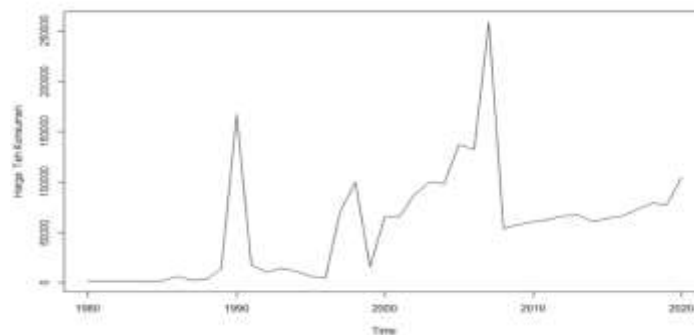
Dalam melakukan estimasi luas areal teh dengan menggunakan model fungsi transfer, peubah inputnya adalah harga konsumen teh. Langkah awal dalam mencari model fungsi transfer adalah penyiapan data, yang berupa eksplorasi data dengan menampilkan plot data produksi teh maupun luas tanaman menghasilkan teh. Data yang digunakan bersifat data series yang dimulai dari tahun 1980 hingga tahun 2020. Data luas areal teh berfluktuasi namun cenderung menurun, sementara data harga konsumen teh berfluktuasi namun cenderung meningkat.

Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (*Autoregressive Integrated Moving Average*). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variabel lain sebagaimana pada model regresi. Model fungsi transfer merupakan suatu model yang menggambarkan nilai prediksi dari suatu time series (deret output atau  $Y_t$ ) berdasarkan nilai-nilai dari deret itu sendiri ( $Y_t$ ) dan berdasarkan pula pada data time series yang mempunyai hubungan (deret input atau  $X_t$ ) dengan deret output. Model fungsi transfer yang bersifat dinamis berpengaruh tidak hanya pada hubungan linier antara waktu ke- $t$  input  $X_t$  dan waktu ke- $t$  output  $Y_t$ , tetapi juga saat input  $X_t$  dengan saat  $t, t+1, \dots, t+k$  pada output  $Y_t$  (Fitriani et al., 2013). Dalam melakukan estimasi dengan fungsi transfer diperlukan dua variabel, yaitu output  $Y_t$  dan input  $X_t$ . Input  $X_t$  merupakan variabel yang diduga paling berpengaruh terhadap output  $Y_t$ . Pada analisis ini dicontohkan estimasi untuk luas areal teh (output  $Y_t$ ) yang diduga dipengaruhi oleh harga konsumen teh.

Eksplorasi data dilakukan dengan menampilkan plot data luas areal maupun harga konsumen. Berdasarkan plot data dapat diketahui pola data series 41 titik yang akan digunakan untuk pemodelan yaitu dimulai dari tahun 1980 hingga tahun 2020. Berdasarkan Gambar 5 dan Gambar 6, terlihat bahwa terdapat data luas areal teh nasional memiliki tren menurun dari tahun ke tahun sedangkan harga konsumen teh berfluktuasi namun cenderung meningkat. Luas areal cenderung turun pada sekitar 10 tahun terakhir, akibat persaingan harga dengan produk teh luar negeri. Luas areal teh nasional maupun harga konsumen teh terindikasi tidak stasioner berdasarkan plotnya.



Gambar 5. Luas Areal Teh Tahun 1980-2020



Gambar 6. Harga Teh Konsumen Tahun 1980-2020

Dari hasil uji Augmented Dickey Fuller yang telah dilakukan differencing dua kali menghasilkan nilai *test-statistic* yaitu -5.7828 lebih kecil dari *critical values* artinya data harga teh konsumen sudah stasioner.

Tabel 8. Hasil Uji Augmented Dickey Fuller Harga Teh Konsumen Setelah Differencing

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
  Min   1Q Median   3Q   Max
-138602 -5320  3395 13858 159616
```

Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
z.lag.1	-1.7686	0.3058	-5.783	2.9e-06 ***
z.diff.lag	0.2332	0.1806	1.291	0.207

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 56750 on 29 degrees of freedom

Multiple R-squared: 0.7325, Adjusted R-squared: 0.714

F-statistic: 39.7 on 2 and 29 DF, p-value: 4.976e-09

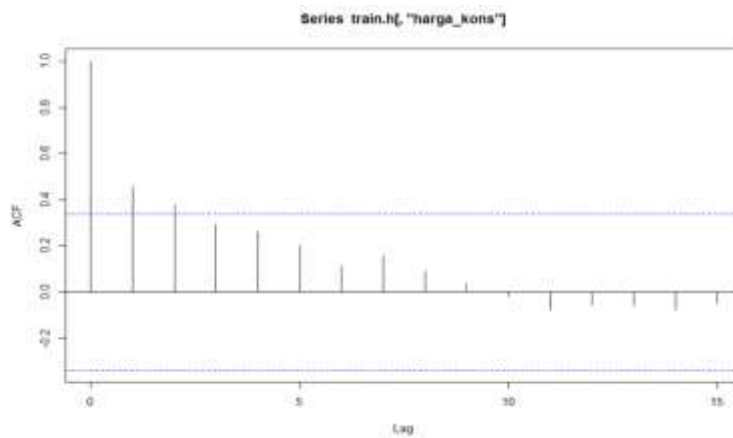
Value of test-statistic is: -5.7828

Critical values for test statistics:

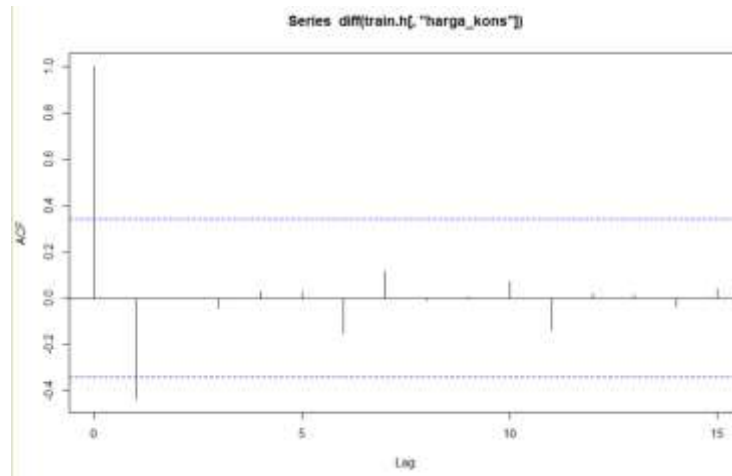
	1pct	5pct	10pct
tau1	-2.62	-1.95	-1.61

Stasioneritas data dapat juga dilihat dari plot ACF dan PACF. Jika lag pada ACF maupun PACF menunjukkan *tail off* maka data tidak stasioner.

Stasioneritas data luas tanaman menghasilkan teh juga didukung dengan plot ACF dan PACF, dimana pada plot tersebut tidak banyak lag yang keluar dari *confidence interval*.



Gambar 7. Plot ACF Harga Konsumen Teh Setelah Differencing



Gambar 8. Plot PACF Harga Konsumen Teh Setelah Differencing

Penelusuran model ARIMA dengan model terbaik dapat menggunakan *script auto arima*, dengan data *training*. Dari hasil auto arima harga konsumen menghasilkan ARIMA (0,1,1) dengan MAPE sebesar 83,71% dengan nilai AIC sebesar 814,18 disajikan pada Tabel 9.

Tabel 9. Model Auto Arima Harga Konsumen Teh

Series: train.h[, "harga_kons"]						
ARIMA(0,1,1)						
Coefficients:						
ma1						
-0.6621						
s.e. 0.1588						
sigma^2 estimated as 2.726e+09: log likelihood=-405.09						
AIC=814.18 AICc=814.58 BIC=817.17						
Training set error measures:						
ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set 6180.476	50656.15	32553.7	-42.43018	83.71481	1.072906	0.01703515

Hasil pengolahan harga konsumen teh pada RStudio yang telah dilakukan differencing dua kali akan menampilkan sepuluh model tentatif dengan model arima selection. Model pertama dengan

nilai  $p=1$  dan  $q=3$  dan model kesepuluh adalah  $p=2$  dan  $q=5$ . Dari kesepuluh model ini dipilih model kesembilan yaitu ARIMA (1,2,1).

Tabel 10. Arima Selection Data Harga Konsumen Teh

	p	q	sbc
[1,]	1	3	712.6363
[2,]	1	4	714.1386
[3,]	2	3	715.6897
[4,]	2	4	716.8344
[5,]	1	5	717.8079
[6,]	4	3	718.0707
[7,]	3	4	718.7175
[8,]	3	3	718.7763
[9,]	1	1	719.1867
[10,]	2	5	720.5936

Hasil pengujian koefisien model ARIMA (1,2,1) menghasilkan nilai AIC sebesar 800,3 ditemukan bahwa  $ar1$  tidak signifikan sedangkan  $ma1$  signifikan pada taraf 1%.

Tabel 11. Pengujian Koefisien Model Arima (1,2,1) Harga Konsumen Teh

```
Call:
arima(x = train.h[, "harga_kons"], order = c(1, 2, 1))
Coefficients:
      ar1      ma1
-0.4058 -1.0000
s.e.  0.1571  0.0872

sigma^2 estimated as 3.081e+09: log likelihood = -397.15, aic = 800.3
> library(lmtest)
> coeftest(modell)

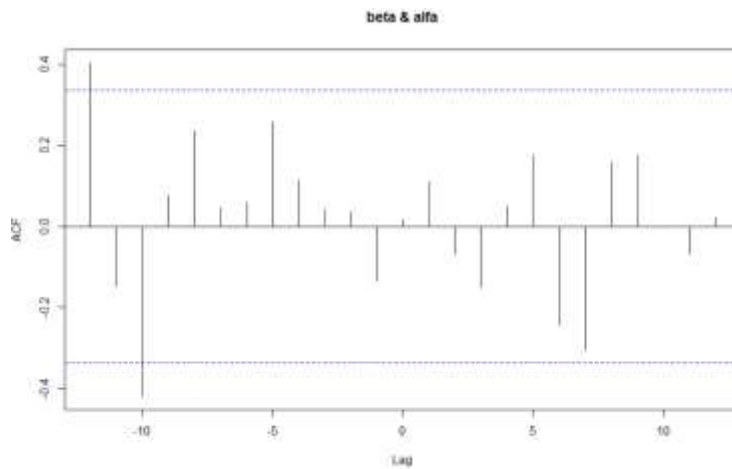
z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1 -0.40580   0.15710 -2.5831 0.009791 **
ma1 -1.00000   0.08719 -11.4692 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tahap selanjutnya dalam penyusunan model fungsi transfer adalah *prewhitening* dan korelasi silang. Korelasi silang menggambarkan struktur hubungan antara  $X_t$  dengan  $Y_t$ . Untuk



mengidentifikasi pengaruh  $X_t$  terhadap  $Y_t$  maka deret  $X_t$  harus stasioner atau sudah distasionerkan. Dalam konteks pemodelan  $X_t$  terhadap  $Y_t$ , untuk membuat  $X_t$  stasioner tidak dengan pembedaan (*differencing*) namun dengan mengambil komponen *white noise* dari  $X_t$  (*prewhitening*). *Prewhitening* dilakukan terhadap deret input  $X_t$  yang didefinisikan sebagai alfa serta deret input  $Y_t$  yang didefinisikan sebagai beta. Hasil output untuk prewhitening dan korelasi silang berupa grafik ACF untuk beta dan alfa.



Gambar 9. Plot korelasi silang Luas Areal Teh dengan Harga Konsumen Teh

Hasil plot korelasi silang digunakan untuk mengidentifikasi ordo  $r, s$ , dan  $b$ . Ordo  $r$  adalah panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$ , ordo  $s$  adalah panjang lag  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$ , dan ordo  $b$  adalah panjang jeda pengaruh  $X_t$  terhadap  $Y_t$ . Identifikasi ordo  $r, s$  dan  $b$  hanya dilihat pada lag yang positif.

Plot korelasi silang diatas menunjukkan bahwa hanya lag 0 yang keluar dari garis signifikansi, maka nilai  $b=0$ . Kemudian, tidak ada tambahan lagi lag yang signifikan maka nilai  $s=0$ . Mengingat data luas areal teh dan harga konsumen teh merupakan data tahunan yang tidak mengandung musiman maka diasumsikan nilai  $r=0$ . Nilai  $b=0$  menunjukkan tidak ada jeda pengaruh antara harga konsumen teh pada waktu  $t$  terhadap luas areal teh pada waktu  $t$ . Nilai  $s=0$  berarti ada korelasi antara luas areal teh dan harga konsumen pada tahun yang sama. Dengan kata lain, dampak dari harga konsumen terhadap luas areal dirasakan pada waktu yang sama ( $t$ ).

Tahap selanjutnya dilakukan pengepasan model, untuk nilai  $r, s$  dan  $b$ . Hasil pengujian fungsi transfer dengan nilai  $r=0, s=0$ , dan  $b=0$  menghasilkan nilai MAPE sebesar 8,77%

Tabel 12. Pengujian Orde r, s, dan b ARIMA (0,0,0) Luas Areal Teh

Regression with ARIMA(0,0,0) errors						
Coefficients:						
intercept	xreg					
129522.927	0.0634					
s.e.	3061.618	0.0393				
sigma^2 estimated as 187863551: log likelihood=-371.08						
AIC=748.17 AICc=748.97 BIC=752.75						
Training set error measures:						
	ME	RMSE	MAE	MPE	MAPE	MASE ACF1
Training set	-3.295659e-11	13297.1	11584.84	-1.008168	8.767603	3.026654 0.8613594

Untuk menghasilkan orde yang paling tepat untuk menentukan orde Arima fungsi transfer dengan melakukan identifikasi model noise. Untuk menghasilkan model terbaik dengan menggunakan *auto-arima* pada R Studio, model maka noise yang disarankan adalah Arima (1,2,1). Model ini ternyata masih kurang tepat, karena menghasilkan MAPE yang cukup besar yaitu 100,00%

Tabel 13. Model Residual Tentatif Luas Areal Teh

Regression with ARIMA(0,0,0) errors						
Coefficients:						
intercept	xreg					
129522.927	0.0634					
s.e.	3061.618	0.0393				
sigma^2 estimated as 187863551: log likelihood=-371.08						
AIC=748.17 AICc=748.97 BIC=752.75						
Training set error measures:						
	ME	RMSE	MAE	MPE	MAPE	MASE ACF1
Training set	-3.295659e-11	13297.1	11584.84	-1.008168	8.767603	3.026654 0.8613594

Tahap berikutnya adalah dilakukan pengepasan model dengan noise. Pengepasan model dengan noise dilakukan dengan menggabungkan nilai r,s,b dan hasil identifikasi model noise yaitu ARIMA(1,2,1). Selanjutnya diperoleh output model fungsi transfer ARIMA (1,2,1) (Tabel 14)

Tabel 14. Model Residual Tentatif ARIMA (1,2,1) Untuk Luas Areal Teh

```
Series: res
ARIMA(1,2,1)
Coefficients:
      ar1    ma1
    -0.1685 -0.8711
s.e.  0.1954  0.1103
sigma^2 estimated as 42238675: log likelihood=-326.18
AIC=658.36  AICc=659.21  BIC=662.75
> coeftest(modelres)
z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1 -0.16849   0.19544 -0.8621  0.3886
ma1 -0.87114   0.11027 -7.8998 2.794e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

### Peramalan Berbasis Fungsi Transfer

Model fungsi transfer dengan model noise ARIMA (1, 2, 1), dilakukan peramalan berbasis nilai aktual dimana produksi teh diestimasi menggunakan data aktual luas tanaman menghasilkan periode 2015-2020. Meskipun data aktual produksi teh periode 2015-2020 telah ada, dilakukan peramalan produksi teh untuk mengecek *performance* model fungsi transfer.

Tabel 15. Pengujian Model Fungsi Transfer ARIMA (1,2,1)

```
Series: train.h[, "Lareal_TOTAL"]
Regression with ARIMA(1,2,1) errors
Coefficients:
      ar1    ma1  xreg
    -0.0626 -0.8398  0.0021
s.e.  0.2114  0.1154  0.0137
sigma^2 estimated as 27647712: log likelihood=-318.66
AIC=645.31  AICc=646.79  BIC=651.17

Training set error measures:
      ME      RMSE     MAE     MPE      MAPE     MASE     ACF1
Training set -179.3917 4856.117 3589.401 -0.05871287 2.660745 0.9377667 0.02689135
```

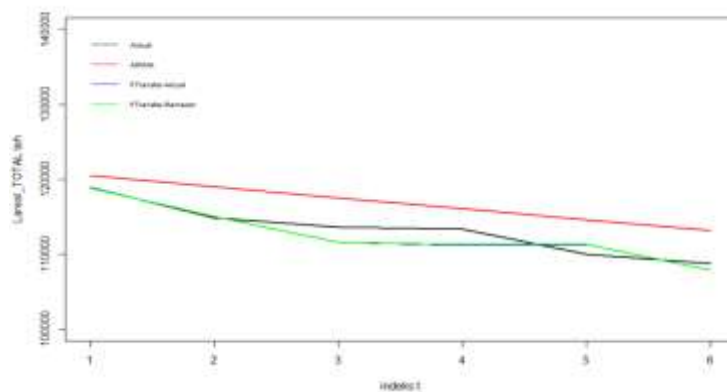
Tujuan melakukan pemodelan fungsi transfer adalah untuk mendapatkan nilai ramalan periode ke depan, yakni produksi teh tahun 2021-2025.

Hasil estimasi produksi teh berbasis fungsi transfer dengan model noise ARIMA (1,2,1) menghasilkan MAPE data training sebesar 2,65% dan MAPE data testing 3,19%.

Tabel 16. Model Fungsi Transfer ARIMA (1,2,1) Luas Areal Teh

ARIMA(1,2,1)							
Coefficients:							
	ar1	ma1					
	-0.0648	-0.8393					
	s.e.	0.2108	0.1150				
	sigma^2 estimated as 26746105: log likelihood=-318.67						
	AIC=643.34 AICc=644.19 BIC=647.73						
Training set error measures:							
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-180.6315	4857.932	3578.238	-0.05977472	2.652660	0.9348501	0.02647468
Test set	-3580.3307	3737.991	3580.331	-3.18711595	3.187116	0.9353969	

Dilakukan perbandingan hasil estimasi terhadap data aktual luas areal teh maupun hasil ramalan. Hasil ramalan tersebut ditampilkan dalam bentuk grafik pada Gambar 10.



Gambar 10. Perbandingan Hasil Ramalan Luas Areal Teh

Dari grafik di atas terlihat jika dibandingkan dengan data aktual luas areal teh 2015-2020 (warna hitam), maka hasil peramalan dengan model ARIMA tanpa fungsi transfer (warna merah) tidak mengikuti pola data aktualnya. Di sisi lain, peramalan dengan fungsi transfer khususnya jika input harga konsumen teh yang digunakan adalah data aktual maka hasil ramalan luas arealnya (warna

biru) sangat menyerupai pola data luas areal aktual selama 6 tahun terakhir. Jika input harga konsumen teh yang digunakan adalah hasil ramalan, maka estimasi luas arealnya (warna hijau) hampir menyerupai pola data asli, meskipun tidak terlalu berimpit seperti estimasi menggunakan data actual.

Berdasarkan nilai MAPE yang dihasilkan maka model fungsi transfer ARIMA dengan input data aktual lebih bagus jika dibandingkan dengan model fungsi transfer ARIMA dengan input data ramalan maupun model ARIMA tanpa fungsi transfer. Namun demikian, karena tujuan penyusunan model adalah untuk mendapatkan angka estimasi tahun 2021 -2025, maka digunakan model Fungsi Transfer berbasis input data ramalan, yakni dengan terlebih dahulu dilakukan estimasi harga konsumen teh lima tahun ke depan menggunakan model ARIMA (1,2,1). Hasil estimasi harga konsumen teh tersebut kemudian menjadi input untuk meramalkan luas areal teh lima tahun ke depan menggunakan model fungsi transfer.

Tabel 17. Hasil Peramalan Model Fungsi Transfer Luas Areal Teh

Time Series: Start = 41 End = 45 Frequency = 1 [1] 106885.76 105045.36 103171.10 101358.43 99524.19
---

Rata-rata pertumbuhan luas areal teh untuk 5 tahun kedepan yaitu periode 2021 hingga 2025 akan mengalami penurunan sebesar 1,77%. Tahun 2021 luas areal teh diestimasi sebesar 106.887 ha. Luas areal teh tahun 2022 sebesar 105.045 ha. Luas areal teh diramalkan terus menurun hingga tahun 2025. Tahun 2023 hasil estimasi luas areal teh sebesar 103.171 ha. Kemudian luas areal teh mengalami penurunan kembali di tahun 2024 dan tahun 2025 masing-masing sebesar 101.358 ha dan 99.524 ha.

Tabel 18. Hasil Estimasi Luas Areal Teh Model Fungsi Trasfer Tahun 2021-2025

Tahun	Luas Areal (Ha)	Pertumbuhan (%)
2021	106.887	
2022	105.045	-1,72
2023	103.171	-1,78
2024	101.358	-1,76
2025	99.524	-1,81
<b>Pertumbuhan (%)</b>		<b>-1,77</b>

### C. Model Vector Auto Regressive (VAR)

Beberapa tahapan yang dilakukan dalam penyusunan model VAR antara lain persiapan data, pembagian data *training* dan *testing*, pemilihan lag dan type, pengajuan asumsi, ramalan data *training*, *testing*, penghitungan mape, dan plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Dibahas juga interpretasi *Impulse Response Function* (IRF) dan *Variance Decomposition*.

Variabel yang digunakan untuk estimasi model VAR antara lain produksi teh dalam satuan ton, luas areal dalam satuan hektare, harga konsumen dalam satuan Rp/Kg, volume ekspor teh dalam satuan ton dan volume impor teh dalam satuan ton. Semua data ini bersumber dari Ditjen Perkebunan kecuali data harga konsumen bersumber dari Badan Pusat Statistik, selain itu series data yang digunakan dimulai dari tahun 1980 hingga 2020.

Komposisi variabel tersebut, komponen konstanta dan trend juga perlu dilakukan pengujian untuk mengetahui signifikan atau tidak keberadaannya. Jika keduanya signifikan, maka komponen tersebut harus dimasukkan ke dalam model VAR dengan type “both”. Jika hanya konstanta yang signifikan, maka trend perlu dikeluarkan dari model VAR dengan model VAR type “const”. Jika hanya trend yang signifikan maka konstanta dikeluarkan dari model menggunakan model VAR type “trend”, dan jika keduanya tidak signifikan, maka type yang digunakan model VAR adalah “none”.

Keberadaan konstanta dan trend dapat dideteksi dari plot data awal, namun terkadang hal tersebut sulit dilakukan. Untuk mengantisipasi hal tersebut, sebaiknya dilakukan uji coba/*trial and error* menggunakan model VAR dari lag  $p=1$  s.d 6 dengan type “both” untuk mengetahui signifikan atau tidak keberadaannya. Untuk panjang lag maksimum bisa dilakukan trial error sampai tidak memungkinkan untuk dilakukan permodelan.

Pada data luas areal teh, setelah dilakukan running model VAR dengan lag  $p=1$  type “both”, diperoleh informasi komponen konstanta dan trend signifikan, sehingga pada pengujian berikutnya dilakukan running model VAR dengan menggunakan lag  $(p) = 1$  s.d lag  $(p)=4$  dengan menyertakan komponen konstanta dan trend. Pemilihan lag  $p$  ditentukan dengan melihat banyaknya variabel yang signifikan dalam lag tersebut sekaligus memastikan harus ada variabel/peubah yang signifikan pada lag terpilih dimaksud. Setelah dilakukan run model ternyata ada 2 kandidat model VAR terbaik yaitu Var (1) type “both” dan Var (2) type “both”. Type “both” artinya persamaan Var itu mengandung constanta dan faktor trend.

Untuk model VAR yang pertama adalah Model VAR (p=1) dengan konstanta dan trend (type=both). Hasil model Var(1) type both, untuk mengestimasi luas areal (t) teh menunjukkan bahwa koefisien constanta dan trend menunjukkan hasil yang signifikan pada tingkat kepercayaan 95%. Untuk variabel lain yang signifikan untuk mengestimasi luas areal (t) antara lain Produksi, luas areal, dan harga konsumen pada lag-1. Ketiga variabel tersebut signifikan dengan tingkat kepercayaan 95%.

Model Var (1) type both ini menghasilkan nilai Adjusted R Square = 82,62%, artinya keragaman produksi dipengaruhi oleh variabel penjelasnya sebesar 82,62%. Nilai p-value untuk model produksi ini sangat kecil atau jauh lebih kecil dari 0,05, sehingga model layak untuk digunakan. Jumlah variabel yang signifikan ada 3 variabel dari total 6 variabel, sehingga model ini layak digunakan dan menjadi kandidat terbaik untuk model VAR.

Tabel 19. Output Model VAR(1) Type “Both” untuk Luas Areal Teh

Estimation results for equation Lareal_TOTAL:				
=====				
Lareal_TOTAL = Produksi_TOTAL.l1 + Lareal_TOTAL.l1 + harga_kons.l1 + vol_ekspor.l1 + vol_impopor.l1 + const + trend				
	Estimate	Std. Error	t value	Pr(> t )
Produksi_TOTAL.l1	1.781e-01	1.355e-01	1.315	0.200
Lareal_TOTAL.l1	8.388e-01	1.372e-01	6.115	1.83e-06 ***
harga_kons.l1	-9.452e-03	1.869e-02	-0.506	0.617
vol_ekspor.l1	-4.978e-02	7.571e-02	-0.657	0.517
vol_impopor.l1	7.380e-02	3.189e-01	0.231	0.819
const	6.580e+03	1.121e+04	0.587	0.562
trend	-3.398e+02	2.351e+02	-1.445	0.160
---				
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				
Residual standard error: 4988 on 26 degrees of freedom				
Multiple R-Squared: 0.8932, Adjusted R-squared: 0.8685				
F-statistic: 36.22 on 6 and 26 DF, p-value: 2.015e-11				

Tabel 20. Output Model VAR(2) Type “Both”

Estimation results for equation Lareal_TOTAL:				
=====				
Lareal_TOTAL = Produksi_TOTAL.11 + Lareal_TOTAL.11 + harga_kons.11 + vol_ekspor.11 + vol_impор.11 + Produksi_TOTAL.12 + Lareal_TOTAL.12 + harga_kons.12 + vol_ekspor.12 + vol_impор.12 + const + trend				
	Estimate	Std. Error	t value	Pr(> t )
Produksi_TOTAL.11	1.730e-01	1.313e-01	1.318	0.20248
Lareal_TOTAL.11	6.836e-01	1.807e-01	3.783	0.00117 **
harga_kons.11	-3.225e-03	2.684e-02	-0.120	0.90554
vol_ekspor.11	-1.287e-01	8.368e-02	-1.538	0.13973
vol_impор.11	-1.558e-01	6.811e-01	-0.229	0.82136
Produksi_TOTAL.12	2.151e-01	1.257e-01	1.711	0.10258
Lareal_TOTAL.12	-3.256e-02	1.941e-01	-0.168	0.86847
harga_kons.12	-1.631e-02	3.359e-02	-0.486	0.63255
vol_ekspor.12	7.755e-02	8.203e-02	0.945	0.35572
vol_impор.12	3.613e-01	8.774e-01	0.412	0.68486
const	7.243e+03	1.074e+04	0.675	0.50763
trend	-6.413e+02	2.929e+02	-2.190	0.04056 *
---				
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1				
Residual standard error: 4237 on 20 degrees of freedom				
Multiple R-Squared: 0.9324, Adjusted R-squared: 0.8952				
F-statistic: 25.07 on 11 and 20 DF, p-value: 2.65e-09				

Hasil model Var(2) type both, untuk mengestimasi Produksi (t) teh menunjukkan bahwa koefisien constanta dan trend menunjukkan hasil yang signifikan pada tingkat kepercayaan 99%. Untuk variabel lain yang signifikan untuk mengestimasi produksi (t) antara lain Produksi (t-1), luas areal (t-1), harga konsumen (t-1), dan volume ekspor (t-1), serta volume impor teh (t-1) dengan tingkat kepercayaan 90% - 95%. Jumlah variabel sebanyak 5 variabel. Model VAR(2) type=both ini juga layak perlu diuji MAPE untuk melihat apakah lebih baik dari model VAR(1) type=both.

Selanjutnya dilakukan pengujian model VAR, dengan nilai p=3, p=4 type both, dan model VAR p=3, p=4 type konstan. Hasil pengujian untuk Var(3) type ‘both’ menghasilkan koefisien constanta dan trend signifikan. Jika dibandingkan dengan p=2, untuk p=3 penambahan variabel yang signifikan relatif kecil. Selanjutnya dilakukan pengujian untuk nilai p=4, ternyata constanta dan trend sudah tidak signifikan, sehingga berhenti pada p=4. Selanjutnya dilakukan pengujian untuk nilai p=3 dan p=4 tetapi type ‘constant’, menunjukkan hasil koefisien constanta tidak signifikan, sehingga berhenti untuk dilakukan pengujian lebih lanjut. Berdasarkan pengujian ini yang yang terbaik adalah untuk nilai p=1 atau p=2 dengan type ‘both’.



Pemeriksaan autokorelasi residual model menggunakan fungsi “serial.test” yang di dalamnya dilakukan pengujian *Portmanteau-and Breusch-Godfrey test*. Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi. Pengujian Jarque-Bera tests untuk menguji kenormalan, hasil pengujian menunjukkan Nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal. Pemeriksaan heteroskedastisitas model menggunakan fungsi “arch.test” yang di dalamnya dilakukan pengujian *ARCH-LM tests*. Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.

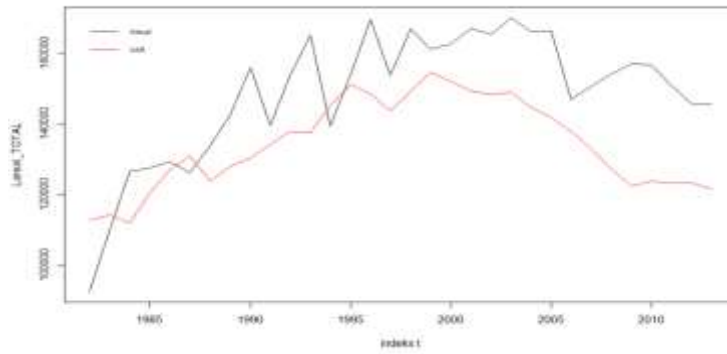
Tabel 21, Ouput Pengujian Asumsi VAR(2) type=both

Portmanteau Test (asymptotic) data: Residuals of VAR object varhsheet.b2 Chi-squared = 285.91, df = 350, p-value = 0.9948 \$JB data: Residuals of VAR object varhsheet.b2 Chi-squared = 31.417, df = 10, p-value = 0.0005006 \$Skewness data: Residuals of VAR object varhsheet.b2 Chi-squared = 18.299, df = 5, p-value = 0.002594 \$Kurtosis data: Residuals of VAR object varhsheet.b2 Chi-squared = 13.118, df = 5, p-value = 0.0223 ARCH (multivariate) data: Residuals of VAR object varhsheet.b2 Chi-squared = 405, df = 1125, p-value = 1
---

Tabel 22. Nilai MAPE untuk Model VAR (2) type=both

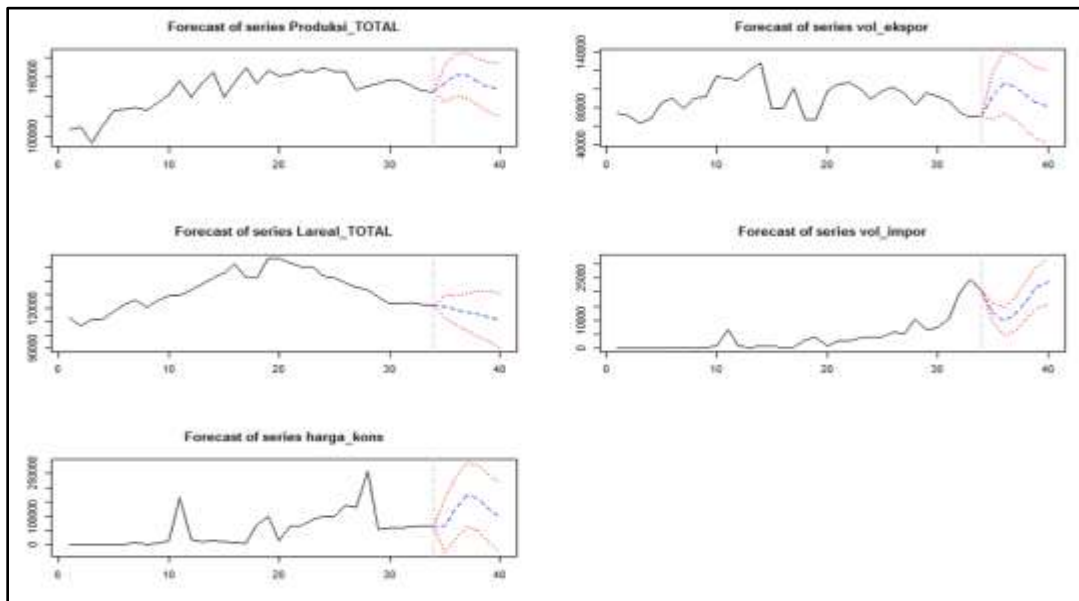
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
1.712	6.007	10.551	10.884	15.634	21.974

Dari grafik di bawah dapat dilihat bahwa pergerakan ramalan pada data testing tidak mengikuti pergerakan data asli/aktual. Sehingga mungkin model VAR (1) type “both” bukan model yang terbaik. Hasil peramalan dengan model VAR (2) ini cenderung lebih tinggi dari data aktual, kecuali di tahun 1987 dan tahun 1994 data aktual lebih tinggi dari data estimasi dengan VAR (2). Untuk data testing ini rata-rata penyimpangannya adalah sebesar 10,88%.



Gambar 11. Plot Ramalan dan Aktual VAR (2) Type “Both”

Keunggulan model VAR ini pada saat dilakukan permalan maka semua variabel dalam sistem akan ikut diramalkan. Hasil peramalan bukan hanya target variabel, tetapi semua variabel yang masuk dalam sistem. Hasil permalan 5 tahun ke depan untuk variabel produksi menunjukkan produksi akan menurun, begitu juga luas areal teh juga cenderung turun. Sebaliknya harga konsumen cenderung meningkat. Volume impor teh juga diperkirakan akan meningkat seiring dengan peningkatan produksi, sementara itu untuk volume impor teh diperkirakan akan tetap stagnan.



Gambar 12. Hasil Estimasi Model VAR (2) Type “Both”

Hasil estimasi luas areal teh pada model VAR (2) Type “Both” cenderung menurun. Rata-rata pertumbuhan untuk 5 tahun kedepan yaitu periode 2021 hingga 2025 akan turun sebesar 1,17%. Tahun 2021 luas areal teh diestimasi sebesar 120.421 ha. Luas Areal teh tahun 2022 sebesar 119.068 ha. Luas areal teh diramalkan terus menurun hingga tahun 2025. Tahun 2023 hasil estimasi produksi teh sebesar 117.701 ha. Kemudian produksi teh mengalami penurunan di tahun 2024 dan tahun 2025 masing-masing sebesar 116.312 ha dan 114.899 ha.

Tabel 23. Hasil Estimasi Produksi Teh VAR (2), Type “Both”

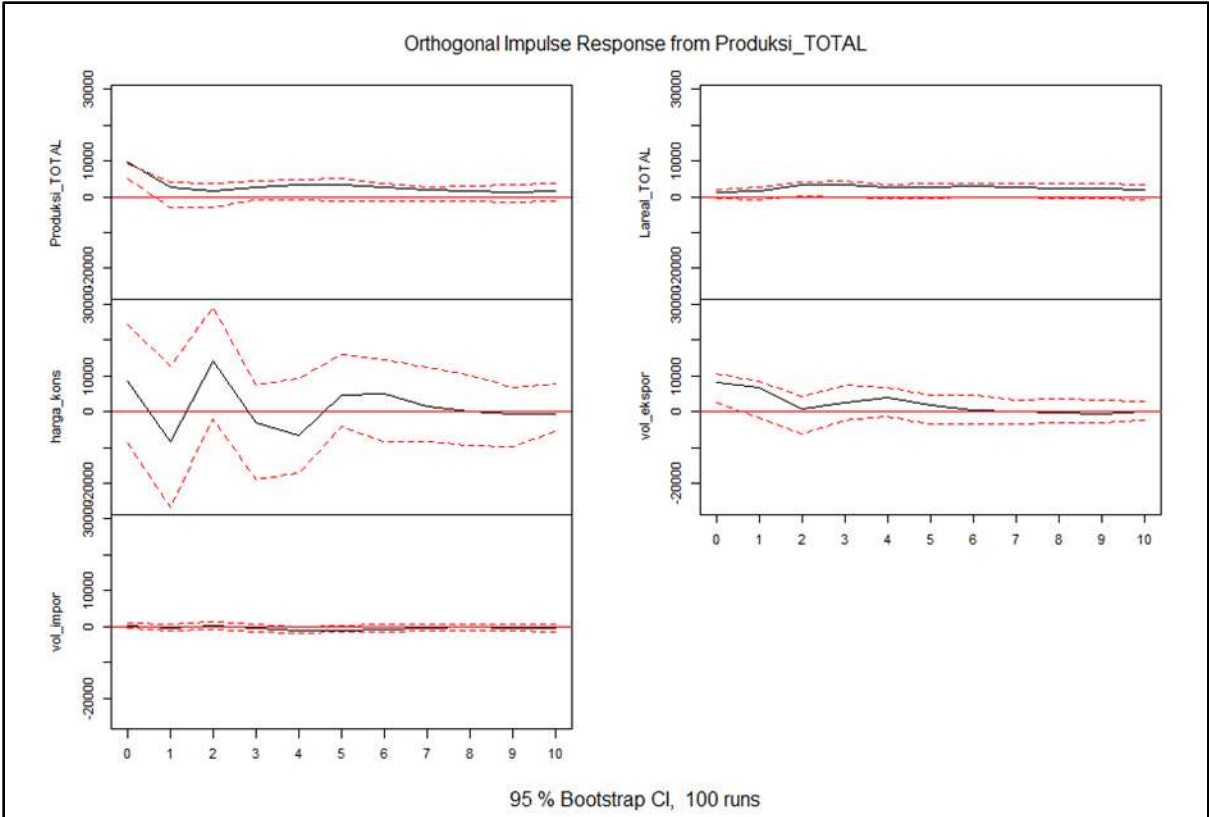
Tahun	Luas Areal (Ha)	Pertumbuhan (%)
2021	120.421	
2022	119.068	-1,12
2023	117.701	-1,15
2024	116.312	-1,18
2025	114.899	-1,21
<b>Pertumbuhan (%)</b>		<b>-1,17</b>

### Interpretasi Impulse Response Function (IRF) dan Variance Decomposition

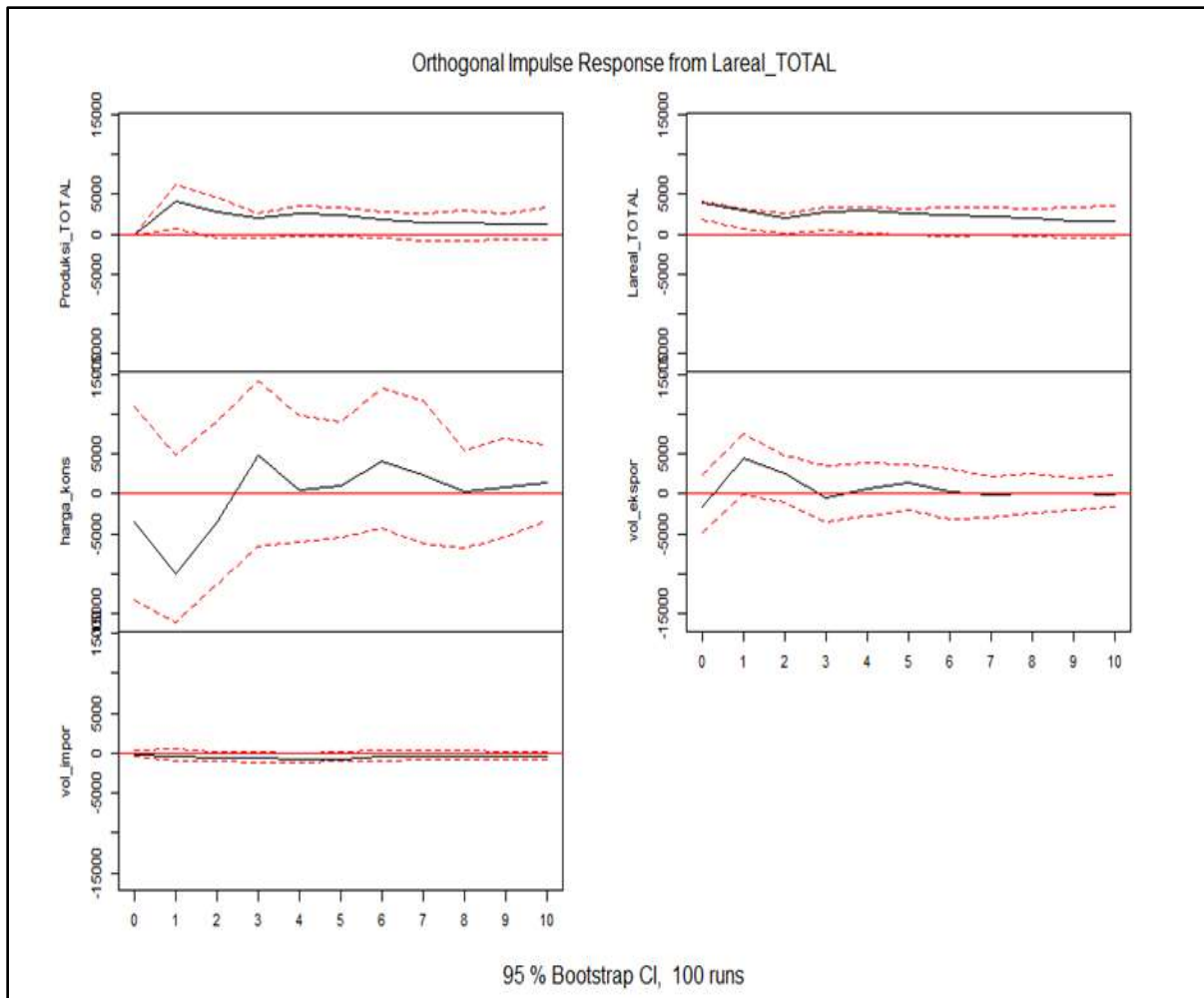
Setelah diperoleh model terbaik, untuk permodelan VAR dapat dilakukan interpretasi lebih mendalam terkait Impulse Response Function (IRF) dan Variance Decomposition. Berikut interpretasi ketiga hal di atas untuk model terbaik VAR (2) type “both”.

**Impulse Response Function** akan menjelaskan bagaimana perubahan atau gejolak yang terjadi pada suatu variabel/peubah di tahun tertentu akan berdampak pada variabel/peubah lain di tahun tertentu dan tahun-tahun setelahnya.

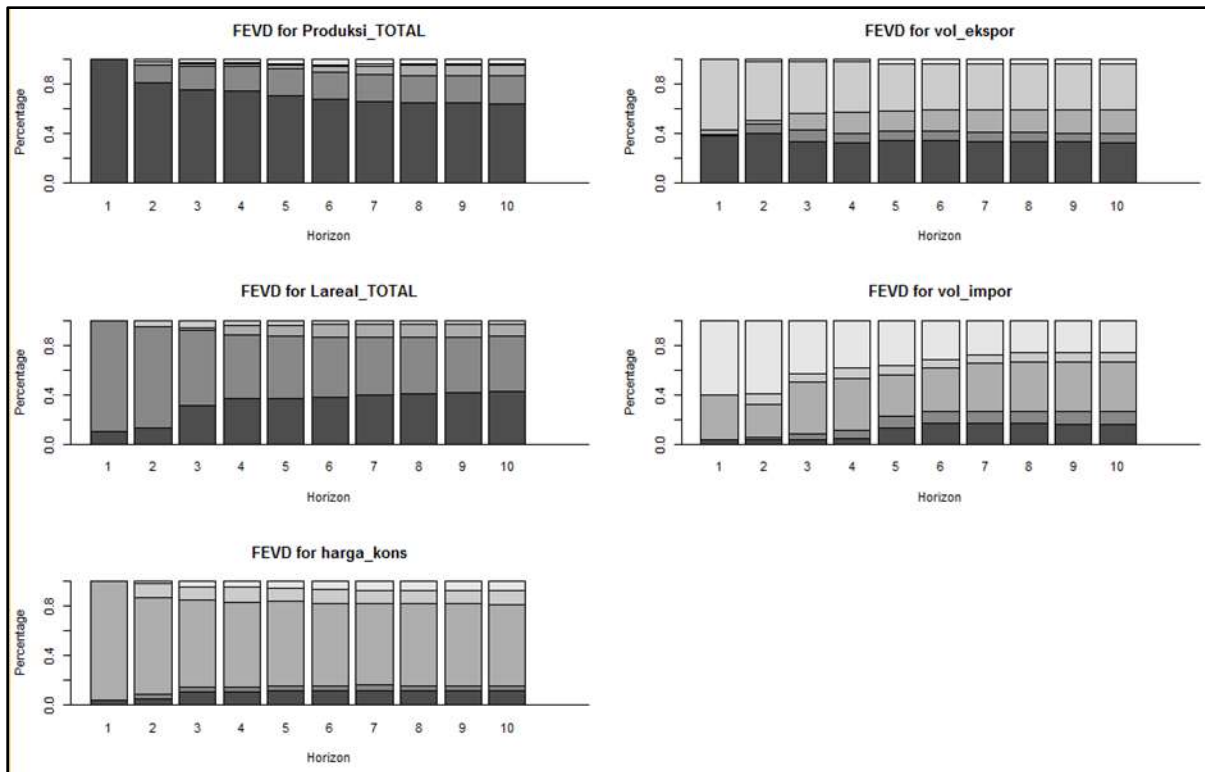
Dari grafik Impulse Response Function produksi model terbaik VAR (2) type “both”, dapat dilihat bahwa jika terjadi perubahan pada produksi di tahun tertentu maka akan berdampak pada produksi itu sendiri sampai 3 tahun ke depan. Dampak tersebut akan hilang setelah 3 tahun. Sedangkan dampak perubahan produksi tidak berdampak pada harga konsumen teh, luas areal, produksi teh, dan volume impor teh, namun masih berdampak pada volume ekspor kedepan. Perubahan produksi berdampak pada volume ekspor sampai dengan dua tahun dan akan hilang dampaknya setelah melewati tahun kedua tersebut.



Gambar 13. Impulse Respon Produksi Teh Model VAR (2) Type “Both”



Gambar 14. Impulse Respon Luas Areal Teh Model VAR (2) Type “Both”



Gambar 14. Dekomposisi Model VAR (2) Type “Both”

#### D. Pemilihan Model Terbaik

Salah satu dasar penentuan model terbaik dapat dilakukan dengan membandingkan nilai MAPE untuk data testing dan training yaitu dengan memilih nilai MAPE yang paling kecil, terutama untuk data testing. Selain MAPE yang terkecil, pola pergerakan ramalan juga harus diperhatikan. Pilihlah plot yang paling berhimpit/bersesuaian dengan data asli/aktual atau dengan kata lain performa hasil ramalan seiring dengan data historisnya.

Penyusunan angka estimasi produksi teh dilakukan dengan cara uji coba model statistik antara lain model ARIMA, Fungsi Transfer dan VAR.

Model pertama adalah ARIMA (0,2,1) menghasilkan MAPE training sebesar 2,63% dan MAPE testing sebesar 3,47%. Hal ini menunjukkan bahwa kemampuan untuk melakukan estimasi dengan model ARIMA ini rata-rata akan mengalami kesalahan sekitar 3,47% lebih tinggi atau 3,47% lebih rendah. Sementara model ARIMA kedua adalah ARIMA (2,2,4) menghasilkan MAPE training sebesar 2,31% dan MAPE testing sebesar 12,70%. Hal ini menunjukkan bahwa kemampuan untuk melakukan estimasi dengan model ARIMA ini rata-rata akan mengalami kesalahan sekitar 12,70%

lebih tinggi atau 12,70% lebih rendah. Model ARIMA ini menghasilkan MAPE yang lebih kecil dibandingkan model Fungsi Transfer dan VAR, sehingga model ARIMA lebih akurat dalam melakukan estimasi.

Metode estimasi yang kedua adalah dengan model fungsi transfer, untuk melakukan estimasi luas areal teh dengan variabel bebas adalah harga konsumen teh. Untuk model fungsi transfer ini menghasilkan MAPE data training 2,65%, sementara untuk MAPE data testing sebesar 3,19%. Hasil estimasi juga menunjukkan angka yang lebih realistis, dengan angka estimasi tahun 2021 sebesar 106.887 ha. Namun untuk estimasi 5 tahun kedepan (tahun 2021-2025) angka pertumbuhannya mengalami penurunan sebesar 1,77%, sementara angka pertumbuhan 5 tahun sebelumnya (tahun 2016-2020) turun sebesar 0,26%.

Untuk model estimasi yang terakhir adalah dengan model VAR (Vector Auto Regressive). Untuk model VAR ini menggunakan 5 variabel yaitu produksi, luas areal, harga konsumen, volume ekspor dan volume impor teh. Model yang terbaik untuk Model VAR adalah nilai  $p=2$  dan  $type="both"$ ,  $p=2$  artinya menggunakan variabel bebas sampai lag-2 dan ada konstanta dan factor trend. Estimasi luas areal teh dengan menggunakan model VAR ini menghasilkan ketelitian yang cukup tinggi yaitu MAPE untuk data training 10,88% dan MAPE untuk data testing 16,44%.

Jika dibandingkan dengan ketiga metode tersebut, maka metode ARIMA menghasilkan MAPE data testing paling kecil dibandingkan dengan model Fungsi Transfer dan VAR. Angka pertumbuhan luas areal teh hasil estimasi 5 tahun kedepan dengan rerata pertumbuhan turun 0,96%, sementara angka pertumbuhan 5 tahun terakhir sebelumnya turun sebesar 0,26%.

Tabel 24. Perbandingan MAPE Luas Areal Teh Indonesia Model Arima, Fungsi Transfer dan VAR

	Pengujian MAPE	Model ARIMA				Fungsi Transfer		Model VAR	
		ARIMA (0,2,1)	(%)	ARIMA (2,2,4)	(%)	Arima (2,2,5) Xreg=Hkonsumen	(%)	VAR (2) type=both	(%)
		MAPE Training	2,63	2,31	2,66	10,88			
	MAPE Testing	3,47	12,70	3,98	16,44				
ATAP	2016	113.617		113.617		113.617		113.617	
	2017	113.307	-0,27	113.307	-0,27	113.307	-0,27	113.307	-0,27
	2018	109.935	-2,98	109.935	-2,98	109.935	-2,98	109.935	-2,98
	2019	108.750	-1,08	108.750	-1,08	108.750	-1,08	108.750	-1,08
	2020	112.307	3,27	112.307	3,27	112.307	3,27	112.307	3,27
Angka Estimasi (AESTI)	2021	111.257	-0,93	110.060	-2,00	106.886	-4,83	120.421	7,22
	2022	110.208	-0,94	109.371	-0,63	105.045	-1,72	119.068	-1,12
	2023	109.158	-0,95	108.363	-0,92	103.171	-1,78	117.701	-1,15
	2024	108.108	-0,96	107.576	-0,73	101.358	-1,76	116.312	-1,18
	2025	107.059	-0,97	106.877	-0,65	99.524	-1,81	114.899	-1,21
Rata-rata Pertumbuhan	ATAP 2016 - 2020		-0,26		-0,26		-0,26		-0,26
	AESTI 2021 - 2025		-0,96		-0,73		-1,77		-1,17

## KESIMPULAN

Model yang digunakan untuk penyusunan angka estimasi luas areal teh antara lain model *Autoregresif Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Autoregressive* (VAR). Sementara data yang terdapat pada statistik perkebunan merupakan hasil sinkronisasi dan validasi yang dilaksanakan secara berjenjang dari tingkat kabupaten/kota sampai tingkat nasional dengan mengacu pada Pedoman Pengelolaan Data Komoditas Perkebunan (PDKP) yang diterbitkan oleh Direktorat Jenderal Perkebunan. Secara periodik, data perkebunan disajikan dalam 3 status angka yaitu Angka Tetap (ATAP), Angka Sementara (ASEM) dan Angka Estimasi (AESTI). Angka Tetap merupakan angka hasil rekapitulasi dari pelaporan yang sudah tetap, sehingga tidak dilakukan estimasi, sedangkan untuk penentuan Angka Sementara dan Estimasi perlu dilakukan estimasi dengan metode estimasi yang paling relevan dan tepat (PDKP, 2013). Data Angka Tetap (ATAP) merupakan data 2 tahun yang lalu ( $n-2$ ), Angka Sementara (ASEM) merupakan data tahun lalu ( $n-1$ ), dan Angka Estimasi (AESTI) merupakan data tahun yang berjalan ( $n$ ).

Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1980 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan. Untuk estimasi luas areal teh alternatif model pertama adalah Model ARIMA. Model ARIMA terbaik adalah ARIMA (0,2,1), menghasilkan MAPE untuk data training 2,63%, dan MAPE data testing 3,47%. Model yang kedua dengan menggunakan Fungsi Transfer dengan variabel input harga konsumen teh, dengan MAPE untuk data training sebesar 2,66% dan MAPE data testing 3,98%. Untuk model yang ketiga model VAR(2) type 'both' ada pengaruh trend dan konstanta, menghasilkan MAPE data training 10,88% dan data MAPE data testing 16,44%.

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen. Model VAR berlaku pada saat nilai setiap variabel dalam sebuah system tidak hanya bergantung pada lag-nya sendiri, namun juga pada nilai lag variabel lain.

Tahapan yang dilakukan dalam penyusunan model VAR adalah sebagai berikut: persiapan data, pembagian data training dan testing, pemilihan lag dan type, pengajuan asumsi, ramalan data training, testing, penghitungan mape, dan plot, pemilihan model terbaik, dan pengepasan model



untuk seluruh data dan peramalannya. Disamping itu dibahas Interpretasi Impulse Response Function (IRF) dan Variance Decomposition.

Berdasarkan perbandingan besarnya MAPE baik data testing maupun data training dan hasil estimasi produksi 5 tahun kedepan, maka model terbaik yang terpilih adalah model ARIMA karena menghasilkan akurasi yang cukup tinggi sehingga MAPE data training sebesar 2,63% dan MAPE data testing sebesar 3,47%. Hasil estimasi luas areal teh nasional dalam satuan hektar untuk tahun 2021 sebesar 106.887 ha, tahun 2022 sebesar 105.045 ha, tahun 2023 sebesar 103.171 ha, tahun 2024 sebesar 101.358 ha, dan tahun 2025 sebesar 99.524 ha. Laju pertumbuhan estimasi luas areal teh nasional selama 5 tahun kedepan rata-rata turun sebesar 1,77%.

## DAFTAR PUSTAKA

- Budiar, Aindra. 2013. Peramalan Produksi TBS Kelapa Sawit Dengan Model Arima dan Model Fungsi Transfer Input Ganda. Institut Pertanian Bogor, Bogor
- Gujarati, D.N. and D.C. Porter, 2010. Dasar-dasar Ekonometrika. Jakarta: Penerbit Salemba Empat, Buku 2, Edisi 5.
- Heizer, J., Render, B. & Munson, C., 2011. Operations Management Sustainability and Supply Chain Management. Boston: Pearson.
- Kementerian Pertanian. 2013. Pedoman Teknis Pengembangan Tanaman Teh Tahun 2014. Jakarta: Direktorat Jenderal Perkebunan Kementerian Pertanian.
- Kementerian Pertanian. 2019. Statistik Perkebunan Indonesia 2018-2020 (Teh). Jakarta: Direktorat Jenderal Perkebunan.
- Mardiyanto, Atqo. 2000. Kajian Peramalan Dengan Model Struktural dan Non Struktural (VAR dan ARIMA). Institut Pertanian Bogor, Bogor
- Makridakis S, Wheelwright SC, McGee VE. 1999. Metode dan Aplikasi Peramalan. Suminto H, penerjemah. Binarupa Aksara, Jakarta.
- M. Firdaus 2020. Aplikasi Ekonometrika dengan E-Views, Stata dan R. Institut Pertanian Bogor (IPB) Press.
- Tsay, R.S., 2014. Multivariate Time Series Analysis With R and Financial Applications. New Jersey: John Wiley & Sons

# KAJIAN METODOLOGI ESTIMASI DATA LUAS AREAL CENGKEH

Vera Junita Siagian, Fungsional Statistisi, Pusdatin Kementan

## ABSTRAK

Pada tahun 2021, Pusdatin bekerjasama dengan Ditjen Perkebunan dan BPS mengkaji 3 metode estimasi yang nantinya didapatkan metode yang lebih akurat, lebih objektif dan lebih baik secara statistik dari metode sebelumnya dan diharapkan hasil estimasinya cenderung mendekati angka aktualnya. Melalui kegiatan pengembangan metode estimasi data perkebunan tahun 2021, diharapkan akan dihasilkan metode yang paling sesuai untuk mengestimasi luas areal cengkeh nasional dengan membandingkan besaran MAPE (Mean Absolut Percentage Error) dan kerealistisan hasil permalan dengan data series sebelumnya

Tujuan dari makalah ini adalah untuk membandingkan efektivitas metode tersebut dalam hasil estimasi luas areal cengkeh dan menentukan metode terbaik dalam estimasi luas areal cengkeh nasional. Metode yang diterapkan adalah ARIMA (*Autoregressive Integrated Moving Average*), Fungsi Transfer dan VAR (*Vector Auto Regression*) dengan menggunakan software R Studio

## ABSTRACT

*In 2021, Center for Agriculture Data and Informasi System (CADIS) together with the Directorate General of Plantation and the BPS-Statistic Indonesia examined 4 methods of estimation that will be obtained more accurately, more objectively and statistically better than the previous method and it is expected that the estimated results tend to be close to the actual number. Through the development of plantation data estimation method in 2020, it is expected that the most suitable method will be produced to estimate national clove production by comparing the amount of MAPE (Mean Absolut Percentage Error) and the reality of the results of the permalan with the previous series data*

*The purpose of this paper is to compare the effectiveness of such methods in the estimated production of cloves and determine the best method in estimating national clove production. The methods applied are ARIMA (Autoregressive Integrated Moving Average), Transfer Function and VAR (Vector Auto Regression) using R Studio software*

## PENDAHULUAN

Cengkeh atau *Syzygium aromaticum* merupakan jenis tumbuhan perdu yang memiliki batang pohon besar dan berkayu keras dari keluarga pohon *Myrtaceae*. Cengkeh adalah tanaman asli Indonesia, banyak digunakan sebagai bumbu masakan pedas di negara-negara Eropa, dan sebagai bahan utama rokok kretek khas Indonesia. Cengkeh merupakan salah satu komoditas sub sektor perkebunan yang sebagian besar (97,43%) diusahakan oleh perkebunan rakyat. Hasil utama tanaman Cengkeh adalah bunganya yang dipanen pada saat kelopak bunga belum mekar. Bunga cengkeh kering merupakan salah satu bahan baku utama untuk rokok kretek yang merupakan rokok khas Indonesia.

Luas areal cengkeh pada 2016-2020 (Angka Sementara) mengalami peningkatan seluas 1,34% per tahun. Dari luasan tersebut sekitar 66.84% diusahakan oleh petani rakyat (perkebunan rakyat) yang dibudidayakan secara monokultur maupun tumpang sari dengan tanaman lainnya. Pada tahun 2020 (Angka Sementara) luas area cengkeh sebesar 574.576 hektar. Untuk 2021 diprediksikan meningkat 0,04% menjadi 574.999 ha. Peningkatan luas area selain karena pendataan ulang oleh petugas kabupaten, juga diluncurkannya program Mengembalikan Kejayaan Rempah Indonesia (Ditjenbun, 2020). Dilihat dari produktivitas, tahun 2021 produktivitas cengkeh nasional diperkirakan sebesar 416 kg/ha. Dibandingkan tahun lalu yaitu tahun 2020 sedikit naik yaitu sebesar 416 kg/ha.

Metode yang dilakukan selama ini untuk menyusun angka estimasi (AESTI) luas areal cengkeh nasional menurut Buku Pedoman Pelaksanaan Pengeolaan Data Komoditas Perkebunan (PDKP) adalah metode pemulusan eksponensial tunggal atau Single Exponential Smoothing (SES) dan pemulusan eksponensial ganda atau DSS (Double Exponential Smoothing). Untuk pemilihan model terbaik yaitu dengan membandingkan besaran MAPE (Mean Absolut Percentage Error) dan kerealistisan hasil permalan dibandingkn dengan data series sebelumnya. Menurut T. Hani Handoko.jkflkdhflhk (2011), *Exponential Smoothing* adalah suatu tipe teknik peramalan rata-rata bergerak yang melakukan penimbangan terhadap data masa lalu dengan cara eksponensial sehingga data paling akhir mempunyai bobot atau timbangan lebih besar dalam rata-rata bergerak.

Pada tahun 2021 ini, Pusdatin mengkaji 3 (tiga) metode estimasi guna menghasilkan angka yang nilainya semakin mendekati nilai sebenarnya. Ke tiga metode yang di kaji yaitu ARIMA (*Autoregressive Integrated Moving Average*), Fungsi Transfer dan VAR (*Vector Auto Regression*) dengan menggunakan software R Studio

## MATERI DAN METODE

Data yang digunakan dalam kajian ini adalah data sekunder yang bersumber dari Direktorat Jenderal Perkebunan dan Badan Pusat Statistik (BPS). Untuk data luas areal dan luas areal cengkeh bersumber dari Direktorat Jenderal Perkebunan sementara untuk data harga produsen dan volume ekspor bersumber dari BPS dengan series data tahunan yaitu dari tahun 1983 sampai tahun 2020.

### Model Arima

ARIMA sering juga disebut metode runtun waktu Box-Jenkins. ARIMA sangat baik ketepatannya untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatan peramalannya kurang baik. Biasanya akan cenderung flat (mendatar/konstan) untuk periode yang cukup panjang. Model Autoregresif Integrated Moving Average (ARIMA) adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari deret waktu (time series) secara statistik berhubungan satu sama lain (dependent). ARIMA sangat baik ketepatannya untuk peramalan jangka pendek, yang tidak membentuk suatu model struktural baik itu persamaan tunggal atau simultan yang berbasis kepada teori ekonomi atau logika. Beberapa keuntungan yang dapat diperoleh dengan menggunakan ARIMA: 1.

- Merupakan model tanpa teori karena variabel yang digunakan adalah nilai lampau dan kesalahan yang mengikutinya.
- Memiliki tingkat akurasi peramalan yang cukup tinggi karena setelah mengalami pengukuran kesalahan peramalan mean absolute error, nilainya mendekati nol.
- Cocok digunakan untuk meramal sejumlah variabel dengan cepat, sederhana, akurat dan murah karena hanya membutuhkan data variabel yang akan diramal.

Model ARIMA menggunakan pendekatan iteratif dalam indentifikasi terhadap suatu model yang ada. Model yang dipilih diuji lagi dengan data masa lampau untuk melihat apakah model tersebut menggambarkan keadaan data secara akurat atau tidak. Suatu model dikatakan sesuai (tepat) jika residual antara model dengan titik-titik data historis bernilai kecil, terdistribusi secara acak dan bebas satu sama lainnya. Pemilihan model terbaik dapat dilakukan dengan membandingkan distribusi koefisien-koefisien autocorrelation (otokorelasi) dari data time series tersebut dengan distribusi teoritis dari berbagai macam mode.

## Model Fungsi Transfer

Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (*Autoregressive Integrated Moving Average*). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variable lain sebagaimana pada model regresi. Model fungsi transfer dengan variabel input lebih dari dua deret berkala disebut dengan model fungsi transfer multivariat. Prosedur pembentukan model fungsi transfer multivariat melalui dua tahap, yaitu pembentukan model fungsi transfer tunggal dari masing-masing input, baru dilakukan pembentukan model fungsi transfer secara simultan dari semua variabel.

Beberapa hal yang berkaitan dengan model fungsi transfer antara lain deret berkala output, disebut  $Y_t$ , yang diperkirakan akan dipengaruhi oleh deret berkala input, disebut  $X_t$ , dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (noise)  $N_t$ . Model fungsi transfer bivariat ditulis dalam dua bentuk umum. Bentuk pertama adalah sebagai berikut: (Makridakis, dkk: 1999:448)

$$Y_t = v(B) X_t + N_t$$

Model Var

## Model VAR (*Vector Auto Regression*)

Model VAR digunakan untuk memproyeksikan sebuah sistem dengan variabel runtut waktu dan untuk menganalisis dampak dinamis dari faktor gangguan yang terdapat dalam sistem variabel tersebut. Pada dasarnya Analisis VAR sama dengan suatu model persamaan simultan, karena dalam Analisis VAR kita mempertimbangkan beberapa variabel endogen secara bersama-sama dalam suatu model. Sebenarnya analisis ini mirip dengan model persamaan simultan biasa. Hanya saja di dalam Analisis VAR masing-masing variabel selain diterangkan oleh nilainya di masa lampau, juga dipengaruhi oleh nilai masa lalu dari semua variabel endogen lainnya dalam model yang diamati. Selain itu, dalam analisis VAR biasanya tidak ada variabel eksogen di dalam model.

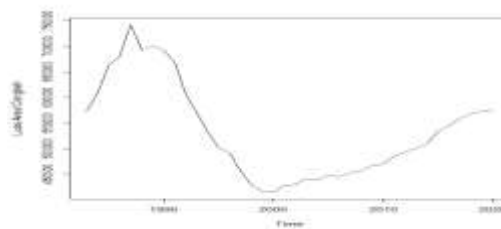
## HASIL DAN PEMBAHASAN

R Studio adalah bahasa pemrograman dan sistem perangkat lunak yang dirancang khusus untuk mengerjakan segala hal terkait komputasi statistik. Bahasa pemrograman ini pertama kali dikembangkan pada tahun 1993 oleh dua orang pakar statistik yaitu Ross Ihaka dan Robert Gentleman di Auckland University, New Zealand. Sampai saat ini, bahasa pemrograman R terus berkembang secara pesat seiring dengan semakin populernya terminologi “*Big Data*” dan meningkatnya kebutuhan perusahaan akan data *scientist* untuk mengolah dan menganalisis data di perusahaan tersebut sebagai dasar pengambilan kebijakan dan mengotomatisasi proses bisnis

menjadi data *driven*. Bahasa pemrograman seperti Python dan R telah menjadi pilihan utama bagi para peneliti maupun praktisi di bidang data *science* untuk mengolah dan menganalisis data baik itu untuk kepentingan penelitian maupun bisnis. Oleh karena itu, bagi seorang pemula di bidang data *science*, R merupakan bahasa pemrograman yang sangat *recommended* untuk dikuasai.

### Model Arima

Syarat utama dalam melakukan pemodelan ARIMA adalah kestasioneran data. Kestasioneran data dapat diketahui secara visual (plot datanya) dan juga dengan uji statistik. Dari plot data gambar 1 terlihat luas areal cengkeh tahun 1983-2020 memiliki trend sehingga terindikasi tidak stasioner.



Gambar 1. Plot L. Areal Cengkeh Tahun 1983-2020

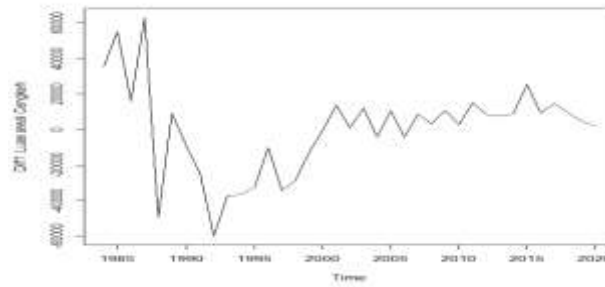
Berdasarkan gambar 1 terlihat bahwa data luas areal cengkeh belum stasioner. Untuk lebih memastikan apakah data stasioner digunakan uji ADF (*Augmented Dickey-Fuller*). Jika test statistics lebih besar dari critical value maka data tersebut tidak stasioner. Berdasarkan hasil uji *Augmented Dickey-Fuller* dapat dilihat seperti hasil dibawah ini.

Tabel 1. Hasil Uji Augmented Dickey-Fuller

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Residuals:
  Min   1Q Median   3Q   Max
-62930 -8936  -68  6825 66812
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.972e+04  3.216e+04  0.613  0.5441
z.lag.1    -4.567e-02  4.935e-02 -0.926  0.3616
tt         1.862e+02  4.511e+02  0.413  0.6825
z.diff.lag  4.237e-01  1.580e-01  2.681  0.0115 *
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 23160 on 32 degrees of freedom
Multiple R-squared:  0.2303, Adjusted R-squared:  0.1581.
F-statistic: 3.192 on 3 and 32 DF, p-value: 0.03669
Value of test-statistic is: -0.9256 0.7132 1.0416
Critical values for test statistics:
  1pct  5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```

Dari hasil uji Augmented Dickey-Fuller diatas terlihat bahwa nilai tes statistic adalah -0.9256 dan nilai critical tau 3 untuk 1% sebesar -4,15, taraf uji 5% sebesar -3,50 dan taraf uji 10% sebesar -3,18. Dari hasil uji Augmented Dickey-Fuller ini dapat disimpulkan data tidak stasioner hingga pada taraf 1%, 5% dan 10% yang berarti tolak  $H_0$ . Suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melakukan *differencing*, jika *differencing* 1 belum stasioner maka dilakukan *differencing* 2 sampai data stasioner. Setelah differencing satu kali, secara visual sudah terlihat data stasioner (Gambar 2) dan untuk lebih memastikan dapat di uji dengan uji ADF.





Gambar 2. Plot L. Areal Cengkeh Tahun 1983-2020

Diff 1

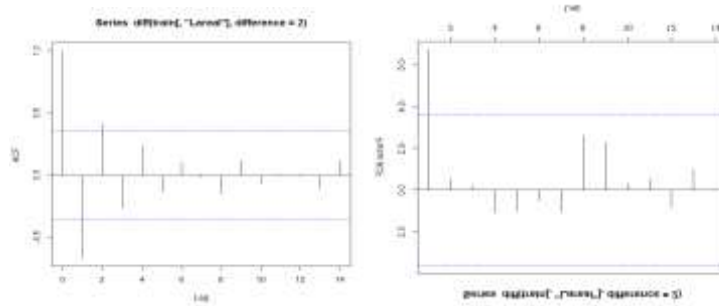
Tabel 2. Hasil Uji Augmented Dickey-Fuller Diff1

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
  Min   1Q Median   3Q   Max
-65834 -7572  972  5843 34216
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
z.lag.1  -0.3805   0.1499  -2.539  0.01603 *
z.diff.lag -0.4811   0.1377  -3.493  0.00138 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 19150 on 33 degrees of freedom
Multiple R-squared:  0.5357,    Adjusted R-squared:  0.5076
F-statistic: 19.04 on 2 and 33 DF, p-value: 3.176e-06
Value of test-statistic is: -2.5386
Critical values for test statistics:
  1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
Critical values for test statistics:
  1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

Hasil uji Augmented Dickey-Fuller *differencing* 1. terlihat bahwa nilai tes statistic adalah -**2.5386** dan nilai critical tau 1 untuk 1% sebesar **-2,62**, taraf uji 5% sebesar -1,95 dan taraf uji 10% sebesar -1,61% sehingga dapat disimpulkan data sudah stasioner pada taraf 5% dan 10%.

### Identifikasi Model

Pada tahap ini akan ditentukan nilai p,d dan q dimana penentuan p dan q dapat dilakukan dengan beberapa cara yaitu dengan melihat korelogram autokorelasi (ACF) dan korelogram autokorelasi parsial (PACF), autoarima serta ar maselect. Sebelum menentukan nilai p,d,dan q terlebih dahulu data luas areal cengkeh dibagi menjadi data training yaitu data dari tahun 1980-2014 dan data testing mulai tahun 2015-2020. Data training digunakan untuk membangun model, sedangkan data testing digunakan untuk validasi dan pengepasan model. Gambar plot ACF dan PACF luas areal cengkeh



Gambar 3. Plot ACF dan PACF Luas areal Cengkeh

Dari gambar plot ACF pada Gambar 3 menunjukkan nyata pada lag=0 , sementara untuk lag=1 dan seterusnya berada dibawah garis selang kepercayaan dan tidak menunjukkan pola *cut off* maupun *tail off* sehingga agak kesulitan untuk mengidentifikasi model ARIMA nya. Apabila terjadi kesulitan mengidentifikasi model tentatif ARIMA menggunakan plot ACF dan PACFnya maka dapat digunakan *autoarima*. Hasil autoarima adalah sebagai berikut:

Tabel 3. Model Arima Berdasarkan Automodel

Series: train[, "Lareal"]

ARIMA(2,1,0)

Coefficients:

ar1 ar2

0.2262 0.5203

s.e. 0.1465 0.1561

sigma^2 estimated as 473834069: log likelihood=-353.03

AIC=712.06 AICc=712.95 BIC=716.36

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -873.3783 20722.26 14146.12 -0.02679157 2.494327 0.700784 0.001571991

dengan *Standard error* sebesar 0.1465 dan nilai MAPE data training sebesar 2,49% seperti terlihat pada tabel diatas. Selain cara diatas dapat juga dijalankan dengan perintah *armaselect*, maka R-Studio akan menunjukkan sepuluh (10) model tentative yang diurutkan berdasarkan nilai SBC terkecil. Sepuluh model tentative yang ditampilkan berdasarkan *armaselect* diff 1 yaitu:

Tabel 4. Model Arima Berdasarkan Model Armaselect

	p	q	sbc
[1,]	2	4	526.8942
[2,]	3	4	529.3674
[3,]	2	5	529.9539
[4,]	4	4	532.2613
[5,]	3	5	533.2519
[6,]	5	4	533.8430
[7,]	5	5	534.5775
[8,]	4	5	536.4695
[9,]	3	1	544.0887
[10,]	0	2	544.3371

Dari hasil model autoarima dan 10 model tentative hasil *armaselect diff 1* dicobakan satu per satu dengan melihat mape terkecil dan kelayakan modelnya serta kelogisan hasil ramalan datanya dan dibandingkan dengan autoarima yang disarankan maka yang dipilih adalah hasil *autoarima*

ARIMA (2,1,0).

Tabel 5. Model Arima Berdasarkan Automodel

Call:

```
arma(x = train[, "Lareal"], order = c(2, 1, 0))
```

Coefficients:

```
      ar1  ar2
0.2262 0.5203
s.e. 0.1465 0.1561
```

sigma^2 estimated as 443253551: log likelihood = -353.03, aic = 712.06

z test of coefficients:

```
Estimate Std. Error z value Pr(>|z|)
ar1 0.22619 0.14650 1.5440 0.1225952
ar2 0.52031 0.15610 3.3332 0.0008584 ***
```

---

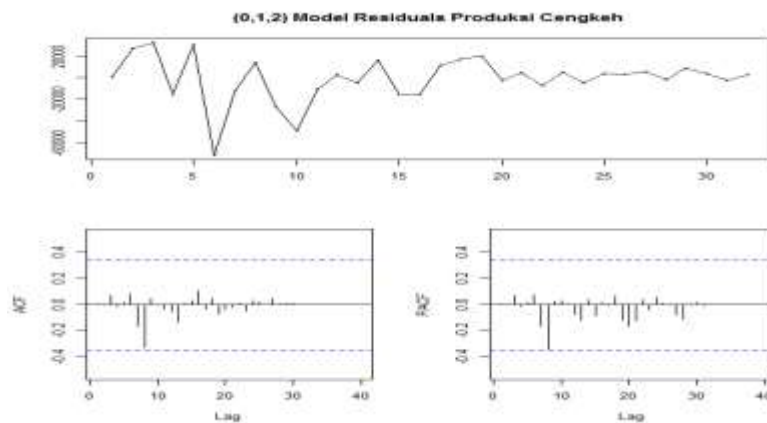
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dari Hasil penelusuran model ARIMA (2,1,0) diatas menunjukkan parameter ar 2 adalah sangat nyata. Kemudian dilihat sebaran sisaan LJung-Box dari lag 5 sampai lag 30. Salah satu

syarat kebaikan model ARIMA adalah sebaran sisaan LJung-Box. Hasil pengujian nilai p-value pada lag 5 sampai dengan lag 30 tidak ada yang signifikan, sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (0,2,1) layak digunakan. Hasil sisaan dengan Uji LJung-Box dan hasil plot sisaan adalah sebagai berikut:

Tabel 6. Uji Sisaan Uji LJung-Box Model Arima (0,2,1)

lags	statistic	df	p-value
5	0.1893069	5	0.9992247
10	6.6881152	10	0.7545241
15	8.0485824	15	0.9218202
20	9.6737143	20	0.9737099
25	10.1985012	25	0.9960836
30	10.6023367	30	0.9995869



Gambar 4. Plot Sisaan Model Arima (0,2,1)

Hasil plot sisaan model ARIMA (0,2,1) menunjukkan bahwa plot ACF dan PACF sudah tidak menunjukkan pola *cut off* atau *tail off* yang merujuk ke model ARIMA tertentu sehingga model ARIMA (0,2,1) sudah cukup layak.

Kemudian dilakukan uji coba peramalan dengan menggunakan data testing. Hasil peramalan data testing selama lima tahun kedepan yaitu tahun 2021 – 2025 tersaji di bawah ini.:

Tabel 7. Hasil Peramalan Data Testing

Time Series:

Time Series:

Start = 34

End = 38

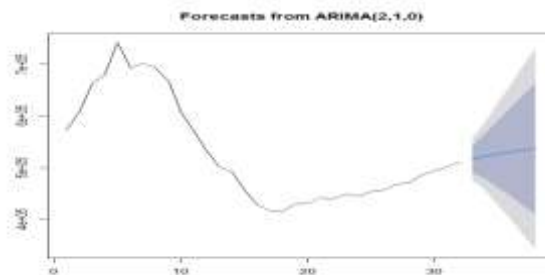
Frequency = 1

[1] 516062.4 521971.3 526371.1 530440.8 533650.6 536494.1

> accuracy(ramalan\_luas areal,test[, "Lareal"])

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	-873.3783	20722.26	14146.12	-0.02679157	<b>2.494327</b>	0.700784	0.001571991
Test set	32338.9593	33350.33	32338.96	5.74158694	<b>5.741587</b>	1.602039	NA

Dari table diatas terlihat nilai MAPE yang yang diperoleh 2,49% untuk data training dan 5,74% untuk data set testing. Plot data dari tahun 1980-2015 (set data training) dan garis berwarna biru merupakan hasil peramalan data set testing (2015-2020), yang dilengkapi dengan arsiran selang kepercayaannya (daerah biru muda).



Gambar 5. Plot Ramalan L.Areal Cengkeh Data Testing

Dengan model ARIMA (2,1,0) dilakukan pengepasan seluruh data dengan hasil sebagai berikut:

Tabel 8. Hasil Pengepasan Seluruh Data

Call:

```
arima(x = data_luas areal[, "Lareal"], order = c(2, 1, 0))
```

Coefficients:

```
ar1 ar2
```

```
0.2312 0.5243
```

```
s.e. 0.1348 0.1433
```

```
sigma^2 estimated as 382645819: log likelihood = -418.57, aic = 843.13
```

Training set error measures:

```
ME RMSE MAE MPE MAPE MASE ACF1
```

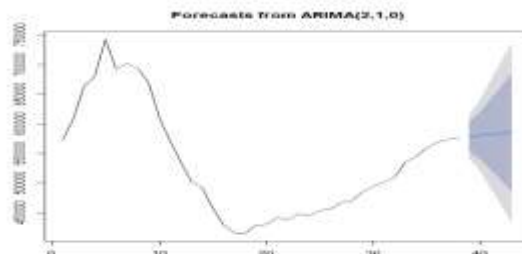
```
Training set -486.195 19302.46 12749.56 0.0293879 2.253672 0.6822797 0.001761793
```

Dari hasil pengepasan untuk seluruh data yaitu tahun 1983-2020 maka dihasilkan MAPE sebesar 2,25% Dengan menggunakan model ARIMA (2,1,0) kemudian dilakukan peramalan untuk tahun 2021-2025. Hasil peramalan dan plot dengan ARIMA (2,1,0) adalah sebagai berikut:

Tabel 9.. Hasil Peramalan Model ARIMA (0,2,1)

Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
39	578789.0	553720.1	603857.8	540449.5
40	580494.0	540731.9	620256.1	519683.1
41	582448.4	522138.1	642758.6	490211.8
42	583794.0	504336.1	663252.0	462273.6
43	585129.7	485077.9	685181.6	432113.6

Untuk plot estimasi luas areal cengkeh tahun 2021-2025 seperti gambar 6.



Gambar 6. Plot Hasil Estimasi Tahun 2020-2025

Tabel 9. Hasil Peramalan Model ARIMA (2,1,0)

Tahun	Produksi (Ton)	Pertumbuhan (%)
2021	578,789	
2022	580,494	0.29
2023	582,448	0.34
2024	583,794	0.23
2025	585,129	0.23
Rata-rata pertumbuhan (%)		0.27
Mape Training	2.49%	
Mape Testing	5.74%	

### Model Fungsi Transfer

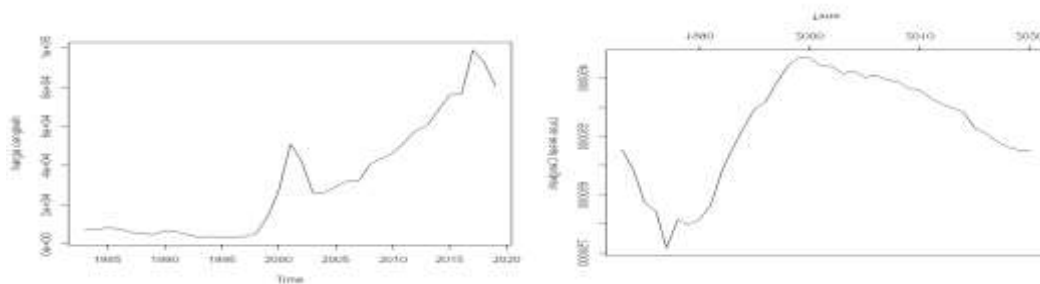
Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (Autoregressive Integrated Moving Average). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variabel lain sebagaimana pada model regresi. Model fungsi transfer merupakan suatu model yang menggambarkan nilai prediksi dari suatu time series (deret output atau  $Y_t$ ) berdasarkan nilai-nilai dari deret itu sendiri ( $Y_t$ ) dan berdasarkan pula pada data time series yang mempunyai hubungan (deret input atau  $X_t$ ) dengan deret output. Model fungsi transfer yang bersifat dinamis berpengaruh tidak hanya pada hubungan linier antara waktu ke- $t$  input  $X_t$  dan waktu ke- $t$  output  $Y_t$ , tetapi juga saat input  $X_t$  dengan saat  $t, t+1, \dots, t+k$  pada output  $Y_t$  (Fitriani et al., 2013). Dalam melakukan estimasi dengan fungsi transfer diperlukan dua variabel, yaitu output  $Y_t$  dan input  $X_t$ . Input  $X_t$  merupakan variabel yang diduga paling berpengaruh terhadap output  $Y_t$ . Bentuk umum fungsi transfer single input ( $X_t$ ) dengan single output ( $Y_t$ ).  $Y_t = v(B)X_t + N_t$ . Pada analisis ini estimasi untuk luas areal cengkeh (output  $Y_t$ ) yang diduga dipengaruhi oleh harga cengkeh di tingkat produsen.

Tahap pembentukan model fungsi transfer untuk luas areal cengkeh adalah sebagai berikut:

- Mempersiapkan deret input ( $X_t$ ) dan output ( $Y_t$ ) yang stasioner.
- Pembagian series data awal menjadi series data training dan testing
- Penentuan Model ARMA untuk deret input.
- Prewhitening deret input dan output
- Mendeteksi dan mengukur hubungan kekuatan antara  $\alpha_t$  dan  $\beta_t$  dengan menggunakan CCF (Cross Correlation Function).

- Penetapan  $(b, r, s)$  yang menghubungkan deret *input* dan deret *output* dimana  $b$  adalah panjang jeda pengaruh  $X_t$  terhadap  $Y_t$ ,  $r$  adalah panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$ ,  $s$  adalah panjang lag  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- Penaksiran awal deret noise  $(\pi_t)$
- Penentuan model ARIMA deret noise
- Estimasi parameter dengan fungsi transfer
- Uji kesesuaian model fungsi transfer
- Menggunakan model fungsi transfer untuk peramalan.

Eksplorasi data dilakukan dengan menampilkan plot data luas areal dan harga cengkeh. Data luas areal dan harga cengkeh dimulai tahun 1983-2020 yaitu sebanyak 38 series akan dibagi menjadi series data training untuk periode tahun 1983-2014 dan series data testing untuk periode tahun 2015-2020. Plot data luas areal dan harga cengkeh dari tahun 1983-2020.



Gambar. 7. Plot data luas areal dan harga cengkeh tahun 1983-2020

Dari plot data harga cengkeh terlihat bahwa adanya peningkatan dari tahun ke tahun sehingga terindikasi data tidak stasioner. Untuk pengujiannya dapat dilakukan dengan uji Augment Dickey-Fuller. Hasil Augmen Dickey Fuller adalah sebagai berikut:



Tabel 10. Hasil Uji Augmented Dickey-Fuller

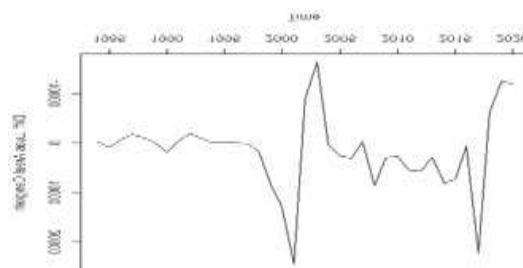
```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Residuals:
    Min     1Q   Median     3Q    Max
-12902.8 -4184.7 -653.6  2357.0 23577.0
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -4222.3391  3059.1298  -1.380  0.17708
z.lag.1      -0.3063    0.1034  -2.963  0.00571 **
tt           768.6126  279.5474   2.749  0.00973 **
z.diff.lag   0.3810    0.1657   2.300  0.02814 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 7135 on 32 degrees of freedom
Multiple R-squared:  0.2652,    Adjusted R-squared:  0.1963
F-statistic: 3.849 on 3 and 32 DF,  p-value: 0.0185
Value of test-statistic is: -2.3877 3.1634 3.7204
Critical values for test statistics:
    1pct  5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```

Dari hasil uji Augmented Dickey-Fuller terlihat bahwa data tidak stasioner karena nilai *tes statistic* lebih besar dari nilai *critical* dimana nilai *tes statistic nya* adalah -2,3877 dan nilai critical tau 3 untuk 1% sebesar -4,15, taraf uji 5% sebesar -3,50 dan taraf uji 10% sebesar -3,18. Dari hasil uji Augmented Dickey-Fuller ini dapat disimpulkan data tidak stasioner hingga pada taraf

1%, 5% dan 10% yang berarti tolak  $H_0$ . Untuk menstasionerkan data sehingga dilakukan differencing 1 kali. Hasil uji Augmented Dickey-Fuller Diff 1 adalah sebagai berikut:

Tabel 11. Hasil Uji Augmented Dickey-Fuller Diff 1

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
    Min     1Q   Median     3Q     Max
-14910.5  -845.8   58.4  4281.0 23453.2
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
z.lag.1   -0.8439    0.2185  -3.863 0.000496 ***
z.diff.lag  0.1898    0.1807   1.050 0.301263
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 7886 on 33 degrees of freedom
Multiple R-squared:  0.3553, Adjusted R-squared:  0.3163
F-statistic: 9.095 on 2 and 33 DF, p-value: 0.0007143
Value of test-statistic is: -3.8626
Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

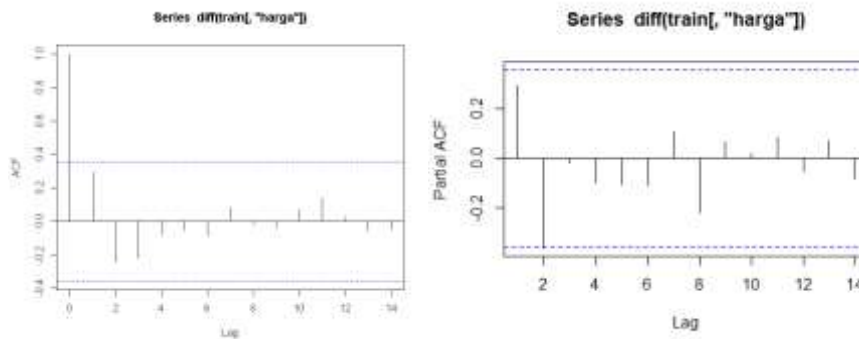


Gambar. 7. Plot data harga cengkeh tahun 1983-2020

Dari hasil uji Augmented Dickey-Fuller setelah di differencing 1 kali terlihat bahwa nilai *tes statistic* sudah lebih kecil dibandingkan nilai *criticalya* dimana nilai *tes statistic* adalah -3.8626 sementara nilai critical tau 1 untuk 1% sebesar -2,62, taraf uji 5% sebesar -1,95 dan taraf uji 10%

sebesar -1,61%. Hal ini dapat disimpulkan data sudah stasioner hingga pada taraf uji 1%, 5% dan 10%. Demikian juga dari gambar diatas menunjukkan data sudah stasioner

Stasioneritas data dapat juga dilihat dari plot ACF dan PACF. Jika lag pada ACF maupun PACF menunjukkan *tail off* maka data tidak stasioner. Berikut adalah plot ACF dan PACF harga Cengkeh setelah differencing 1.



Gambar 8. Plot ACF dan PACF Harga Cengkeh

Berdasarkan plot di atas, terlihat bahwa plot ACF cenderung cut off dimana hanya lag 0 yang keluar dari *confidence interval*. Plot PACF menunjukkan cut off pada lag 2, ditunjukkan berada di luar garis *confidence interval*. Hal tersebut memperkuat hasil uji ADF yaitu data harga cengkeh telah stasioner.

**Pencarian Model Tentatif Variabel Input**

Untuk mencari model tentatif variable input (harga), dapat dilakukan dengan penelusuran menggunakan model ARIMA. Model terbaik dapat dipilih melalui *autoarima* yang tersedia pada R Studio maupun *arma selec*. Data yang digunakan untuk memilih model terbaik adalah series data trainingnya. Hasil dari *autoarima* adalah sebagai berikut:

Tabel 12. Hasil Model *Autoarima*

Series: train.h[, "harga"]

ARIMA(0,1,1)

Coefficients:

ma1

0.5260

s.e. 0.1604

sigma^2 estimated as 37926105: log likelihood=-324.28

AIC=652.56 AICc=652.97 BIC=655.49

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 1387.974 5968.882 3513.897 3.377669 **14.66832** 0.8460786 -0.0624651

Hasil *autoarima* yang disarankan untuk harga cengkeh adalah ARIMA (0,1,1) dengan Standard error sebesar 0,1604, nilai MAPE data training sebesar 14,67% dan nilai AIC 652.56 seperti terlihat pada tabel diatas.

Selanjutnya dilakukan pengujian *coeftest* pada model ARIMA(0,1,1) untuk mengetahui signifikansi dari koefisien MA pada model tersebut, output *coeftest*nya adalah sebagai berikut:

Tabel 13. Hasil Coeftes untuk Arima ( 0.1,1)

Call:

arima(x = train[, "harga"], order = c(0, 1, 1))

Coefficients:

ma1

0.5051

s.e. 0.1682

sigma^2 estimated as 37573843: log likelihood = -314.48, aic = 632.97

z test of coefficients:

Estimate Std. Error z value Pr(>|z|)

ma1 0.50513 0.16825 3.0023 0.002679 \*\*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

ME RMSE MAE MPE MAPE MASE ACF1

Training set 1387.974 5968.882 3513.897 3.377669 **14.66832** 0.8460786 -0.0624651

Test set 5733.072 12422.430 9861.094 5.228701 **11.19244** 2.3743615 NA

Berdasarkan output di atas terlihat bahwa komponen  $ma_1$  signifikan, pada taraf 1%, 5% dan 0% dengan mape 15,18% dan mape data training sebesar 14,66% dan mape data testing 11,19%. Untuk mencari model terbaik dapat dilakukan dengan perintah *armaselect*. Selain menggunakan *script auto arima* model tentatif dapat juga dijalankan dengan perintah *armaselect*. R-Studio akan menunjukkan sepuluh (10) model tentatif yang diurutkan berdasarkan nilai AIC terkecil. dari ordo  $p=1$  s/d 5 dan  $q=1$  s/d 5. Sepuluh model tentatif yang ditampilkan pada lag 1 untuk *arma select* adalah sebagai berikut:

Tabel 14. Hasil Model Armaselect Diff1

	p	q	sbc
[1,]	5	5	453.8245
[2,]	1	5	513.4337
[3,]	3	5	514.1288
[4,]	2	5	515.2249
[5,]	4	5	515.7160
[6,]	0	5	517.6856
[7,]	3	4	526.7095
[8,]	4	4	527.3080
[9,]	5	4	530.7385
[10,]	2	4	534.4641

Model terbaik adalah model yang memiliki nilai SBC terkecil. Model pertama yang disarankan adalah nilai  $p=5$  dan  $q=5$ . Karena data harga Cengkeh di differencing satu kali berarti  $d=1$ , artinya model yang direkomendasikan adalah ARIMA (5,1,5). Model terbaik kedua adalah ARIMA (1,1,5) sampai model terbaik ke sepuluh. Untuk mengetahui model terbaik dapat dilihat dari nilai AIC, mape dan koefisien ar dan ma nya nyata. Meskipun ARIMA (5,1,5) diurutan pertama dengan sbc yang lebih kecil, belum tentu merupakan terbaik sehingga perlu diujicobakan semua model yang disajikan oleh *armaselect*.

Setelah di uji cobakan dari 10 model tentatif diatas, mape data testing yang didapatkan diatas 10% sehingga dicobakan untuk lag 2. Model tentatif dari hasil *armaselect* untuk lag 2 adalah:

Tabel 15. Hasil Model Armaselect Diff2

	p	q	sbc
[1,]	5	5	346.0552
[2,]	3	5	448.7169
[3,]	4	5	452.0989
[4,]	2	5	463.9719
[5,]	5	4	479.0701
[6,]	1	5	490.6346
[7,]	3	4	494.7115
[8,]	4	4	497.6300
[9,]	2	4	502.0853
[10,]	1	3	506.5748

Dari 10 model tentatif *armaselect* lag 2 yang telah diuji cobakan maka ARIMA yang dipilih adalah ARIMA (1,2,3)

Tabel 16. Hasil Coefftest Arima (1,2,3)

Call:

```
arima(x = train[, "harga"], order = c(1, 2, 3))
```

Coefficients:

```
      ar1    ma1    ma2    ma3
      0.7048 -1.3891 -0.1349 0.5771
s.e. 0.1827 0.2493 0.3494 0.2019
sigma^2 estimated as 31896033: log likelihood = -304.12, aic = 618.23
```

z test of coefficients:

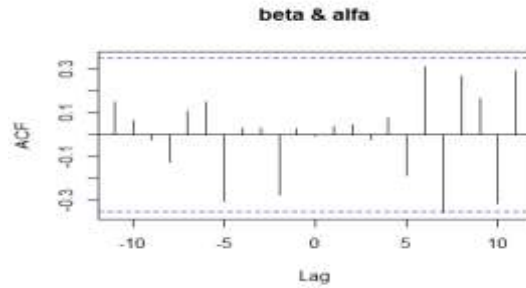
```
Estimate Std. Error z value Pr(>|z|)
ar1 0.70480 0.18270 3.8577 0.0001145 ***
ma1 -1.38905 0.24926 -5.5726 2.509e-08 ***
ma2 -0.13486 0.34937 -0.3860 0.6994874
ma3 0.57707 0.20191 2.8581 0.0042616 **
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dari hasil z test arima (1,2,3) terlihat ordo ar 1 dan ma 1 signifikan pada taraf 1%, 5% dan 10%, sementara untuk ordo ma 3 signifikan pada taraf 1% dan 5%. Oleh karena ARIMA (1,2,2) memiliki AIC yang lebih kecil maka selanjutnya dapat digunakan untuk peramalan luas areal cengkeh.

Tahapan selanjutnya untuk penyusunan model fungsi transfer ini adalah dengan prewhitening dan korelasi silang. Korelasi silang menggambarkan struktur hubungan antara  $X_t$  dengan  $Y_t$ . Untuk mengidentifikasi pengaruh  $X_t$  terhadap  $Y_t$  maka deret  $X_t$  harus stasioner atau sudah distasionerkan. Dalam konteks pemodelan  $X_t$  terhadap  $Y_t$ , untuk membuat  $X_t$  stasioner tidak dengan pembedaan (*differencing*) namun dengan mengambil komponen *white noise* dari  $X_t$  (*prewhitening*). *Prewhitening* dilakukan terhadap deret input  $X_t$  yang didefinisikan sebagai alfa serta deret input  $Y_t$  yang didefinisikan sebagai beta. Hasil grafik ACF untuk beta dan alfa.



Gambar 9. Plot Korelasi Silang

Berdasarkan grafik ACF di atas, terlihat bahwa tidak ada yang keluar dari garis signifikansi,. Karena lag signifikan pertama kali pada lag 0 maka nilai  $b=0$ . Selain itu, tidak ada tambahan lagi lag yang signifikan maka nilai  $s=0$ . Mengingat data luas areal Cengkeh dan harga Cengkeh merupakan data tahunan yang tidak mengandung musiman maka diasumsikan nilai  $r=0$ . Nilai  $b=0$  menunjukkan tidak ada jeda pengaruh antara harga pada waktu  $t$  terhadap luas areal pada waktu  $t$ . Nilai  $s=0$  berarti ada korelasi antara luas areal dan harga Cengkeh pada tahun yang sama. Dengan kata lain, dampak dari harga cengkeh terhadap luas areal dirasakan pada waktu yang sama ( $t$ ).

Tahap selanjutnya dilakukan pengepasan model, untuk nilai  $r,s$  dan  $b$ . Hasil pengujian fungsi transfer dengan nilai  $r=0$ ,  $s=0$ , dan  $b=0$  menghasilkan nilai MAPE yaitu 14,36%

Tabel 17. Pengujian Ordo  $r,s$  dan  $b$  ARIMA (0,0,0)

Series: train.h[, "Lareal"]

Regression with ARIMA(0,0,0) errors

Coefficients:

intercept    xreg  
580828.92   -2.1002

s.e.    22838.39   0.6912

sigma^2 estimated as 8.063e+09: log likelihood=-422.17

AIC=850.33   AICc=851.16   BIC=854.82

Training set error measures:

ME   RMSE   MAE   MPE   MAPE   MASE   ACF1

Training set 1.270033e-10 87028.64 75334.46 -2.543732 14.35679 3.701433 0.9093961

Untuk menghasilkan orde yang paling tepat untuk menentukan orde Arima fungsi transfer dengan melakukan identifikasi model noise. Untuk menghasilkan model terbaik dengan menggunakan auto-arima pada R Studio, model noise yang disarankan adalah Arima (1,1,0).

Model ini ternyata masih kurang tepat, karena menghasilkan MAPE yang cukup besar yaitu 55,60%.

Tabel 18. Model Residual Tentatif Arima (1,1,0)

Series: res

ARIMA(1,1,0)

Coefficients:

ar1

0.4203

s.e. 0.1646

sigma<sup>2</sup> estimated as 859385505: log likelihood=-374.14

AIC=752.29 AICc=752.7 BIC=755.22

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 2287.344 28413.05 22063.81 22.82852 **55.6033** 0.8401541 -0.09808737

Dari hasil auto arima menghasilkan mape 55,60%, karena mape yang disarankan auto arima cukup besar maka di lakukan pencarian arima dengan armaselect. Hasil tentative armaselect residual differencing 1 adalah:

Tabel 18. Model Residual Tentatif Diff 1

	p	q	sbc
[1,]	5	5	532.4180
[2,]	4	5	535.7740
[3,]	5	4	548.3926
[4,]	4	4	582.9419
[5,]	0	5	601.8163
[6,]	1	4	603.8326
[7,]	2	4	604.2768
[8,]	1	5	604.9863
[9,]	3	4	607.7040
[10,]	2	5	608.4614

Model tentatif pertama yang disarankan untuk data residual adalah ARIMA (5,1,5). Selanjutnya ke 10 model hasil armaselect di ujitobakan. Setelah semua di ujitobakan tidak ada yang signifikan sehingga di cari model tentative differencing 2. Model arima yang digunakan adalah ARMA (0,2,1)



Tabel 19. Tabel 25. Model Residual ARIMA(0,2,1) Series: res

ARIMA(0,2,1)

Coefficients:

ma1

-0.5555

s.e. 0.1619

sigma^2 estimated as 1.044e+09: log likelihood=-353.75

AIC=711.49 AICc=711.94 BIC=714.29

Dengan Arima (0,2,1) menghasilkan nilai AIC 711.49 dan standar error 0.1619

Kemudian dilakukan uji z test untuk melihat signifikan ordo ma 1.

z test of coefficients:

Estimate Std. Error z value Pr(>|z|)

ma1 -0.55546 0.16193 -3.4302 0.0006032 \*\*\*

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dari hasil z test terlihat bahwa ma 1 sangat signifikan pada tingkat kepercayaan 99%. Selanjutnya model tersebut didefinisikan sebagai model res.

Kemudian dilakukan pengepasan model noise dengan menggabungkan nilai r,s,b dan hasil identifikasi model noise yaitu ARIMA(0,2,1).

Tabel 20. Model Fungsi Transfer ARIMA (0,2,1)

Series: train[, "Lareal"]

Regression with ARIMA(0,2,1) errors

Coefficients:

ma1 xreg

-0.5234 0.3253

s.e. 0.1232 0.5915

sigma^2 estimated as 605136332: log likelihood=-345.01

AIC=696.01 AICc=696.94 BIC=700.22

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -2045.649 23010.74 14242.13 -0.1139878 2.54235 0.7055405 -0.2410441

Kemudian dilakukan uji z test

z test of coefficients:

```
Estimate Std. Error z value Pr(>|z|)
mal -0.52339 0.12315 -4.2499 2.139e-05 ***
xreg 0.32534 0.59147 0.5501 0.5823
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Berdasarkan hasil uji z test terlihat untuk koefisien mal signifikan akan tetapi untuk komponen xreg tidak signifikan.

**Peramalan berbasis fungsi transfer, data Input nilai Aktual**

Berdasarkan model fungsi transfer dengan model noise ARIMA (0,2,1) dilakukan peramalan berbasis nilai aktual dimana luas areal cengkeh diestimasi menggunakan data aktual harga cengkeh periode 2015-2020.

Tabel 22. Peramalan Dengan Data Input Nilai Aktual

Series: test[, "Lareal"]

Regression with ARIMA(0,2,1) errors

Coefficients:

```
mal xreg
-0.5234 0.3253
s.e. 0.0000 0.0000
```

sigma<sup>2</sup> estimated as 605136332: log likelihood=-37.28

AIC=76.56 AICc=78.56 BIC=75.94

Training set error measures:

```
ME RMSE MAE MPE MAPE MASE ACF1
Training set -520.6975 2136.84 1709.339 -0.09131647 0.3010311 0.2130336 -0.1742654
```

Peramalan luas areal cengkeh periode 2015-2020 menggunakan fungsi transfer ARIMA (0,2,1) dengan data input harga nilai aktual menghasilkan MAPE 0,30%

Tabel 23. Peramalan Dengan Data Input Nilai Ramalan

Series: test[, "Lareal"]

Regression with ARIMA(0,2,1) errors

Coefficients:

mal xreg

-0.5234 0.3253

s.e. 0.0000 0.0000

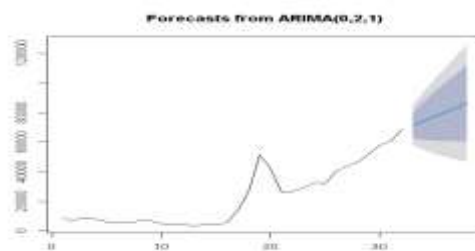
sigma<sup>2</sup> estimated as 605136332: log likelihood=-39.95

AIC=81.9 AICc=83.9 BIC=81.28

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -1789.074 4137.67 3403.304 -0.3089904 10.5980777 0.4241511 0.302475



Gambar 10. Peramalan Dengan Nilai Ramalan

### Pengepasan Model Arima Output

Pengepasan model ARIMA output dimaksudkan untuk membandingkan hasil ramalan luas areal baik berdasarkan data training maupun data testing (2015-2020). Model ARIMA biasa, tanpa fungsi transfer, terbaik untuk variabel luas areal yang dipilih adalah ARIMA (2,1,0). Model ARIMA (2,1,0) tersebut merupakan ARIMA biasa dan bukan hasil dari fungsi transfer. Hasil ramalan model ARIMA tanpa fungsi transfer sbb:

Tabel 23. Pengepasan Model Arima (2,1,0)

Series: train[, "Lareal"]

ARIMA(2,1,0)

Coefficients:

ar1 ar2

0.2262 0.5203

s.e. 0.1465 0.1561

sigma^2 estimated as 473834069: log likelihood=-353.03

AIC=712.06 AICc=712.95 BIC=716.36

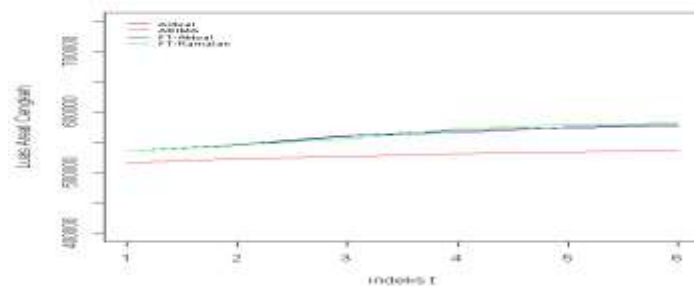
Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -873.3783 20722.26 14146.12 -0.02679157 12.494327 0.700784 0.001571991

Test set 32338.9593 33350.33 32338.96 15.74158694 5.741587 1.602039 NA

Hasil peramalan luas areal cengkeh dengan ARIMA biasa atau ARIMA tanpa fungsi transfer menunjukkan mape data training sebesar 12,49% dan mape data testing sebesar 15,74%



Gambar 11. Perbandingan Hasil Ramalan Produksi Tahun 2015-2020

Tabel Hasil Ramalan Luas Areal Cengkeh 2021-2025

Tahun	Produksi (Ton)	Pertumbuhan (%)
2021	575,465	
2022	575,691	0.04
2023	581,180	0.95
2024	595,715	2.50
2025	594,769	-0.16
Rata-rata pertumbuhan (%)		0.83
Mape Traini	12,49%	
Mape Testir	15,74%	

### Model VAR

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini tidak perlu menentukan mana variabel endogen dan variabel eksogen karena semua variabel merupakan variabel endogen. Dalam pemodelan luas areal cengkeh menggunakan tiga variabel yaitu luas areal cengkeh, harga cengkeh dan volume ekspor cengkeh dengan series tahun 1983-2020

Tahapan-tahapan penyusunan VAR adalah : persiapan data, pembagian data training dan testing, pemilihan lag dan type, pengajuan asumsi, ramalan data training, testing, penghitungan MAPE, dan plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Terakhir menggunakan *Impulse Response Function* (IRF) untuk menelusuri guncangan atau shock suatu variabel terhadap variabel lainnya.

Series data yang digunakan mulai tahun 1983-2000. Data ini akan dibagi menjadi 2 set yaitu set data training mulai tahun 1983 - 2014 dan set data testing mulai tahun 2015 -2020. Untuk pemilihan lag (p) dan bentuk type yang digunakan perlu dilakukan pengujian. Jika keduanya signifikan, maka komponen tersebut termasuk ke type “both”. Jika konstantanya yang signifikan maka termasuk model VAR type “const”. Jika keduanya tidak signifikan maka type yang digunakan adalah type “non”. Untuk pemilihan lag p nya dilakukan uji coba, trial and error dengan menggunakan model VAR dari lag p=1 s.d 5 dengan type “both” untuk mengetahui signifikan atau tidak keberadaannya. Model VAR(1) adalah model Vector Autoregressive berorde 1, artinya variabel bebas dari model tersebut hanyalah satu nilai lag dari variabel tak bebasnya. Sebelum melakukan estimasi para-meter pada model ini terlebih dahulu harus

diputuskan berapa maksimum panjang lag yang digunakan dengan melihat banyaknya yang signifikan yaitu dari mulai lag p=1-5 baik untuk type cons, both dan non.

Setelah di uji cobakan dari lag p=1-5 baik untuk type cons, both dan non maka ada 2 type yang dipilih dengan memperhatikan signifikannya yaitu p=1 type “both” dan p=1 type “cons”

Tabel 24. Output VAR p=1 type “both”:

Endogenous variables: Luas areal, harga, volek

Deterministic variables: const

Sample size: 32

Log Likelihood: -980.406

Roots of the characteristic polynomial:

1.04 0.6263 0.07462

Call:

VAR(y = cengkeh[1:33, c(2, 3, 4)], p = 1, type = "both")

Estimation results for equation Luas areal:

=====

Luas areal = Luas areal.l1 + harga.l1 + volek.l1 + const + trend

	Estimate	Std. Error	t value	Pr(> t )
Luas areal.l1	0.4906	0.1687	2.909	0.00717 **
harga.l1	0.4727	0.2179	2.169	0.03904 *
volek.l1	-0.9431	0.5234	-1.802	0.08273 .
const	29408.0389	9991.1534	2.943	0.00660 **
trend	247.6390	542.9977	0.456	0.65199

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 12130 on 27 degrees of freedom

Multiple R-Squared: 0.7062, Adjusted R-squared: 0.6626

F-statistic: 16.22 on 4 and 27 DF, p-value: 6.954e-07

Estimation results for equation harga:

=====

harga = Luas areal.l1 + harga.l1 + volek.l1 + const + trend

	Estimate	Std. Error	t value	Pr(> t )
Luas areal.l1	-0.07974	0.08865	-0.900	0.3763
harga.l1	0.83875	0.11455	7.322	7.08e-08 ***
volek.l1	-0.09951	0.27510	-0.362	0.7204
const	741.74381	5251.47728	0.141	0.8887
trend	651.06672	285.40648	2.281	0.0306 *

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 6373 on 27 degrees of freedom

Tabel 25. Output VAR type p=1 type “cons”:

VAR Estimation Results:

=====  
 Endogenous variables: Luas areal, harga, volek

Deterministic variables: const

Sample size: 32

Log Likelihood: -980.406

Roots of the characteristic polynomial:

1.04 0.6263 0.07462

Call:

VAR(y = cengkeh[1:33, c(2, 3, 4)], p = 1, type = "const")

Estimation results for equation Luas areal:

=====  
 Luas areal = Luas areal.l1 + harga.l1 + volek.l1 + const

Estimate Std. Error t value Pr(>|t|)

Luas areal.l1 0.5171 0.1561 3.314 0.002551 \*\*

harga.l1 0.5457 0.1458 3.743 0.000832 \*\*\*

volek.l1 -0.8489 0.4740 -1.791 0.084134 .

const 29666.1619 9833.0190 3.017 0.005386 \*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 11950 on 28 degrees of freedom

Multiple R-Squared: 0.7039, Adjusted R-squared: 0.6722

F-statistic: 22.19 on 3 and 28 DF, p-value: 1.469e-07

Estimation results for equation harga:

=====  
 harga = Luas areal.l1 + harga.l1 + volek.l1 + const

Estimate Std. Error t value Pr(>|t|)

Luas areal.l1 -0.01005 0.08925 -0.113 0.911

harga.l1 1.03067 0.08337 12.363 7.35e-13 \*\*\*

volek.l1 0.14837 0.27104 0.547 0.588

const 1420.37388 5622.87326 0.253 0.802

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

## Pengujian Asumsi

Selanjutnya dilakukan pengujian asumsi  $p=1$  type both dan  $p=1$  type cons". untuk non autokorelasi, normalitas, dan homoskedastisitas pada sisaan model VAR  $p=1$  type "both" dan VAR  $p=1$  type "cons"

VAR  $p=1$  type both

- Pemeriksaan autokorelasi residual model menggunakan fungsi "serial.test" yang di dalamnya dilakukan pengujian Portmanteau-and Breusch-Godfrey test. Hasil pemeriksaan autokorelasi adalah

Portmanteau Test (asymptotic) data: Residuals of VAR object varhsheet.b1 Chi-squared = 103.69, df = 135, p-value = 0.9791. Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi.

- Pengujian Jarque-Bera tests untuk menguji kenormalan. JB-Test (multivariate) data: Residuals of VAR object varhsheet.b1

Chi-squared = 86.902, df = 6, p-value < 2.2e-16

\$Skewness

Skewness only (multivariate)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 23.992, df = 3, p-value = 2.508e-05

\$Kurtosis

Kurtosis only (multivariate)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 62.91, df = 3, p-value = 1.403e-13

0.2567 5.2827 9.6832 12.5374 17.0834 40.0469

hasil pengujian menunjukkan nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal.

- Pemeriksaan heteroskedastisitas model menggunakan fungsi "arch.test" yang di dalamnya dilakukan pengujian ARCH- LM tests .

ARCH (multivariate)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 162, df = 180, p-value = 0.8281

Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.

Var  $p=1$  type "cons"



- Pengujian Portmanteau-and Breusch-Godfrey test. Hasil pemeriksaan autokorelasi adalah data: Residuals of VAR object varhsheet1  
Chi-squared = 104.79, df = 135, p-value = 0.9747  
Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi.  
JB-Test (multivariate)  
Pengujian Jarque-Bera tests untuk menguji kenormalan.  
data: Residuals of VAR object varhsheet1  
Chi-squared = 184.92, df = 6, p-value < 2.2e-16  
\$Skewness  
    Skewness only (multivariate)  
    data: Residuals of VAR object varhsheet1  
    Chi-squared = 43.075, df = 3, p-value = 2.372e-09  
    \$Kurtosis  
    Kurtosis only (multivariate)  
    data: Residuals of VAR object varhsheet1  
    Chi-squared = 141.84, df = 3, p-value < 2.2e-16  
    hasil pengujian menunjukkan Nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal.
- Pemeriksaan heteroskedastisitas model menggunakan fungsi “arch.test” yang di dalamnya dilakukan pengujian ARCH- LM tests .  
ARCH (multivariate)  
data: Residuals of VAR object varhsheet1  
Chi-squared = 162, df = 180, p-value = 0.828  
Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.  
Selanjutnya dilakukan peramalan data, baik untuk data training maupun untuk data testing sekaligus dilakukan penghitungan MAPE. MAPE (Mean Absolute Percentage Error) merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan data hasil peramalan. Untuk menguji suatu model lebih baik dengan model yang lain, maka dilakukan pengujian model dengan membandingkan Nilai MAPE baik untuk data training maupun data testing.  
Untuk MAPE VAR p=1 type both  
Mape data testing

Min. 1st Qu. Median Mean 3rd Qu. Max.  
 2.441 6.838 9.246 **6.845** 12.111 13.586

Mape data training

Min. 1st Qu. Median Mean 3rd Qu. Max.  
 0.2567 5.2827 9.6832 **10.5374** 17.0834 40.0469

Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data actual dengan data hasil ramalan pada data testing adalah 8,84%

MAPE VAR p=1 type cons

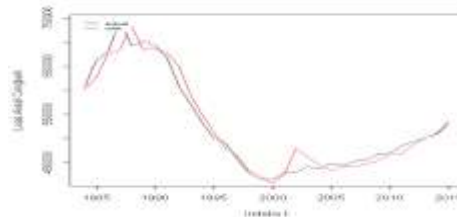
Mape data testing

Min. 1st Qu. Median Mean 3rd Qu. Max.  
 2.435 2.739 3.908 6.253 5.194 16.990

Mape data testing

Min. 1st Qu. Median Mean 3rd Qu. Max.  
 0.9324 5.1170 9.3525 12.7886 16.1274 39.5442

Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data actual dengan data hasil ramalan pada data testing adalah 6,25%. Jika dibandingkan nilai MAPEnya maka nilai MAPE terkecil adalah p=1 type both sehingga yang digunakan VAR p=1 type “both”  
 Kemudian dilakukan plot data ramalan dan data actual



Gambar 12. Plot data ramalan dan data actual

Kemudian dilakukan peramalan untuk tahun 2021-2025 dengan VAR ==1 type both

Tabel 26. Hasil Ramalan Tahun 2021-2021

Tahun	Produksi (Ton)	Pertumbuhan (%)
2021	564,122	
2022	535,853	-5.01
2023	500,641	-6.57
2024	465,051	-7.11
2025	433,139	-6.86
Rata-rata pertumbuhan (%)		-6.39
Mape Traini	12,62%	
Mape Testir	7,54%	

### Pemilihan Model Terbaik Estimasi Luas Areal Cengkeh

Salah satu dasar penentuan model terbaik dapat dilakukan dengan membandingkan nilai MAPE untuk data testing dan training yaitu dengan memilih nilai MAPE yang paling kecil, terutama untuk data testing. Selain MAPE yang terkecil, pola pergerakan ramalan juga harus diperhatikan. Dipilih plot yang paling berhimpit/bersesuaian dengan data asli/aktual atau dengan kata lain performa hasil ramalan seiring dengan data historisnya.

Tabel 27. Perbandingan MAPE Model Arima, F Transfer dan VAR

No	Uraian Model	Mape		Hasil Estimasi Luas Areal Cengkeh (Ha)					Pertumb.
		Testing	Training	2021	2022	2023	2024	2025	%
1	ARIMA (2,1,0)	2,49	5,74	578789	580494	582448	583794	585129	0,27
2	Fungsi Transfer	12,49	15,74	575465	575691	581180	599715	594765	0,84
3	VAR $\rho=1$ both	6,84	10,53	564122	535853	500641	465051	433139	-6,39

### KESIMPULAN

Dari 3 model yang diuji cobakan maka model yang terbaik adalah ARIMA (2,1,0). Model Arima ini ini menghasilkan ketelitian yang cukup tinggi yaitu MAPE untuk data training 5,74% dan MAPE untuk data testing 2,49%. MAPE untuk data testing dan data training paling kecil dibandingkan dengan model-model lainnya, sehingga model ARIMA (2,1,0) disimpulkan paling kecil kesalahan estimasinya atau model terbaik dibandingkan model fungsi transfer dan model VAR. Jika dilihat pertumbuhan selama 5 tahun yaitu tahun 2021-2025 sebesar 0,27%

### DAFTAR PUSTAKA

- Kementerian Pertanian. 2013. Pedoman Pelaksanaan Pengelolaan Data Komoditas Perkebunan (PDKP). Dirjen Perkebunan – Kementerian Pertanian
- Kementerian Pertanian. 2021. Statistik Perkebunan Indonesia 2019-2021 (Cengkeh). Jakarta: Direktorat Jenderal Perkebunan.



## MODEL ESTIMASI PRODUKSI KARET NASIONAL 2022 – 2025 MENGGUNAKAN MODEL ARIMA, FUNGSI TRANSFER DAN VAR (*Vector Auto Regressive*)

**Mohammad Chafid**

*Center for Agricultural Data and Information System-Ministry of Agriculture  
Jalan Harsono RM No. 3 Gd D Lt 4, Ragunan-Jakarta Selatan, Indonesia  
E-mail: mohammad.hafidz1@gmail.com*

### ABSTRAK

Komoditas karet merupakan salah satu komoditas strategis penghasil devisa negara karena sebagian besar produksi karet nasional untuk diekspor. Status angka statistik perkebunan terdiri dari Angka Tetap, Angka Sementara dan Angka Estimasi. Tujuan penulisan ini adalah mencari model alternatif lain untuk menyusun angka estimasi Produksi karet sehingga akurasi menjadi lebih baik yang ditandai dengan semakin kecilnya MAPE baik untuk data training maupun testing.

Model yang digunakan untuk menyusun angka estimasi produksi karet meliputi, Metode Arima, Metode Fungsi Transfer dengan peubah input harga karet dunia, Metode VAR (*Vector Autoregressive*) dengan variabel produksi, luas areal, harga karet sheet dalam negeri, harga karet dunia, volume ekspor karet dan volume impor karet. Sumber data yang digunakan untuk variabel produksi karet, harga karet nasional, volume ekspor dan impor karet (1970-2021) berasal dari Ditjen Perkebunan. Untuk variabel harga karet dunia (1971 - 2020) berasal dari World Bank. Run model menggunakan software RStudio.

Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1970 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan. Model ARIMA terbaik adalah ARIMA (2,1,0), menghasilkan MAPE untuk data training 4,74%, dan MAPE data testing 7,19%. Model ARIMA (1,2,3) juga menghasilkan MAPE yang cukup baik, yaitu MAPE training 4,15% dan MAPE testing 7,30%. Untuk model yang kedua dengan menggunakan Fungsi Transfer ARIMA (2,1,0) dengan variabel input harga karet dunia, menghasilkan MAPE untuk data training sebesar 3,67% dan MAPE data testing 4,26%. Untuk model Fungsi Transfer lainnya adalah ARIMA (0,1,1) menghasilkan MAPE data training sebesar 3,40% dan MAPE data testing 4,43%. Untuk model yang ketiga model VAR(2) type 'both' ada pengaruh konstanta dan trend, menghasilkan MAPE data training 1,10% dan data testing 6,02%. Model tentatif VAR adalah VAR(4) type 'both', menghasilkan MAPE data training 0,64% dan data testing 6,46%.

Berdasarkan perbandingan besarnya MAPE baik data testing maupun data training dan hasil estimasi produksi 5 tahun kedepan, maka model terbaik yang terpilih adalah model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet dunia karena menghasilkan akurasi yang paling tinggi sehingga MAPE rata-rata data testing sebesar 4,26%. Hasil estimasi produksi karet nasional untuk model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet dunia untuk tahun 2021 sebesar 2.919.246 ton dan tahun 2022 sebesar 2.962.994 ton, tahun 2023 sebesar 3.005.103 ton, tahun 2024 sebesar 3.045.418 ton, dan tahun 2025 sebesar 3.044.920 ton. Laju pertumbuhan estimasi produksi karet nasional selama 5 tahun kedepan (2021 – 2025) rata-rata 1,06% per tahun. Namun jika mempertimbangkan program yang telah dijalankan oleh direktorat teknis terkait untuk meningkatkan areal dan produksi karet maka ada optimisme estimasi produksi tahun 2021 sebesar 3.045.340 ton dan tahun 2022 sebesar 3.135.344 ton, yaitu berdasarkan hasil pemodelan dengan ARIMA (1,2,3).

**Kata Kunci :** *Produksi Karet, Arima, Fungsi Transfer, VAR (Vector Autoregressive)*

## PENDAHULUAN

Karet (*Hevea brasiliensis*) termasuk dalam genus *Hevea* dari familia *Euphorbiaceae*, yang merupakan pohon kayu tropis yang berasal dari hutan Amazon. Di dunia, setidaknya 2.500 spesies tanaman diakui dapat memproduksi lateks, tetapi *Hevea brasiliensis* saat ini merupakan satu-satunya sumber komersial produksi karet alam. Karet alam mewakili hampir separuh dari total produksi karet dunia karena sifat unik mekanik, seperti ketahanan sobek, dibandingkan dengan karet sintetis.

Data statistik perkebunan yang disajikan merupakan hasil sinkronisasi dan validasi yang dilaksanakan secara berjenjang dari tingkat kabupaten sampai tingkat nasional dengan mengacu pada Pedoman Pengelolaan Data Komoditas Perkebunan yang diterbitkan oleh Direktorat Jenderal Perkebunan. Secara periodik, data perkebunan disajikan dalam 3 status angka yaitu Angka Tetap (ATAP), Angka Sementara (ASEM) dan Angka Estimasi (AESTI). Angka Tetap merupakan angka hasil rekapitulasi dari pelaporan yang sudah tetap, sehingga tidak dilakukan estimasi, sedangkan untuk penentuan Angka Sementara dan Estimasi perlu dilakukan estimasi dengan metode estimasi yang paling relevan dan tepat (PDKP, 2013). Data Angka Tetap (ATAP) merupakan data 1 tahun yang lalu ( $n-1$ ), Angka Sementara (ASEM) merupakan data tahun berjalan ( $n$ ), dan Angka Estimasi (AESTI) merupakan data tahun satu tahun kedepan ( $n+1$ ).

Metode estimasi yang digunakan adalah Metode *Exponential Smoothing* (Peramalan Pemulusan Eksponensial) yang merupakan salah satu kategori metode time series yang menggunakan pembobotan data masa lalu secara eksponensial. Dalam kategori ini terdapat dua metode yang umum dipakai yaitu metode Pemulusan Eksponensial Tunggal (*Single Exponential Smoothing*) dan metode Pemulusan Eksponensial Ganda (*Double Exponential Smoothing*). Pemilihan model *Single Exponential Smoothing* atau *Double Exponential Smoothing* harus mempertimbangkan nilai *Mean Absolute Percentage Error (MAPE)* serta kerealistisan hasil estimasi bila dibandingkan dengan series data sebelumnya. MAPE adalah pengukur tingkat akurasi (ketepatan) nilai dugaan yang dihasilkan oleh model dalam bentuk presentase. Model yang mempunyai nilai MAPE lebih kecil dianggap sebagai model yang lebih baik. Keunggulan dari metode estimasi ini adalah dapat digunakan untuk meramalkan data yang berisi trend atau pola musiman. Namun metode estimasi ini juga memerlukan keahlian khusus dalam menginterpretasikan hasil estimasi yang diperoleh (PDKP, 2013).

Metode untuk menghasilkan angka estimasi (AESTI) yang diliris oleh Ditjen. Perkebunan pada waktu tahun berjalan ( $n$ ) perlu dikaji kembali, agar didapatkan metode yang lebih akurat, lebih objektif dan lebih baik secara statistik dari metode sebelumnya, sehingga tingkat kesalahannya lebih kecil. Guna merumuskan kebijakan produksi, pasokan dan distribusi komoditas karet sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Metode ramalan produksi karet dapat digunakan dengan menggunakan beberapa metode statistik yang biasa digunakan untuk melakukan peramalan.

Pada analisis ini akan dikaji metode ARIMA, Fungsi Transfer dan VAR dalam melakukan pemodelan dan peramalan produksi karet nasional.

Oleh karenanya, tujuan dari disusunnya analisis ini adalah:

1. Melakukan analisis dan peramalan data produksi karet nasional menggunakan model ARIMA, Fungsi Transfer dan VAR.
- m. Menentukan metode terbaik dalam meramal data produksi karet nasional.

## METODOLOGI

Data yang digunakan dalam melakukan pemodelan produksi karet nasional adalah data series produksi karet nasional tahun 1970 - 2020. Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1970 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan.

Peubah yang diasumsikan mempengaruhi besaran produksi karet untuk Model Fungsi Transfer adalah harga karet sheet lokal. Harga karet lokal diasumsikan akan berpengaruh terhadap produksi, karena jika terjadi kenaikan harga karet maka produksi karet cenderung akan terus ditingkatkan, kondisi sebaliknya maka harga karet cenderung stagnan, pertumbuhan produksi rendah, atau cenderung turun.

Peubah yang berpengaruh terhadap produksi untuk pemodelan VAR adalah produksi, harga karet domestik, harga karet dunia, volume ekspor dan impor karet. Harga karet dunia dan harga karet domestik diduga kuat berpengaruh pada tingkat produksi. Untuk volume ekspor dan impor juga berpengaruh terhadap produksi, dimana semakin tinggi permintaan ekspor karet maka produksi cenderung akan semakin tinggi.

Peramalan data produksi karet, dilakukan uji coba beberapa metode yakni mengkaji metode ARIMA, Regresi, Fungsi Transfer dan VAR (Vector Autoregression), menggunakan software *R Studio*.

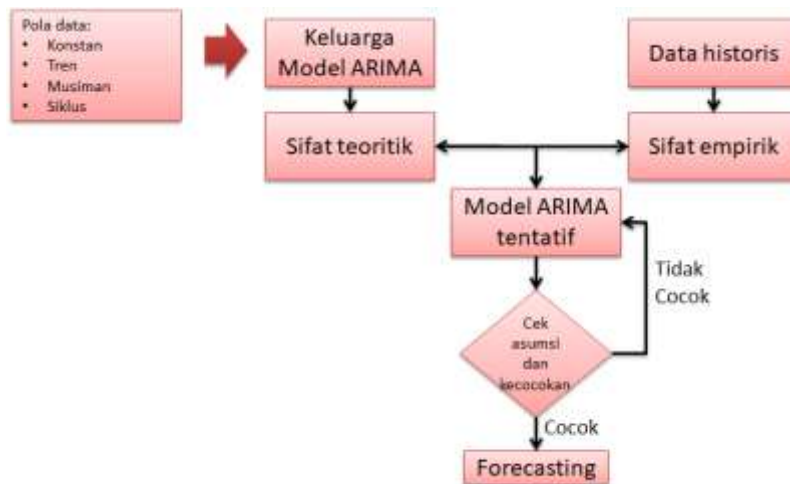
### 1. ARIMA (*Autoregressive Integrated Moving Average*)

Metode Autoregressive Integrated Moving Average (ARIMA) atau biasa disebut juga sebagai metode Box-Jenkins merupakan metode yang secara intensif dikembangkan oleh George Box dan Gwilym Jenkins pada tahun 1970 (Iriawan, 2006).

ARIMA sering juga disebut metode runtun waktu Box-Jenkins. ARIMA sangat baik ketepatannya untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatan peramalannya kurang baik. Biasanya akan cenderung flat (mendatar/konstan) untuk periode yang cukup panjang.

Model *Autoregresif Integrated Moving Average* (ARIMA) adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari deret waktu (time series) secara statistik berhubungan satu sama lain (dependent).

Model ARIMA terdiri dari tiga langkah dasar, yaitu tahap identifikasi, tahap penaksiran dan pengujian, dan pemeriksaan diagnostik. Selanjutnya model ARIMA dapat digunakan untuk melakukan peramalan jika model yang diperoleh memadai.



Gambar 1. Prosedur Peramalan Model Arima (Box- Jenkins)

Hal yang perlu diperhatikan adalah bahwa kebanyakan deret berkala bersifat stasioner. Stasioner berarti tidak terdapat pertumbuhan atau penurunan pada data. Data secara kasarnya harus horizontal sepanjang sumbu waktu. Dengan kata lain, fluktuasi data berada di sekitar suatu nilai rata-rata yang konstan, tidak tergantung pada waktu dan varians dari fluktuasi tersebut pada pokoknya tetap konstan setiap waktu. Suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melakukan differencing. Yang dimaksud dengan differencing adalah menghitung perubahan atau selisih nilai observasi. Nilai selisih yang diperoleh dicek lagi apakah stasioner atau tidak. Jika belum stasioner maka dilakukan differencing lagi. Jika varians tidak stasioner, maka dilakukan transformasi logaritma.

Model Box-Jenkins (ARIMA) dibagi kedalam 3 kelompok, yaitu: model autoregressive (AR), moving average (MA), dan model campuran ARIMA (autoregressive moving average) yang mempunyai karakteristik dari dua model pertama. ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampau. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

- $Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t
- $Y_{t-p}$  = data *time series* pada kurun waktu ke (t-p)
- $\mu$  = suatu konstanta
- $\theta_1, \theta_2, \dots, \theta_p$  = parameter-parameter model
- $\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

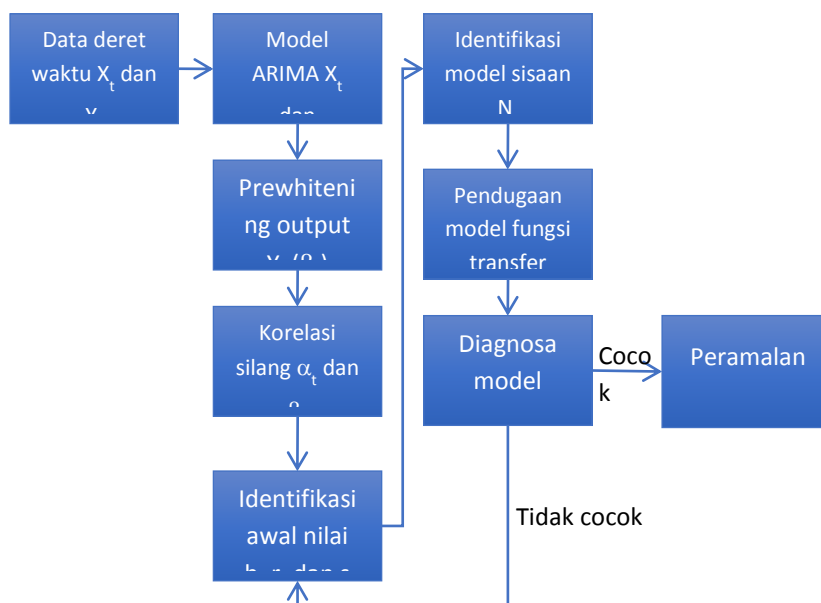
Model dikatakan baik jika nilai error bersifat random, artinya sudah tidak mempunyai pola tertentu lagi. Dengan kata lain model yang diperoleh dapat menangkap dengan baik pola data yang ada. Untuk melihat kerandaman nilai error dilakukan pengujian terhadap nilai koefisien autokorelasi dari error, dengan menggunakan salah satu dari dua statistik berikut, yaitu Uji Q-Box and Pierce dan uji Ljung-Box.



### m. Model Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke- $t$ , tetapi juga pada waktu  $t+1, t+2, \dots, t+k$ . Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara peubah *input* dan peubah *output*.

Tujuan pemodelan fungsi transfer adalah untuk menetapkan model yang sederhana, yang menghubungkan deret output ( $Y_i$ ) dengan deret input ( $X_i$ ) dan gangguan/noise( $n_i$ ). Wei (1994) juga menjelaskan bahwa di dalam fungsi transfer terdapat rangkaian output yang mungkin dipengaruhi oleh rangkaian multiple input. Pada kasus single input peubah, dapat menggunakan metode korelasi silang yang dianjurkan oleh Box and Jenkins (1976). Teknik ini juga dapat digunakan ketika terdapat single input peubah yang lebih dari satu selama antar variable input tidak berkorelasi silang. Jika beberapa atau semua peubah input berkorelasi silang maka teknik *prewhitening* atau metode korelasi silang tidak dapat digunakan secara langsung. Alasan utama bagi perlunya suatu perencanaan atau peramalan adalah adanya tenggang waktu pengambilan keputusan yang dapat berkisar dari beberapa hari atau sampai beberapa tahun. Pada analisis fungsi transfer untuk peramalan deret berkala univariate, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (noise). Deret input mempengaruhi deret output melalui sebuah fungsi transfer yang mendistribusikan pengaruhnya secara dinamis melalui beberapa periode waktu yang akan datang dengan persentase tertentu yang disebut sebagai bobot respons impuls atau bobot fungsi transfer.



Gambar 2. Langkah-langkah melakukan pemodelan Fungsi Transfer

Model umum Fungsi Transfer:

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t \quad \dots\dots\dots(7)$$

Dimana:

- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$
- $q \rightarrow$  ordo MA bagi noise  $N_t$

**n. Model Vector Autoregression (VAR)**

*Vector Autoregression* (VAR) dikemukakan pertama kali oleh Christopher Sims (1980). Sims mengembangkan model ekonometri dengan mengabaikan pengujian asumsi secara apriori. VAR dikembangkan oleh Sims sebagai kritik atas metode simultan. Jumlah peubah yang besar dan klasifikasi endogen dan eksogen pada metode simultan merupakan dasar dari kritik tersebut. Penggunaan pendekatan struktural atas pemodelan persamaan simultan biasanya menerapkan teori ekonomi di dalam usahanya untuk mendeskripsikan hubungan antar peubah yang ingin diuji. Akan tetapi sering ditemukan bahwa teori ekonomi saja ternyata tidak cukup kaya di dalam menyediakan spesifikasi yang ketat dan tepat atas hubungan dinamis antar peubah (Gujarati, 2010). Model VAR merupakan jalan keluar atas permasalahan ini melalui pendekatan non-struktural atau tidak mendasarkan pada teori ekonomi tertentu (*atheoretical*). Metode VAR memperlakukan seluruh peubah secara simetris tanpa memperlakukan peubah dependen dan independen atau dengan kata lain model ini memperlakukan seluruh peubah sebagai peubah endogen, karena pada kenyataannya suatu peubah dapat bertindak sebagai endogen atau dependen dari suatu peubah yang lainnya (Gujarati, 2010).

Kelebihan dalam penggunaan metode VAR (Gujarati, 2010):

- i. Kemudahan dalam penggunaan, tidak perlu mengkhawatirkan tentang penentuan peubah endogen dan peubah eksogen.
- j. Kemudahan dalam estimasi, metode *Ordinary Least Square* (OLS) dapat diaplikasikan pada tiap persamaan secara terpisah.
- k. *Forecast* atau peramalan yang dihasilkan pada beberapa kasus ditemukan lebih baik daripada yang dihasilkan oleh model persamaan simultan yang kompleks.
- l. Perangkat estimasi yang digunakan adalah *Impulse Respon Function* (IRF) untuk melacak respon dari peubah dependen dalam sistem VAR terhadap *shock* dari *error term* dan *Variance Decomposition* yang memberikan informasi mengenai pentingnya masing-masing *error term* dalam mempengaruhi peubah-peubah dalam VAR.

Di sisi lain, terdapat beberapa kritik terhadap model VAR menyangkut permasalahan berikut (Gujarati, 2010) :

- 11) Model VAR merupakan model yang *atheoretic* atau tidak berdasarkan teori, hal ini tidak seperti pada persamaan simultan. Pada persamaan simultan, pemilihan peubah yang akan dimasukkan dalam persamaan memegang peranan penting dalam mengidentifikasi model.
- 12) Pada model VAR penekanannya terletak pada *forecasting* atau peramalan sehingga model ini kurang cocok digunakan dalam menganalisis kebijakan.
- 13) Permasalahan yang besar dalam model VAR adalah pada pemilihan *lag length* atau panjang lag yang tepat. Karena semakin panjang lag, maka akan menambah jumlah parameter yang akan bermasalah pada *degrees of freedom*.
- 14) Peubah yang tergabung pada model VAR harus stasioner. Apabila tidak stasioner, perlu dilakukan transformasi bentuk data, misalnya melalui *first difference*.
- 15) Sering ditemui kesulitan dalam menginterpretasi tiap koefisien pada estimasi model VAR, sehingga sebagian besar peneliti melakukan interpretasi pada estimasi fungsi *impulse respon*.

### m. Kriteria Pemilihan Model Terbaik

Untuk menguji kebaikan suatu model ada beberapa kriteria yang digunakan. Pada buku pedoman teknis ini kriteria yang digunakan adalah MAPE. Model time series/arima, model regresi, model fungsi transfer, dan model VAR dapat digunakan untuk melakukan estimasi variabel untuk beberapa tahun ke depan. Untuk model *time series* baik analisis ARIMA, model regresi, model fungsi transfer, dan model VAR, ukuran kelayakan model berdasarkan nilai kesalahan dengan menggunakan statistik MAPE (*mean absolute percentage error*) atau kesalahan persentase absolut rata-rata yang diformulasikan sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right| \cdot 100$$

Dimana :  $X_t$  adalah data aktual

$F_t$  adalah nilai ramalan.

Semakin kecil nilai MAPE maka model yang diperoleh semakin baik, karena makin mendekati nilai aktual.

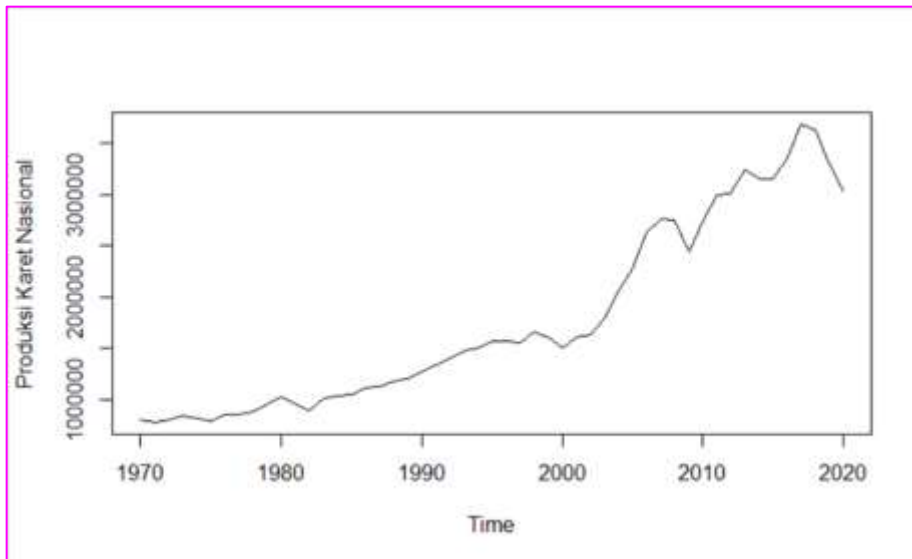
### f. Sumber Data

Data yang digunakan untuk analisis ini bersumber dari Ditjen. Perkebunan, Kementerian Pertanian, meliputi data : produksi karet dalam bentuk karet kering satuan ton tahun 1970 – 2020, produksi karet satuan hektar tahun 1970 – 2020, harga riil karet sheet dalam satuan Rp/kg tahun 1971 – 2020, volume ekspor dan impor karet nasional satuan ton tahun 1970 – 2020. Disamping itu juga ada data harga karet dunia yang bersumber dari World Bank, tahun 1971 – 2020.

## HASIL DAN PEMBAHASAN

### E. Model ARIMA

Eksplorasi data produksi karet nasional dalam bentuk karet kering berupa data tahunan dari tahun 1970 sampai 2020, seperti yang terlihat pada Gambar 3. Pada Gambar 3 terlihat produksi karet pada tahun 1970 sebesar 802 ribu ton, pada awalnya produksi karet terus naik secara landai sehingga pada tahun 1980 produksi karet mencapai 1,02 juta ton. Setelah tahun 1980 pertumbuhan produksi cenderung lebih cepat, pertumbuhan produksi karet tahun 1980 – 2000 rata-rata naik sebesar 2,06%/tahun. Pada tahun 1999 sampai 2000 produksi karet nasional cenderung terus turun, tetapi sejak tahun 2005 sampai 2020 produksi kembali meningkat. Jika tahun 2005 produksi karet sebesar 2,27 juta ton, maka tahun 2020 produksi karet sudah mencapai 3,03 juta ton, atau rata-rata pertumbuhan 2005 – 2020 mencapai 2,45%/tahun atau lebih tinggi pertumbuhannya dibandingkan periode sebelumnya. Berdasarkan Gambar 3 juga bisa terlihat bahwa data cenderung terus meningkat, sehingga data belum stasioner karena masih mengalami perubahan seiring perubahan waktu.



Gambar 3. Perkembangan Produksi Karet Nasional Tahun 1970 - 2020

Dalam melakukan pemodelan produksi karet menggunakan model Autoregressive Integrated Average (ARIMA), data yang digunakan adalah periode tahun 1970 sampai 2020. Periode data tersebut kemudian dipisahkan menjadi data set training dan testing. Perlunya pemisahan data training dan testing adalah untuk menguji tingkat akurasi dalam melakukan peramalan. Panjang series data pada data set training adalah tahun 1970 sampai 2014, sementara dataset testing adalah periode 2015 sampai 2020 (6 titik). Dataset training digunakan untuk melakukan penyusunan model, sementara dataset testing digunakan untuk validasi model.

Uji kestasioneran data seperti yang disyaratkan apabila melakukan pemodelan ARIMA dilakukan secara visual menggunakan hasil plot data maupun uji formal statistik. Gambar 3 menunjukkan tidak ada fluktuasi yang muncul secara regular setiap bulannya, sehingga dapat

disimpulkan bahwa data belum stasioner atau nilai rata-rata dan varian dari data time series karet mengalami perubahan secara stokastik sepanjang waktu atau sebagian ahli menyatakan rata-rata dan variannya belum konstan (Narchrowi dan Hardius Usman, 2006).

Tabel 1. Hasil Uji Augmented Dickey-Fuller Produksi Karet

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-309880  -59383  -6717   41199  348971

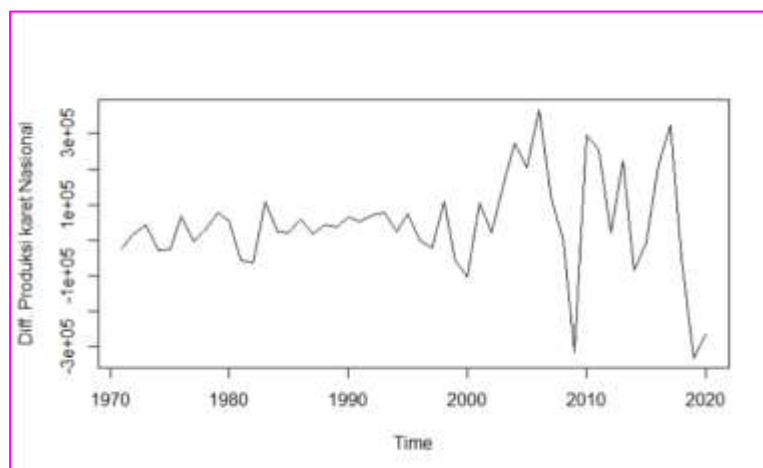
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6.131e+04  4.118e+04   1.489   0.1435
z.lag.1     -1.598e-01  6.542e-02  -2.442   0.0186 *
tt           9.614e+03  4.109e+03   2.340   0.0238 *
z.diff.lag   3.516e-01  1.454e-01   2.419   0.0197 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 129800 on 45 degrees of freedom
Multiple R-squared:  0.1877, Adjusted R-squared:  0.1335
F-statistic: 3.466 on 3 and 45 DF,  p-value: 0.02382

Value of test-statistic is: -2.4424 2.7928 2.9867

Critical values for test statistics:
      1pct   5pct  10pct
tau3  -4.04  -3.45  -3.15
phi2   6.50   4.88   4.16
phi3   8.73   6.49   5.47
```

Hal ini diperkuat oleh hasil uji formal statistik yaitu dengan uji Augmented Dickey-Fuller yang mengindikasikan bahwa data produksi karet adalah belum stasioner, terlihat dari hasil uji tes statistik sebesar  $-2,44$  sementara nilai kritis pada tingkat kepercayaan  $95\% = -3,45$  (nilai tau3) atau lebih kecil dari nilai uji statistik sehingga  $H_0$  diterima, atau data produksi karet belum stasioner. Oleh karena itu, selanjutnya data produksi karet dilakukan pembedaan (differencing) satu kali. Hasil plot setelah dilakukan differencing satu kali seperti pada Gambar 4.



Gambar 4. Data Produksi Karet Nasional Setelah Differencing 1

Hasil plot produksi karet setelah dilakukan differencing 1, menunjukkan bahwa data sudah terlihat stasioner untuk rata-rata. Hal ini juga didukung dengan uji Augmented Dickey-Fuller yang mengindikasikan bahwa data produksi karet setelah differencing 1 sudah stasioner, terlihat dari hasil uji tes statistik sebesar  $-4,04$  sementara nilai kritis pada tingkat kepercayaan  $95\% = -$

1,95 (tau1) dan tingkat kepercayaan 99% = -2,62 (tau1) atau lebih besar dari nilai uji statistik sehingga sehingga  $H_0$  ditolak, atau data produksi karet setelah differencing 1 sudah stationer.

Tabel 2. Hasil Uji Augmented Dickey-Fuller Produksi Karet Differencing 1

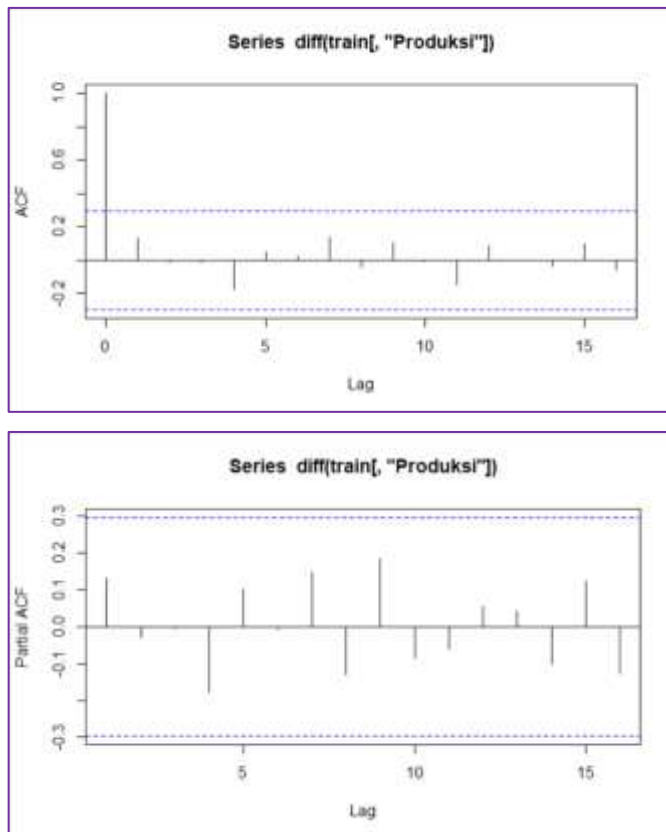
```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
    Min       1Q   Median       3Q      Max
-296432 -32611  28843  77721  426302
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1    -0.7260     0.1797  -4.040 0.000201 ***
z.diff.lag  0.1461     0.1568   0.932 0.356303
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 138500 on 46 degrees of freedom
Multiple R-squared:  0.3071, Adjusted R-squared:  0.277
F-statistic: 10.19 on 2 and 46 DF,  p-value: 0.0002164

Value of test-statistic is: -4.0404

Critical values for test statistics:
      1pct      5pct     10pct
tau1 -2.62 -1.95 -1.61
```

Identifikasi ordo AR dan MA untuk data produksi karet sebelum ada differencing menunjukkan pola ACF *cut off* pada lag ke-0, sementara pola PACF tidak ada yang signifikan. Hal ini menunjukkan bahwa benar data karet sudah signifikan, seperti pola ACF dan PACF lebih jelas terlihat pada Gambar 5.



Gambar 5. Plot ACF dan PACF Produksi Karet Differencing 1

Pada Gambar 5 menunjukkan bahwa baik plot ACF maupun plot ACF setelah *Differencing* 1 keduanya sudah ada *cut off*. Identifikasi ordo AR dan MA seperti tersaji pada Gambar 5

menunjukkan pola ACF *cut off* setelah lag 0, sementara pola PACF menunjukkan tidak ada yang signifikan, sehingga model tentatif belum dapat ditentukan.

Tabel 3. Model Arima Tentatif Berdasarkan Automodel

```
Series: train[, "Produksi"]
ARIMA(0,1,0) with drift

Coefficients:
    drift
    53432.73
s.e.      17346.15

sigma^2 estimated as 1.355e+10:  log likelihood=-575.17
AIC=1154.35  AICC=1154.64  BIC=1157.92

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 16.63806 113773.3 78017 -1.134623 4.715679 0.877731 0.1313069
```

Pengamatan secara visual pada plot ACF dan PACF sulit menentukan orde ARIMA, setelah dilakukan run model dengan menggunakan *auto arima* maka orde ARIMA yang disarankan adalah ARIMA (0,1,0), artinya model ARIMA tentatif terbaik untuk melakukan estimasi produksi karet nasional adalah untuk orde AR nilai  $p=0$ , untuk orde MA nilai  $q=0$ , dan difference  $d=1$ . Berdasarkan Tabel 3 dengan menggunakan ARIMA (0,1,0) maka untuk data training, akan menghasilkan MAPE = 4,71% artinya data berdasarkan model arima akan menyimpang hasil estimasi rata-rata sekitar -4,71% sampai +4,71% dari data aktual.

Disamping metode pemilihan model Arima berdasarkan automodel, digunakan juga metode lain untuk mendapatkan orde ARIMA terbaik, yaitu dengan metode *Arima selection*. Pada metode ini akan dikeluarkan beberapa model terbaik. Setelah dilakukan pemilihan model metode terbaik tetap pada *differencing 1*. Model tentatif pertama menurut metode ini adalah ARIMA (1,1,0) menghasilkan nilai sbc paling kecil yaitu sebesar 1051,91, model tentative kedua ARIMA (2,1,0) dan seterusnya seperti pada Tabel 4.

Tabel 4. Model Arima Tentatif Berdasarkan Arima Selection Differencing 1

	p	q	sbc
[1,]	1	0	1051.917
[2,]	2	0	1058.030
[3,]	3	0	1063.379
[4,]	4	0	1068.693
[5,]	5	0	1072.135
[6,]	1	1	1074.014
[7,]	2	1	1079.598
[8,]	1	2	1079.616
[9,]	3	1	1083.418
[10,]	2	2	1085.123

Metode *Arima selection* juga digunakan untuk mencari model Arima terbaik pada Differencing 2. . Pada metode ini akan dikeluarkan beberapa model tentatif terbaik. Model tentatif pertama menurut metode ini adalah ARIMA (1,2,0) menghasilkan nilai sbc paling kecil yaitu sebesar 1024,01, model tentative kedua ARIMA (2,2,0) dan seterusnya seperti pada Tabel 5.

Tabel 5. Model Arima Tentatif Berdasarkan Arima Selection Differencing 2

	p	q	sbk
[1,]	1	0	1024.013
[2,]	2	0	1024.351
[3,]	4	0	1026.749
[4,]	3	3	1026.818
[5,]	0	0	1026.864
[6,]	3	0	1028.705
[7,]	5	0	1029.974
[8,]	4	3	1030.654
[9,]	1	3	1032.073
[10,]	3	4	1032.154

Untuk menghasilkan model tentative terbaik dari sepuluh kombinasi order ARIMA pada differncing 1, dan sepuluh kombinasi order ARIMA pada differncing 2. Untuk memilih model tentative yang terbaik harus diperbandingkan koefisien MA dan AR apakah signifikan atau tidak. Disamping itu dilakukan pengujian MAPE untuk data training dan data testing. Model ARIMA terbaik adalah jika MAPE data training dan data testing terkecil. Hasil pengolahan perbandingan MAPE dan signifikansi koefisien seperti terlihat pada Tabel 6.

Tabel 6. Pengujian Signifikansi Koefisien dan MAPE untuk Model ARIMA

Model	Signifikansi	MAPE Training	MAPE Testing
<b>Differencing 1</b>			
<b>ARIMA (1,1,0)</b>	ar1 signifikan	4.75	7.53
<b>ARIMA (2,1,0)</b>	ar1 signifikan	4.74	7.19
<b>ARIMA (0,1,1)</b>	ma1 signifikan	4.76	7.71
<b>Differencing 2</b>			
<b>ARIMA (3,2,3)</b>	ar1 ma1 ma2 ma3 signifikan	4.11	7.57
<b>ARIMA (4,2,3)</b>	ar1 ma2 signifikan	4.14	7.86
<b>ARIMA (1,2,3)</b>	ar1 ma2 signifikan	4.15	7.29
<b>ARIMA (3,2,4)</b>	ar1 ar2 ar3 ma1 ma3 ma4 signifikan	4.25	7.25

Berdasarkan hasil pengujian pada Tabel 6 maka model tentative terbaik untuk differencing 1 peramalan Produksi Karet adalah ARIMA (2,1,0) dengan komponen ar1 signifikan, komponen ar2 tidak signifikan. Model ARIMA (1,1,0) dan ARIMA (0,1,1) meskipun data training relative sama, tetapi data testing lebih besar dari ARIMA (2,1,0) sehingga tidak terpilih. Hasil pengujian keakuratan model dalam melakukan peramalan, ARIMA (2,1,0) menghasilkan MAPE data training terkecil yaitu sebesar 4,74% dan MAPE data testing sebesar 7,19%.

Untuk model tentative pada Differencing 2 yang terbaik adalah ARIMA (1,2,3) karena menghasilkan komponen ar1 dan ma1 yang signifikan, serta hasil pengujian MAPE menghasilkan MAPE data training 4,15 dan data testing 7,29 seperti terlihat pada Tabel 6.

### Model ARIMA (2,1,0)

Selanjutnya dilakukan pengujian model tentative ARIMA (2,1,0) apakah koefisien sudah signifikan dan bagaimana perbandingan data training dan data testing. Untuk model ARIMA (2,1,0) koefisien ar1 sebesar 0,267 dan koefisien ini signifikan pada taraf alpha 10% dan koefisien ar2 = 0,079 tidak signifikan. Sehingga model ARIMA (2,1,0) layak digunakan (Tabel 7).



Tabel 7. Uji Koefisien Model Arima (2,1,0)

```
Call:
arima(x = train[, "Produksi"], order = c(2, 1, 0))

Coefficients:
      ar1      ar2
 0.2674  0.0797
s.e. 0.1511  0.1554

sigma^2 estimated as 1.456e+10:  log likelihood = -577.32,  aic = 1160.65

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  0.26737   0.15114   1.7690  0.0769 .
ar2  0.07966   0.15537   0.5127  0.6082
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tabel 8. Perbandingan MAPE untuk ARIMA (2,1,0)

```
Time Series:
Start = 46
End = 51
Frequency = 1
[1] 3148599 3140662 3138174 3136877 3136332 3136082
> accurasy(ramalan_arima, test[, "Produksi"])
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 33912.5 119334.2 81306.31 1.893713 4.741353 0.9147375 -0.1005921
Test set    219360.3 321917.5 253338.82 6.074193 7.191676 2.8501908      NA
```

Setelah dilakukan pengujian koefisien model untuk ARIMA (2,1,0) ternyata komponen ar1 signifikan, dan komponen ar2 tidak signifikan, maka dilakukan pengujian MAPE untuk data training dan data testing. Hasil menunjukkan jika menggunakan model ARIMA (2,1,0) akan menghasilkan data training sebesar 4,74%. Setelah dilakukan pengujian dengan cara meramal 6 tahun kedepan yaitu tahun 2015 – 2020, maka hasil ramalan atau data testing menghasilkan MAPE 7,19%. Hal ini menunjukkan bahwa model ARIMA (2,1,0) jika digunakan untuk peramalan maka rata-rata hasil ramalan hanya menyimpang dengan nilai mutlak 7,19%.

Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan Ljung-Box. Hasil pengujian nilai p-value pada lag 5 sampai dengan lag 30 tidak ada yang signifikan yang ditunjukkan dengan nilai p-value tidak ada yang lebih kecil dari nilai 0.05 (kepercayaan 95%), sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (2,1,0) layak digunakan.

Tabel 9. Uji Sisaan Ljung Box Model Arima (2,1,0)

```
lags statistic df  p-value
 5  2.559639  5  0.7674877
10  4.345178 10  0.9304349
15  6.816083 15  0.9625537
20  7.602043 20  0.9941903
25  8.457486 25  0.9991764
30 13.895486 30  0.9946440
```

Selanjutnya dilakukan pengepasan model untuk seluruh data. Untuk Model ARIMA (2,1,0) koefisien ar1 sebesar 0,411 dan koefisien ar2= -0,140. Jika melakukan run model ARIMA (2,1,0) untuk seluruh data yaitu dari tahun 1970 – 2020 maka akan dihasilkan MAPE sebesar 4,82%. Hal ini menunjukkan bahwa antara data estimasi dengan data actual akan berbeda rata-rata berkisar antara -4,82% sampai +4,82%. Untuk metode estimasi dengan bias masih dibawah 5% dianggap masih cukup baik dan akurat.

Tabel 10. Model Arima (2,1,0) untuk Seluruh Data

```
Series: karet[, "Produksi"]
ARIMA(2,1,0)
Coefficients:
      ar1      ar2
    0.4115  -0.1405
s.e.  0.1406   0.1471

sigma^2 estimated as 1.84e+10: log likelihood=-660.91
AIC=1327.81  AICC=1328.33  BIC=1333.55

Training set error measures:
Training set  ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
31474.45  131596.9  92216.15  1.772206  4.828593  0.9046419  -0.05350715
```

Tabel 11. Ouput Peramalan Model Arima (2,1,0) untuk Produksi Karet

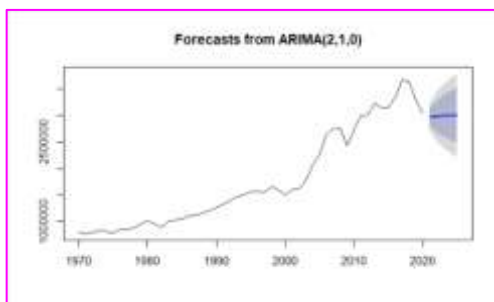
Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	2974904	2801066	3148743	2709041	3240768
2022	2986303	2685595	3287011	2526410	3446196
2023	2999765	2608464	3391065	2401322	3598207
2024	3003702	2543407	3463998	2299741	3707664
2025	3003432	2485060	3521804	2210650	3796213

Dengan menggunakan model ARIMA (2,1,0) menghasilkan angka estimasi produksi karet untuk 5 tahun ke depan. Hasil Estimasi dengan model ARIMA ini pada tahun 2021 produksi karet nasional sebesar 2,974 juta ton. Pada tahun 2022 produksi karet diestimasi akan naik sebesar 0,38% menjadi 2,986 juta ton. Pada tahun 2023 sampai 2024 juga menunjukkan produksi karet nasional mengalami kenaikan secara perlahan masing -masing sebesar 0,45% dan 0,13%. Sementara pada tahun 2025 produksi karet cenderung tetap atau sama dengan tahun 2024. Jika dibandingkan pertumbuhan produksi karet selama 5 tahun terakhir (tahun 2015 -2020) dengan menggunakan Angka Tetap rata-rata produksi turun sebesar 2,20% per tahun, sementara hasil estimasi lima tahun kedepan (2021 – 2025) rata-rata pertumbuhan juga turun hanya 0,22% per tahun atau lebih rendah dari data historisnya. Hal ini terjadi karena beberapa tahun terakhir harga karet dunia terus turun, sehingga perluasan areal sangat kecil pertumbuhannya, jika harga karet dunia meningkat maka pertumbuhan produksi karet diduga akan lebih besar.

Tabel 12. Hasil Estimasi Produksi Karet dengan Model ARIMA (2,1,0)

Tahun	Estimasi Produksi (Ton)	Pertumbuhan (%)
2020	3.037.348	
2021	2.974.904	(2,06)
2022	2.986.303	0,38
2023	2.999.765	0,45
2024	3.003.702	0,13
2025	3.003.432	(0,01)
Rata-rata pertumbuhan 2021 - 2025		(0,22)

Pada Gambar 6 terlihat pada tahun pertama hasil estimasi atau tahun 2021, produksi karet turun 2,06%. Namun pada tahun 2022 – 2024 produksi karet cenderung naik secara perlahan, dan pada tahun 2025 produksi karet cenderung konstan.



Gambar 6. Hasil Estimasi Produksi Karet Tahun 2021 – 2025 Model Arima (2,1,0)

## Model ARIMA (1,2,3)

Model tentative yang kedua adalah ARIMA (1,2,3). Selanjutnya dilakukan pengujian model tentative ARIMA (1,2,3) apakah koefisien sudah signifikan dan bagaimana perbandingan data training dan data testing. Untuk model ARIMA (1,2,3) koefisien ar1 sebesar -0,77 dan koefisien ma2 sebesar -0,895, keduanya signifikan pada taraf alpha 0,1%. Sementara koefisien ma1 dan ma3 tidak signifikan. Model ARIMA (1,2,3) layak layak dipertimbangkan digunakan (Tabel 13).

Tabel 13. Uji Koefisien Model Arima (1,2,3)

```
Call:
arima(x = train[, "Produksi"], order = c(1, 2, 3))

Coefficients:
      ar1      ma1      ma2      ma3
-0.7759  0.0390 -0.8952  0.0085
s.e.    0.1952  0.2731  0.1496  0.1853

sigma^2 estimated as 1.258e+10:  log likelihood = -562.33,  aic = 1134.66
z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1 -0.7758723  0.1952438 -3.9739 7.072e-05 ***
ma1  0.0389838  0.2730568  0.1428  0.8865
ma2 -0.8952405  0.1496353 -5.9828 2.193e-09 ***
ma3  0.0085203  0.1853471  0.0460  0.9633
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tabel 14. Perbandingan MAPE untuk ARIMA (1,2,3)

```
Time Series:
Start = 46
End = 51
Frequency = 1
[1] 3243645 3316490 3401884 3477542 3560754 3638104
> accuracy(ramalan_arima, test[, "Produksi"])
      Training set      Test set
ME      18456.47      -80922.02
RMSE    109655.7      300133.4
MAE      70001.3      238528.7
MPE      1.059132     -2.957646
MAPE     4.14669      7.29509
MASE     0.7875503    2.6835699
ACF1     -0.05083289   NA
```

Setelah dilakukan pengujian koefisien model untuk ARIMA (1,2,3) ternyata komponen ar1 signifikan, dan komponen ma2 signifikan, maka dilakukan pengujian MAPE untuk data training dan data testing. Hasil menunjukkan jika menggunakan model ARIMA (1,2,3) akan menghasilkan data training sebesar 4,14%. Setelah dilakukan pengujian dengan cara meramal 6 tahun kedepan yaitu tahun 2015 – 2020, maka hasil ramalan atau data testing menghasilkan MAPE 7,29%. Hal ini menunjukkan bahwa model ARIMA (1,2,3) jika digunakan untuk peramalan maka rata-rata hasil ramalan hanya menyimpang dengan nilai mutlak 7,19%.

Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan Ljung-Box. Hasil pengujian nilai p-value pada lag 5 sampai dengan lag 30 tidak ada yang signifikan yang ditunjukkan dengan nilai p-value tidak ada yang lebih kecil dari nilai 0.05 (kepercayaan 95%), sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (1,2,3) layak digunakan.

Tabel 15. Uji Sisaan Ljung Box Model Arima (1,2,3)

```
lags statistic df      p-value
5      2.559639  5      0.7674877
10     4.345178  10     0.9304349
15     6.816083  15     0.9625537
20     7.602043  20     0.9941903
25     8.457486  25     0.9991764
30    13.895486  30     0.9946440
```

Selanjutnya dilakukan pengepasan model untuk seluruh data. Untuk Model ARIMA (1,2,3) koefisien ar1 sebesar 0,473 dan koefisien ma2= -0,166. Jika melakukan run model ARIMA (1,2,3) untuk seluruh data yaitu dari tahun 1970 – 2020 maka akan dihasilkan MAPE sebesar 4,45%. Hal ini menunjukkan bahwa antara data estimasi dengan data actual akan berbeda rata-rata berkisar antara -4,45% sampai +4,45%. Untuk metode estimasi dengan bias masih dibawah 5% dianggap masih cukup baik dan akurat.

Tabel 16. Model Arima (1,2,3) untuk Seluruh Data

```
Series: karet[, "Produksi"]
ARIMA(1,2,3)
Coefficients:
      ar1      ma1      ma2      ma3
s.e.  0.4735  -1.1240  -0.1663  0.2908
s.e.  0.6559  0.6638  0.5066  0.2447
sigma^2 estimated as 1.819e+10:  log likelihood=-648.16
AIC=1306.32  AICc=1307.71  BIC=1315.78
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 15060.14 126694.2 86379.59 0.9403439 4.459598 0.8473851 -0.02431017
```

Tabel 17. Ouput Peramalan Model Arima (1,2,3) untuk Produksi Karet

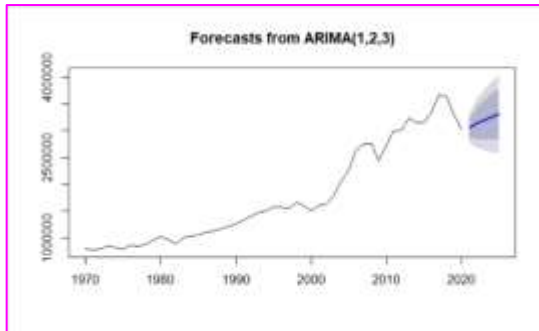
Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	3045340	2870817	3219863	2778431	3312249
2022	3135344	2839565	3431123	2682989	3587699
2023	3202055	2832612	3571499	2637041	3767070
2024	3257738	2830295	3685182	2604021	3911456
2025	3308200	2830262	3786138	2577256	4039144

Dengan menggunakan model ARIMA (1,2,3) menghasilkan angka estimasi produksi karet untuk 5 tahun ke depan. Hasil Estimasi dengan model ARIMA ini pada tahun 2021 produksi karet nasional sebesar 3,045 juta ton atau naik 0,26% dibandingkan tahun 2020. Pada tahun 2022 produksi karet diestimasi akan naik sebesar 2,96% menjadi 3,135 juta ton. Pada tahun 2023 sampai 2024 juga menunjukkan produksi karet nasional mengalami kenaikan secara perlahan masing-masing sebesar 2,13% dan 1,74%. Sementara pada tahun 2025 produksi karet masih meningkat sebesar 1,55%. Jika dibandingkan pertumbuhan produksi karet selama 5 tahun terakhir (tahun 2015 -2020) dengan menggunakan Angka Tetap rata-rata produksi turun sebesar 2,20% per tahun, sementara hasil estimasi lima tahun kedepan (2021 – 2025) rata-rata pertumbuhan naik 1,73% per tahun atau lebih tinggi dari data historisnya. Hal ini terjadi karena beberapa tahun terakhir harga karet dunia terus turun, diharapkan pada lima tahun kedepan harga karet sudah mulai membaik, jika harga karet dunia meningkat maka pertumbuhan produksi karet diduga akan positif.

Tabel 18. Hasil Estimasi Produksi Karet dengan Model ARIMA (1,2,3)

Tahun	Estimasi Produksi (Ton)	Pertumbuhan (%)
2020	3.037.348	
2021	3.045.340	0,26
2022	3.135.344	2,96
2023	3.202.055	2,13
2024	3.257.738	1,74
2025	3.308.200	1,55
Rata-rata pertumbuhan 2021 - 2025		1,73

Pada Gambar 7 terlihat pada tahun pertama hasil estimasi atau tahun 2021, produksi karet naik perlahan 0,26%. Namun pada tahun 2022 – 2023 produksi karet cenderung naik secara lebih cepat dengan pertumbuhan di atas 2%, dan pada tahun 2024 - 2025 produksi karet cenderung naik lebih lambat, pertumbuhan 1,55% sampai 1,74%.



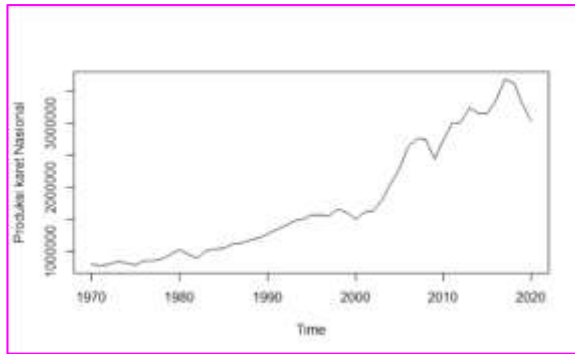
Gambar 7. Hasil Estimasi Produksi Karet Tahun 2021 – 2025 Model Arima (1,2,3)

## F. Model Fungsi Transfer

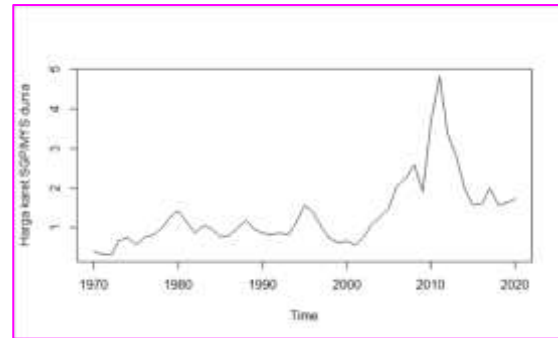
Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (*Autoregressive Integrated Moving Average*). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variabel lain sebagaimana pada model regresi. Model fungsi transfer merupakan suatu model yang menggambarkan nilai prediksi dari suatu time series (deret output atau  $Y_t$ ) berdasarkan nilai-nilai dari deret itu sendiri ( $Y_t$ ) dan berdasarkan pula pada data time series yang mempunyai hubungan (deret input atau  $X_t$ ) dengan deret output. Model fungsi transfer yang bersifat dinamis berpengaruh tidak hanya pada hubungan linier antara waktu ke- $t$  input  $X_t$  dan waktu ke- $t$  output  $Y_t$ , tetapi juga saat input  $X_t$  dengan saat  $t, t+1, \dots, t+k$  pada output  $Y_t$  (Fitriani et al., 2013). Dalam melakukan estimasi dengan fungsi transfer diperlukan dua variabel, yaitu output  $Y_t$  dan input  $X_t$ . Input  $X_t$  merupakan variabel yang diduga paling berpengaruh terhadap output  $Y_t$ . Pada analisis ini dicontohkan estimasi untuk produksi karet (output  $Y_t$ ) yang diduga dipengaruhi oleh harga karet dunia.

Dalam bahasan ini, model fungsi transfer variabel output produksi karet akan melibatkan variabel harga karet dunia sebagai variabel input. Langkah awal dalam mencari model fungsi transfer menggunakan software RStudio adalah penyiapan data, baik output  $Y_t$  (produksi karet) maupun input  $X_t$  (harga karet dunia). Produksi karet dalam satuan hektar bersumber dari Direktorat Jenderal Perkebunan, Kementerian Pertanian, sedangkan harga karet dunia dalam satuan US\$/Kg bersumber dari World Bank. Banyaknya series data yang digunakan baik untuk produksi maupun harga karet dunia harus seragam. Pada bahasan ini, series data yang digunakan yaitu 1970-2020 atau sebanyak 51 observasi.

Eksplorasi data dilakukan dengan menampilkan plot data produksi maupun harga karet dunia. Berdasarkan plot data dapat diketahui pola data series 51 tahun yang akan digunakan untuk pemodelan. Berdasarkan Gambar 8 dan Gambar 9, terlihat bahwa terdapat data produksi karet nasional memiliki tren meningkat dari tahun ke tahun sedangkan harga karet dunia cenderung berfluktuatif. Harga karet dunia cenderung turun pada sekitar 10 tahun terakhir, akibat persaingan harga dengan produk karet sintetis. Produksi karet nasional maupun harga karet dunia terindikasi tidak stasioner berdasarkan plotnya.



Gambar 8. Plot Data Produksi Karet, 1970-2020



Gambar 9. Plot Data Harga Karet Dunia, 1970-2020

Tahapan penyusunan model Fungsi Transfer produksi karet dengan variabel input harga karet dunia adalah sebagai berikut :

- h. Pembagian series data awal menjadi series data *training* dan *testing*
- i. Pemeriksaan kestasioneran
- j. Pencarian model tentatif untuk variabel input
- k. *Prewhitening* dan korelasi silang
- l. Identifikasi model noise
- m. Pengepasan model
- n. Peramalan berbasis fungsi transfer

Data produksi dan harga karet tahun 1970-2020 sebanyak 51 series akan dibagi menjadi series data *training* untuk periode 1970-2014 dan series data *testing* untuk periode 2015-2020. Selanjutnya dilakukan uji kestasioneran data untuk data input  $X_t$  yaitu harga karet dunia menggunakan uji Augmented Dickey-Fuller (ADF).

Hipotesis pada uji ADF ini adalah:

$H_0$  : data tidak stasioner

$H_1$  : data stasioner

Nilai test-statistik yang lebih besar dari critical values (nilai tau3), baik untuk taraf 1%, 5% maupun 10% menunjukkan bahwa  $H_0$  gagal ditolak, atau series data harga karet dunia belum stasioner. Oleh karena itu akan dilakukan pembedaan/differencing satu kali dan kemudian dilakukan uji ADF. Output hasil uji Augmented Dickey-Fuller sebelum dilakukan *differencing* menunjukkan belum stasioner. Setelah dilakukan Differencing 1 hasil uji menunjukkan sudah stasioner seperti pada Tabel 19.

Tabel 19. Hasil uji Kestationeran untuk Harga Karet Dunia Setelah Differencing 1

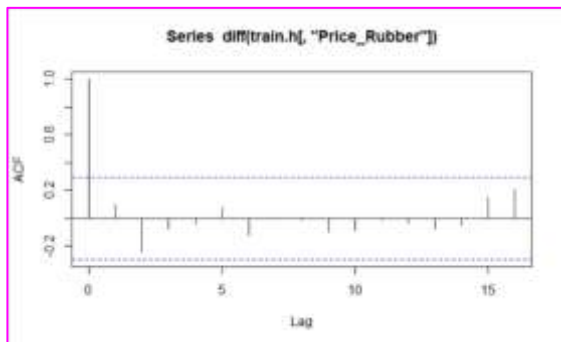
```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
    Min       1Q   Median       3Q      Max
-1.15540 -0.14969  0.02092  0.24884  1.90565
Coefficients:
            Estimate Std. Error t value Pr(>|t|)
z.lag.1     -1.1379    0.2185   -5.207 6.11e-06 ***
z.diff.lag   0.2675    0.1610    1.662  0.104
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.4799 on 40 degrees of freedom
Multiple R-squared:  0.462,    Adjusted R-squared:  0.4351
F-statistic: 17.17 on 2 and 40 DF,  p-value: 4.133e-06

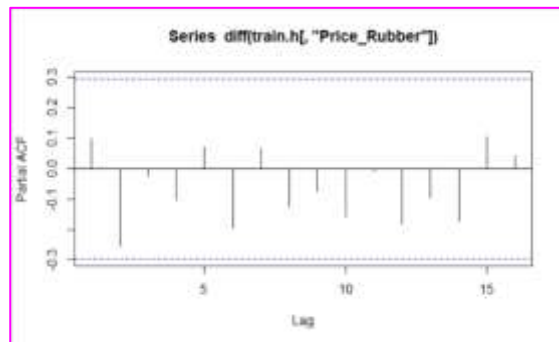
Value of test-statistic is: -5.2068
Critical values for test statistics:
    1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

Uji ADF pada data yang telah dilakukan *differencing* satu kali menunjukkan bahwa nilai *test-statistic* yaitu -5,206 lebih kecil dari *critical values* (nilai tau1) baik 1% maupun 5% menunjukkan bahwa  $H_0$  ditolak yang berarti data harga karet dunia sudah stasioner pada differencing 1.

Stasioneritas data dapat juga dilihat dari plot ACF dan PACF. Jika lag pada ACF maupun PACF menunjukkan *tail off* maka data tidak stasioner. Berikut adalah output untuk menunjukkan plot ACF dan PACF harga karet dunia setelah differencing 1.



Gambar 10. Plot ACF Harga Karet Dunia dengan *Differencing* satu kali



Gambar 11. Plot PACF Harga Karet Dunia dengan *Differencing* satu kali

Plot ACF cenderung cut off dimana hanya lag 0 yang keluar dari *confidence interval*. Plot PACF tidak menunjukkan cut off, ditunjukkan tidak ada yang melewati di luar garis *confidence interval*. Hal tersebut memperkuat hasil uji ADF yaitu data harga karet dunia telah stasioner, karena sudah *cut off*.

Selain itu, plot ACF dan PACF juga dapat digunakan untuk mengidentifikasi ordo AR dan MA, dengan cara melihat lag yang nyata atau keluar dari garis *confidence interval*. Namun demikian, pada plot ACF diatas, terlihat nyata pada lag=0 yang sering terjadi pada data riil, sehingga sulit untuk mengidentifikasi ordo AR dan MA yang tepat melalui plot ACF dan PACF.



Pencarian model tentatif variabel input harga karet sheet dilakukan melalui penelusuran menggunakan model ARIMA. Model terbaik dapat dipilih menggunakan script *auto arima* yang tersedia pada RStudio. Data yang digunakan untuk memilih model terbaik adalah series data *training*.

Tabel 20. Model Auto Arima Harga Karet Dunia

```
Series: train.h[, "Price_Rubber"]
ARIMA(0,1,0)

sigma^2 estimated as 0.2266: log likelihood=-29.78
AIC=61.55 AICc=61.65 BIC=63.34

Training set error measures:
Training set  ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
0.03432016  0.4707504  0.3058313  0.3449166  20.70712  0.9778067  0.0975218
```

Berdasarkan pemilihan orde ARIMA menggunakan *automodel* menyarankan bahwa model terbaik untuk harga karet dunia adalah ARIMA (0,1,0) dengan MAPE data training yang masih sangat tinggi yaitu sebesar 20,7%. Model ARIMA (0,1,0) untuk data harga karet dunia tersebut menghasilkan nilai AIC = 61,55 sebagaimana output yang disajikan pada Tabel 18. Oleh karena model tentatif ARIMA (0,1,0) maka tidak dapat dilakukan pengujian koefisien ar maupun ma karena order AR dan order MA nilainya 0. Hasil estimasi jika menggunakan ARIMA (0,1,0) akan cenderung konstan, karena yang berpengaruh hanya differencing saja. Oleh karena model ARIMA (0,1,0) hasil peramalan kurang memuaskan maka perlu dicari model tentative lainnya. Pada Tabel 19 output yang dihasilkan untuk memilih model tentative terbaik untuk model ARIMA faktor input  $X_t$  yaitu harga karet sheet nasional, dengan *option arima selection*.

Tabel 21. Model Arima Tentatif untuk Harga Karet Sheet (Differencing 1)

```

p q      sbc
[1,] 0 0 -64.54096
[2,] 1 0 -60.22690
[3,] 2 0 -58.65021
[4,] 3 0 -54.38460
[5,] 4 0 -51.42937
[6,] 5 0 -48.47765
[7,] 0 1 -41.87272
[8,] 0 2 -41.70634
[9,] 1 2 -38.87341
[10,] 1 1 -38.13811
```

Hasil output R-Studio akan menunjukkan sepuluh model tentatif dimana idealnya model terbaik adalah model yang memiliki nilai SBC terkecil. Model ARIMA yang direkomendasikan ditunjukkan dari nilai p,d,q. Sebagai contoh model pertama dengan nilai p=0 dan q=0. Karena data harga karet sheet telah dilakukan *differencing* satu kali berarti d=1, artinya model yang direkomendasikan adalah ARIMA (0,1,0). Model tentative lainnya adalah ARIMA (1,1,0), ARIMA(2,1,0), dan seterusnya sampai ARIMA (1,1,1). Untuk mengetahui apakah model ARIMA (1,1,0) lebih baik dari hasil *auto.arima* yaitu ARIMA (0,1,2) maka dapat dibandingkan dari nilai AIC dan koefisien ar1.

Selanjutnya dilakukan pemilihan model ARIMA terbaik untuk faktor input Harga Karet Dunia. Pengujian dilakukan dengan memperhatikan tingkat signifikansi koefisien ar atau ma dan



pengujian dengan menggunakan data training dan data testing untuk peubah harga karet dunia. Hasil pengujian model ARIMA tentatif pada differencing1 terlihat pada Tabel 20.

Tabel 22. Model Tentatif Harga Karet Sheet pada Diiferencing 1

Model ARIMA	Koefisien Signifikansi	Koefisien Tidak Signifikan	AIC
ARIMA (1,1,0)	-	-	63,04
ARIMA (2,1,0)	ar2	ar1	62,24
ARIMA (3,1,0)	ar2	ar1 dan ar3	64,2
ARIMA (4,1,0)	ar2	ar1 ar3 ar4	65,04
ARIMA (0,1,1)	-	-	62,55
ARIMA (0,1,2)	ma2	ma1	60,22
ARIMA (1,1,2)	-	ar1 ma1 ma2	61,87
ARIMA (1,1,1)	ar1 ma1	-	60,5

Hasil pengujian model terbaik untuk harga karet dunia, pada differencing 1 masing-masing nilai ar dan ma ada yang signifikan ada yang tidak signifikan, dan nilai AIC bervariasi dengan rentang 60,22 sampai 65,04. Berdasarkan Tabel 20, nilai AIC yang terkecil adalah ARIMA(0,1,2) tetapi setelah dilakukan pengujian signifikansi menunjukkan bahwa ma2 signifikan, tetapi ma1 tidak signifikan. Model tentatif lain yang terbaik adalah ARIMA (1,1,1) karena menghasilkan nilai AIC terkecil kedua yaitu sebesar 60,5, pengujian signifikansi untuk ARIMA(1,1,1) menunjukkan bahwa ar1 maupun ma1 signifikan. Berdasarkan pertimbangan nilai koefisien yang signifikan dan nilai AIC maka model terbaik untuk estimasi harga karet dunia adalah model ARIMA(1,1,1).

Tabel 23. Pengujian Koefisien Model ARIMA (1,1,1) Harga Karet Dunia

```
Call:
arima(x = train.h[, "Price_Rubber"], order = c(1, 1, 1))

Coefficients:
      ar1      ma1
 -0.6158  0.8995
s.e.    0.1888  0.1032

sigma^2 estimated as 0.1997:  log likelihood = -27.25,  aic = 60.5
> library(lmtest)
> coeftest(model1)

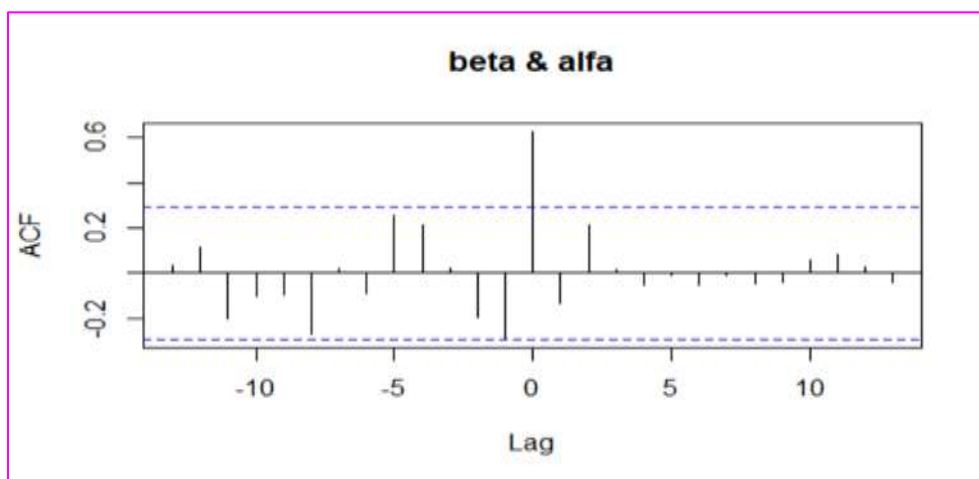
z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1 -0.61577    0.18875 -3.2623  0.001105 **
ma1  0.89946    0.10320  8.7160 < 2.2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model ARIMA (1,1,1) menghasilkan nilai AIC = 60,5 relatif kecil dibandingkan model ARIMA pada order lainnya. Berdasarkan Tabel 21, model ARIMA (1,1,1) menunjukkan bahwa ar1 signifikan pada taraf 1%, dan ma1 signifikan pada taraf 0,1%. Hal ini menunjukkan semua

komponen pada model ARIMA (1,1,1) signifikan. Oleh karena ARIMA (1,1,1) semua koefisien komponen ar dan ma signifikan dan nilai AIC relative kecil, maka selanjutnya yang digunakan untuk peramalan harga karet dunia sebagai faktor input model fungsi transfer untuk meramalkan produksi karet nasional menggunakan ARIMA (1,1,1).

Tahap selanjutnya untuk penyusunan model fungsi transfer ini adalah *prewhitening* dan korelasi silang. Korelasi silang menggambarkan struktur hubungan antara  $X_t$  dengan  $Y_t$ . Untuk mengidentifikasi pengaruh  $X_t$  terhadap  $Y_t$  maka deret  $X_t$  harus stasioner atau sudah distasionerkan. Dalam konteks pemodelan  $X_t$  terhadap  $Y_t$ , untuk membuat  $X_t$  stasioner tidak dengan pembedaan (*differencing*) namun dengan mengambil komponen *white noise* dari  $X_t$  (*prewhitening*). *Prewhitening* dilakukan terhadap deret input  $X_t$  yang didefinisikan sebagai alfa serta deret input  $Y_t$  yang didefinisikan sebagai beta. Hasil output untuk pre-whitening dan korelasi silang berupa grafik ACF untuk beta dan alfa.



Gambar 12. Plot korelasi silang Produksi Karet dengan Harga Karet Dunia

Hasil plot korelasi silang digunakan untuk mengidentifikasi ordo r, s, dan b. Ordo r adalah panjang lag Y periode sebelumnya yang masih mempengaruhi  $Y_t$ , ordo s adalah panjang lag X periode sebelumnya yang masih mempengaruhi  $Y_t$ , dan ordo b adalah panjang jeda pengaruh  $X_t$  terhadap  $Y_t$ . Identifikasi ordo r, s dan b hanya dilihat pada lag yang positif.

Plot korelasi silang diatas menunjukkan bahwa ada yang signifikan ditunjukkan dengan ada plot yang keluar dari garis signifikansi yaitu pada lag ke-0, maka nilai  $b=0$ . Kemudian, tidak ada tambahan lagi lag yang signifikan, ditunjukkan tidak ada garis yang keluar dari batasnya, maka nilai  $s=0$ . Mengingat data produksi karet dan harga karet dunia merupakan data tahunan yang tidak mengandung musiman maka diasumsikan nilai  $r=0$ . Nilai  $b=0$  menunjukkan tidak ada jeda pengaruh antara harga karet dunia pada waktu t terhadap produksi karet pada waktu t. Nilai  $s=0$  berarti ada korelasi antara produksi dan harga karet sheet pada tahun yang sama. Dengan kata lain, dampak dari harga karet sheet terhadap produksi dirasakan pada waktu yang sama (t).

Hasil pengujian dengan Identifikasi Awal nilai  $r=0$   $s=0$  dan  $b=0$ , menghasilkan nilai MAPE yang cukup baik yaitu sebesar 10,82. Hal ini menunjukkan bahwa dengan ARIMA (0,0,0) sudah cukup baik untuk meramalakan produksi karet, seperti pada Tabel 24.

Tabel 24. Fungsi Transfer Identifikasi Awal ( $r = 0$ ,  $s=0$ , dan  $b=0$ ) ARIMA (0,0,0)

```
Series: train.h[, "Produksi"]
Regression with ARIMA(0,0,0) errors

Coefficients:
      intercept          xreg
      709974      662854.07
s.e.      109912      69496.04

sigma^2 estimated as 1.885e+11:  log likelihood=-646.98
AIC=1299.95  AICC=1300.54  BIC=1305.37

Training set error measures:
Training set  -2.518269e-07  424354.3  358241.2  -8.35368  25.32734  4.030396
              ACF1
Training set  0.7202991
```

Untuk menghasilkan orde yang paling tepat untuk menentukan orde Arima fungsi transfer dengan melakukan identifikasi model noise. Untuk menghasilkan model terbaik dengan menggunakan *auto-arima* pada R Studio, model maka noise yang disarankan adalah ARIMA (1,1,2). Tahap selanjutnya dilakukan pengepasan model, dengan model noise. Hasil pengujian fungsi transfer dengan nilai  $r=0$ ,  $s=0$ , dan  $b=0$  dengan model noise ARIMA (1,1,2) menghasilkan nilai MAPE yang cukup besar yaitu 81,47%, sehingga perlu dicari model alternatif lain. Pemilihan model noise tentative, terlihat seperti Tabel 25.

Tabel 25. Pengujian Orde  $r=0$ ,  $s=0$ , dan  $b=0$  ARIMA Noise (1,1,2) Produksi Karet

```
Series: res
ARIMA(1,1,2)

Coefficients:
      ar1      ma1      ma2
      0.3982  -0.0430  -0.6897
s.e.      0.2140   0.1516   0.1106

sigma^2 estimated as 5.511e+10:  log likelihood=-605.6
AIC=1219.2  AICC=1220.23  BIC=1226.34

Training set error measures:
Training set  36679.8  224071.8  137437.8  8.449663  81.46728  0.7988704  -0.02136444
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
```

Tabel 26. Model Residual/Noise Tentatif ARIMA untuk Produksi Karet

```
[1,] p q      sbc
[2,] 4 0 1100.604
[3,] 2 0 1102.204
[4,] 1 0 1102.315
[5,] 3 0 1105.596
[6,] 5 0 1105.656
[7,] 4 1 1117.477
[8,] 1 2 1118.589
[9,] 0 1 1119.625
[10,] 2 2 1119.626
```

Oleh karena model *autoarima* disarankan differencing tingkat 1, maka solusinya akan dicari model alternative noise pada differencing 1. Model alternatif yang diberikan untuk *model noise* seperti Tabel 25. Model tentatif pertama yang disarankan untuk data residual adalah ARIMA (0,1,0). Selanjutnya model tersebut didefinisikan sebagai *modelres* atau *model residual* dan dilihat signifikansi AR. Oleh karena model noise ARIMA (1,1,2) sudah diujicobakan dan menghasilkan MAPE yang tinggi, maka selanjutnya dilakukan uji untuk model noise untuk residual dengan Arima (2,1,0), Arima (1,1,0), Arima (0,1,1), dan Arima (2,1,2) (Tabel 25).

Tabel 25. Pemilihan Model Arima Residual untuk Fungsi Transfer

Model Residual	ARIMA	Signifikansi ARIMA	Signifikansi Xreg	MAPE Training
ARIMA (2,1,0)		ar1 signifikan	Xreg signifikan	3,67
ARIMA (0,1,1)		ma1 signifikan	Xreg signifikan	3,39

Setelah dilakukan pengujian model noise, maka ada 2 model noise tentative terbaik yaitu model Noise Arima (2,1,0) dan Arima (0,1,1). Berdasarkan Tabel 25, model Noise Arima (2,1,0) menghasilkan komponen ar1 signifikan, komponen Xreg signifikan, dan MAPE data training cukup kecil yaitu 3,67. Untuk model Noise Arima(0,1,1) menghasilkan komponen ma1 signifikan, komponen Xreg signifikan, dan menghasilkan MAPE data training lebih kecil yaitu sebesar 3,39. Berdasarkan identifikasi ini maka model fungsi transfer untuk Produksi Karet Nasional adalah pada nilai  $r=0$ ,  $b=0$ , dan  $b=0$  dengan model Noise ARIMA (2,1,0) dan model Noise Arima (0,1,1) dan komponen Xreg adalah harga karet dunia.

Tahap berikutnya adalah dilakukan pengepasan model dengan noise. Pengepasan model dengan noise dilakukan dengan menggabungkan nilai  $r=0$ ,  $s=0$ ,  $b=0$  dan hasil identifikasi model noise yaitu ARIMA(2,1,0). Selanjutnya diperoleh output model fungsi transfer ARIMA (2,1,0) seperti pada Tabel 26.

Tabel 26. Model Fungsi Transfer ARIMA (2,1,0) Untuk Produksi Karet

```
Series: train.h[, "Produksi"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
 0.6750 -0.2342 187425.91
s.e. 0.1582 0.1850 28455.52

sigma^2 estimated as 8.272e+09: log likelihood=-563.51
AIC=1135.02 AICC=1136.04 BIC=1142.15

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 24616.98 86812.56 62296.23 1.347919 3.671636 0.7008643 -0.1062611
> coeftest(tf.arima1)

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1  6.7496e-01 1.5822e-01 4.2660 1.990e-05 ***
ar2 -2.3419e-01 1.8495e-01 -1.2662 0.2054
xreg 1.8743e+05 2.8456e+04 6.5866 4.499e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model fungsi transfer produksi karet dengan ARIMA(2,1,0) menghasilkan MAPE yang sangat baik karena nilainya sangat kecil yaitu sebesar 3,67%. Untuk mengetahui *performance* dari model fungsi transfer tersebut dilakukan pengujian *coeftest* untuk mengetahui tingkat signifikansi dari masing- masing koefisien. Berdasarkan hasil uji Z tes untuk fungsi transfer ARIMA (2,1,0), untuk koefisien ar1 signifikan dengan tingkat kepercayaan 99%, dan komponen fungsi transfer xreg yaitu variabel harga karet dunia juga signifikan di tingkat kepercayaan 99%.

Untuk membandingkan kelayakan model , maka akan dicobakan model fungsi transfer dengan model tentatif komponen noise ARIMA (0,1,1). Hasil output untuk fungsi transfer dengan model noise ARIMA (0,1,1) seperti pada Tabel 27.

Tabel 27. Model Tentatif Fungsi Transfer ARIMA (0,1,1) Untuk Produksi Karet

```
Series: train.h[, "Produksi"]
Regression with ARIMA(0,1,1) errors

Coefficients:
      ma1      xreg
    0.6936 198990.47
s.e. 0.1473 31604.35

sigma^2 estimated as 8.103e+09: log likelihood=-563.68
AIC=1133.36 AICc=1133.96 BIC=1138.71

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 26863.24 86964.56 58341.48 1.443582 3.397469 0.6563714 -0.04849772
> coefstest(tf.arimal)

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ma1 6.9361e-01 1.4730e-01 4.7088 2.492e-06 ***
xreg 1.9899e+05 3.1604e+04 6.2963 3.048e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

MAPE yang dihasilkan dari fungsi transfer ARIMA (0,1,1) adalah 3,39 lebih kecil dari ARIMA (2,1,0), jadi lebih baik dari MAPE fungsi transfer ARIMA(2,1,0). Selanjutnya dilakukan pengujian *coefstest* untuk fungsi transfer ARIMA(0,1,1) untuk melihat signifikansi dari komponen ma. Dari hasil pengujian di atas, ditemukan bahwa komponen ma signifikan pada tingkat kepercayaan 99%, dan komponen fungsi transfer xreg juga signifikan pada tingkat kepercayaan 99%. Berdasarkan hasil penelusuran model dan uji yang telah dilakukan maka disimpulkan bahwa ada 2 model tentative terbaik model fungsi transfer produksi karet ARIMA (2,1,0) dan model ARIMA fungsi transfer (0,1,1).

### Peramalan Berbasis Fungsi Transfer

Berdasarkan model fungsi transfer dengan model noise ARIMA (2, 1, 0), dilakukan peramalan berbasis nilai aktual dimana produksi karet diestimasi menggunakan data aktual harga karet dunia periode 2015-2020. Meskipun data aktual produksi karet periode 2015-2020 telah ada, dilakukan peramalan produksi karet untuk mengecek *performance/kelayakan* model fungsi transfer.

Tabel 28. Pengujian MAPE Model Fungsi Transfer ARIMA (2,1,0) Dengan Faktor Input Harga Karet Dunia Data Aktual tahun 2015 - 2020

```
Series: test.h[, "Produksi"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
    0.675 -0.2342 187425.9
s.e. 0.000 0.0000 0.0

sigma^2 estimated as 8.272e+09: log likelihood=-67.62
AIC=137.24 AICc=138.57 BIC=136.85

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set -22185.85 157443.5 123349.6 -0.7320572 3.631104 0.5235062
```

Peramalan produksi karet periode 2015-2020 menggunakan fungsi transfer ARIMA (2,1,0) dengan input harga karet dunia **data aktual** menghasilkan MAPE testing 3,63%.

Tujuan melakukan pemodelan fungsi transfer adalah untuk mendapatkan nilai ramalan periode ke depan, yakni produksi karet tahun 2021 - 2025. Karena data series input harga karet dunia tersedia hingga tahun 2020, maka perlu dilakukan peramalan harga karet dunia terlebih dahulu atau dengan kata lain peramalan produksi dilakukan berbasis nilai ramalan harga karet.

Oleh karenanya, untuk pengujian data testing terlebih dahulu dilakukan estimasi harga karet dunia periode 2021 - 2025 menggunakan model ARIMA (1,1,1) sebagaimana yang telah diperoleh dari tahap pencarian model tentatif untuk variabel input, sebagai variabel input harga karet sheet. Pemilihan variabel input harga karet sheet karena secara tidak langsung berpengaruh juga pada produksi karet nasional. Selanjutnya dilakukan peramalan produksi karet dengan fungsi transfer ARIMA (2, 1, 0) sebagai model terbaik berdasarkan tahapan pengepasan model dengan noise. Peramalan produksi dengan fungsi transfer ARIMA (2,1,0) menggunakan **nilai ramalan** harga karet dunia yang telah diramal dengan model ARIMA (1,1,1).

Tabel 29. Pengujian MAPE Model Fungsi Transfer ARIMA (2,1,0) Dengan Faktor Input Harga Karet Dunia Nilai Ramalan tahun 2015 – 2020.

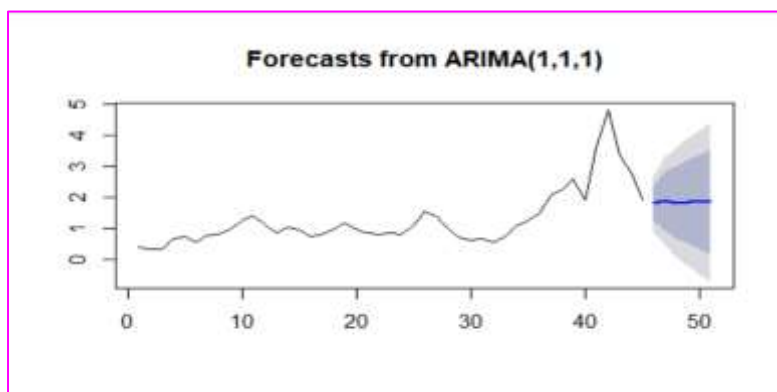
```
Series: test.h[, "Produksi"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
      0.675  -0.2342 187425.9
s.e.      0.000   0.0000   0.0

sigma^2 estimated as 8.272e+09:  log likelihood=-68.04
AIC=138.08  AICC=139.42  BIC=137.69

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -20468.93 171325 147654.8 -0.6657773 4.264563 0.6266595 0.252738
```

Estimasi produksi karet berbasis fungsi transfer dengan model fungsi transfer ARIMA (2,1,0) selama 6 tahun terakhir (2015-2020) menggunakan input harga karet dunia hasil **angka ramalan** ARIMA (1,1,1) menghasilkan MAPE testing 4,26%. Nilai MAPE yang kurang dari 5% ini untuk penyusunan estimasi sudah cukup baik, karena kesalahannya sangat kecil.



Gambar 13. Hasil Peramalan Data Testing Faktor Input Harga Karet Dunia 2015 - 2020

Setelah dilakukan peramalan produksi karet baik menggunakan input (harga karet dunia) data aktual maupun ramalan, tahapan berikutnya adalah pengepasan model arima output. Pengepasan model ARIMA output dimaksudkan untuk membandingkan hasil ramalan produksi baik berdasarkan data training (1970-2014) maupun data testing (2015-2020). Model ARIMA biasa, tanpa fungsi transfer, terbaik untuk variabel produksi yang terpilih adalah ARIMA (1,2,3). Model ARIMA (1,2,3) tersebut merupakan ARIMA biasa dan bukan hasil dari fungsi transfer. Hasil ramalan model ARIMA tanpa fungsi transfer dan model ARIMA dengan fungsi transfer selanjutnya akan dibandingkan.

Tabel 30. Pengujian MAPE Model ARIMA (1,2,3) Tanpa Fungsi Tranfer Untuk Faktor Output (Produksi Karet) tahun 2015 – 2020.

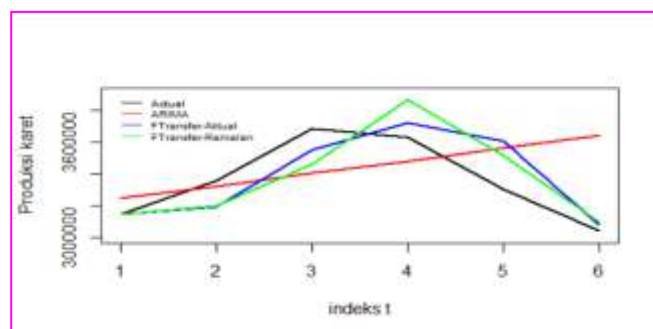
```
Series: train.h[, "Produksi"]
ARIMA(1,2,3)
Coefficients:
      ar1      ma1      ma2      ma3
      -0.7759  0.0390 -0.8952  0.0085
s.e.      0.1952  0.2731  0.1496  0.1853

sigma^2 estimated as 1.387e+10:  log likelihood=-562.33
AIC=1134.66  AICC=1136.28  BIC=1143.46

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 18456.47 109655.7 70001.3 1.059132 4.14669 0.7875503 -0.05083289
> ramalan.output <- forecast(model.output,6)
> accuracy(ramalan.output,test.h[, "Produksi"])
      ME      RMSE      MAE      MPE      MAPE      MASE
Training set 18456.47 109655.7 70001.3 1.059132 4.14669 0.7875503
Test set     -80922.02 300133.4 238528.7 -2.957646 7.29509 2.6835699
      ACF1
Training set -0.05083289
Test set     NA
```

Hasil peramalan produksi karet dengan ARIMA biasa atau ARIMA tanpa fungsi transfer menunjukkan bahwa hasil ramalan jika menggunakan data training menghasilkan MAPE = 4,15%. Di sisi lain, MAPE yang dihasilkan dari data testing yaitu 7,29%. Artinya jika model yang digunakan hanya ARIMA tanpa fungsi transfer maka MAPE data testing cenderung lebih tinggi dibandingkan dengan MAPE fungsi transfer ARIMA (2,1,0).

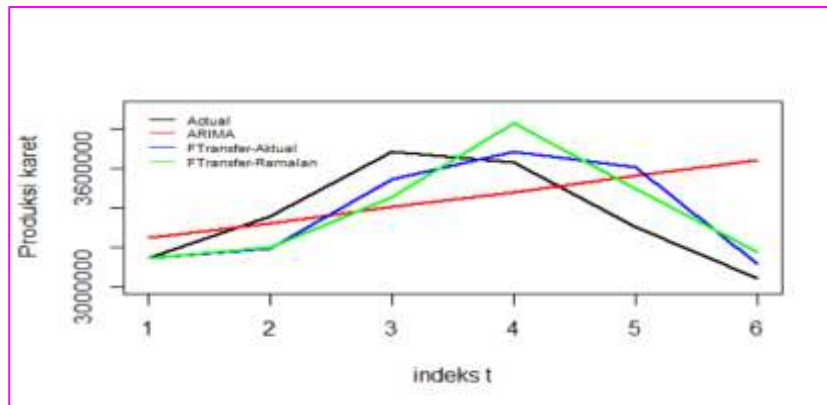
Untuk membandingkan ketepatan model estimasi, dilakukan perbandingan hasil estimasi terhadap data aktual produksi karet pada tahun 2015-2020 (data testing). Hasil ramalan yang dibandingkan yaitu ramalan dengan Arima biasa, model ARIMA(1,2,3), model fungsi transfer ARIMA(2,1,0) dimana input harga karet dunia yang digunakan adalah data aktual maupun ramalan. Selanjutnya hasil ramalan dibandingkan dalam sebuah grafik seperti Gambar 15.



Gambar 14. Perbandingan Hasil Ramalan Produksi Karet Tahun 2015-2020 Menggunakan Model ARIMA (1,2,3) dan Model Fungsi Transfer ARIMA (2,1,0)



Dari grafik di atas terlihat jika dibandingkan dengan data aktual produksi karet 2015-2020 (warna hitam), maka hasil peramalan dengan model ARIMA tanpa fungsi transfer (warna merah) tidak mengikuti pola data aktualnya dan cenderung naik terus. Sebaliknya, peramalan dengan fungsi transfer khususnya jika input harga karet dunia yang digunakan adalah data aktual maka hasil ramalan produksinya (warna biru) sangat menyerupai pola data produksi sebenarnya selama 6 tahun terakhir. Jika input harga karet dunia yang digunakan adalah hasil ramalan, maka estimasi produksinya (warna hijau) juga menghasilkan angka estimasi yang sangat baik karena makin mendekati nilai produksi sebenarnya (data aktual), hal ini juga terlihat di grafik data hasil ramalan dan data actual hampir berimpit pada tahun ke-1 sampai tahun ke-6.



Gambar 15. Perbandingan Hasil Ramalan Produksi Karet Tahun 2015-2020 Menggunakan Model ARIMA (1,2,3) dan Model Fungsi Transfer ARIMA (0,1,1)

Hal yang dilakukan untuk run model Fungsi Transfer dengan ARIMA(0,1,1). Dari grafik terlihat pada Gambar 17, jika dibandingkan dengan data aktual produksi karet 2015-2020 (warna hitam), maka hasil peramalan dengan model ARIMA tanpa fungsi transfer (warna merah) tidak mengikuti pola data aktualnya dan cenderung naik terus. Sebaliknya, peramalan dengan fungsi transfer ARIMA (0,1,1) khususnya jika input harga karet dunia yang digunakan adalah data aktual maka hasil ramalan produksinya (warna biru) sangat menyerupai pola data produksi sebenarnya selama 6 tahun terakhir. Jika input harga karet dunia yang digunakan adalah hasil ramalan, maka estimasi produksinya (warna hijau) juga menghasilkan angka estimasi yang sangat baik karena makin mendekati nilai produksi sebenarnya (data aktual), hal ini juga terlihat di grafik data hasil ramalan dan data actual hampir berimpit pada tahun ke-1 sampai tahun ke-6. Berdasarkan pengamatan visual model Fungsi Transfer ARIMA (0,1,1) sepertinya antara data aktual dan data estimasi lebih dekat atau berimpit, dibandingkan model Fungsi Transfer ARIMA (2,1,0).

Hasil identifikasi model untuk meramalkan produksi karet dengan ke-3 cara diatas menghasilkan nilai MAPE masing-masing sebagai berikut:



Tabel 31. Perbandingan MAPE untuk Arima biasa dan Arima Fungsi Transfer

No	Model	MAPE Training	MAPE Testing
1	ARIMA (2,1,0) (tanpa Fungsi transfer)	4,15%	7,29%
2	Fungsi Transfer dengan ARIMA (2,1,0) input data aktual	3,67%	3,63%
3	Fungsi Transfer dengan ARIMA(2,1,0) input data ramalan	3,67%	4,26%
4	Fungsi Transfer dengan ARIMA (0,1,1) input data aktual	3,40%	3,65%
5	Fungsi Transfer dengan ARIMA(0,1,1) input data ramalan	3,40%	4,43%

Berdasarkan nilai MAPE yang dihasilkan maka model fungsi transfer ARIMA(2,1,0) dengan input data aktual lebih kecil jika dibandingkan dengan model fungsi transfer ARIMA (2,1,0) dengan input data ramalan maupun model ARIMA(2,1,0) tanpa fungsi transfer. Namun demikian, karena tujuan penyusunan model adalah untuk mendapatkan angka ramalan tahun 2021 - 2025, maka digunakan model Fungsi Transfer berbasis input data ramalan, yakni dengan terlebih dahulu dilakukan estimasi harga karet dunia lima tahun ke depan menggunakan model ARIMA (1,2,3). Hasil estimasi harga karet dunia tersebut kemudian menjadi input untuk meramalkan produksi karet lima tahun ke depan menggunakan model fungsi transfer ARIMA (2,1,0) dan sebagai variabel input harga karet sheet. Hasil ramalan produksi karet disajikan pada Tabel 26 dan Tabel 27.

Jika dibandingkan dengan Fungsi Transfer dengan order ARIMA(0,1,1), maka model ARIMA (0,1,1) menghasilkan MAPE data training yang lebih baik, sementara untuk MAPE data testing dengan input harga karet dunia nilai actual tidak berbeda secara signifikan antara model Fungsi Transfer ARIMA (2,1,0) dan model Fungsi Transfer ARIMA (0,1,1). Sementara untuk data testing dengan input harga karet dunia nilai ramalan, maka model Fungsi Transfer ARIMA (2,1,0) lebih baik karena memiliki MAPE yang lebih kecil dibandingkan ARIMA (0,1,1).

Selain mencari model terbaik untuk meramalkan produksi karet, akan diestimasi juga produksi karet lima tahun ke depan (2021-2025) menggunakan fungsi transfer ARIMA (2,1,0) dan fungsi transfer ARIMA (0,1,1) dengan menggunakan seluruh data. Setelah dilakukan run model Fungsi Transfer dengan ARIMA (2,1,0) untuk seluruh data (tidak lagi menggunakan data training), maka dihasilkan MAPE untuk seluruh data sebesar 1,21%, nilai ini juga relative kecil, sehingga estimasi lima tahun kedepan cukup akurat.

Tabel 32. Hasil Peramalan Produksi Karet Model Fungsi Transfer Arima (2,1,0) Tahun 2021 – 2025

```
Series: dataestimasi[, "Produksi"]
Regression with ARIMA(2,1,0) errors

Coefficients:
      ar1      ar2      xreg
0.7676 -0.3535 191279.63
s.e. 0.1342 0.1440 27112.01

sigma^2 estimated as 1.041e+10: log likelihood=-646.39
AIC=1300.77 AICC=1301.66 BIC=1308.42

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 22860.64 97963.16 69617.96 1.235742 3.679147 0.6829532 -0.0247545
Time Series:
Start = 52
End = 56
Frequency = 1
[1] 2919246 2962994 3005103 3045418 3044920
```

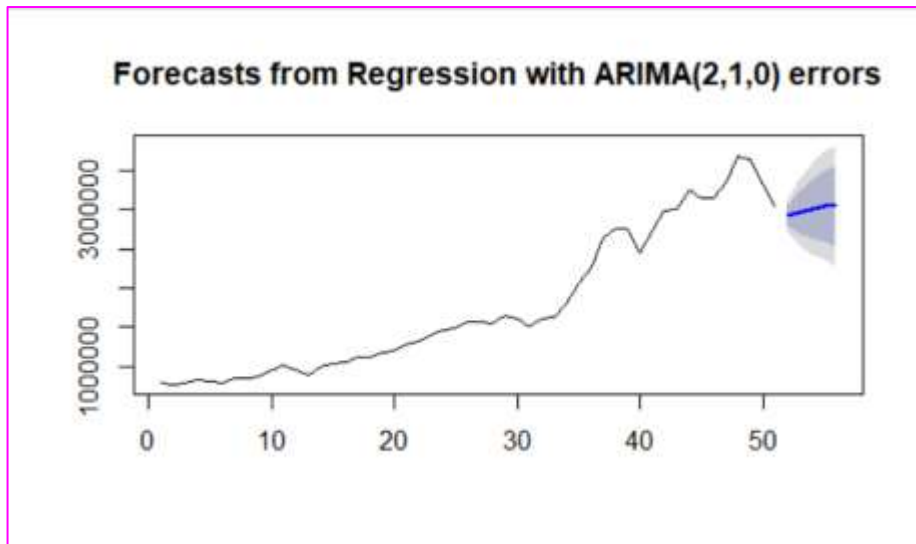
Tabel 33. Hasil Peramalan Produksi Karet Tahun 2021-2025 Model Fungsi Transfer ARIMA (2,1,0)

Tahun	Produksi (Ha)	Pertumbuhan (%)
2020	3,037,348	
2021	2,919,246	(3.89)
2022	2,962,994	1.50
2023	3,005,103	1.42
2024	3,045,418	1.34
2025	3,044,920	(0.02)
Rata-rata Pertumbuhan 2021 -2025		0.07

Tahun 2020 : Angka Tetap (Ditjenbun)

Tahun 2021 – 2025 : Estimasi Berdasarkan Model Fungsi Transfer

Pada tahun 2020 Angka Tetap (ATAP) untuk produksi karet sebesar 3.037.348 ton (Ditjen Perkebunan, 2021). Hasil peramalan dengan fungsi transfer ARIMA (2,1,0) menunjukkan bahwa selama lima tahun ke depan produksi karet akan meningkat rata-rata 0,07% per tahun. Tahun 2021 produksi karet diestimasi sebesar 2.919.246 ton atau turun 3.89% dibandingkan tahun 2020. Hasil estimasi tersebut hampir sama dibandingkan Angka Estimasi Ditjenbun untuk produksi karet tahun 2021 adalah sebesar 3.121.475 ton atau lebih rendah 6,93%. Produksi karet diramalkan terus meningkat hingga tahun 2025, dengan pertumbuhan yang semakin melambat, hingga pada tahun 2025 produksinya diestimasi sebesar 3.044.920 ton. Pada Gambar 17 menunjukkan bahwa hasil peramalan selama lima tahun ke depan (2021 – 2025) produksi karet masih tumbuh positif meskipun pertumbuhan kecil.



Gambar 16. Plot Peramalan Produksi Karet Nasional Fungsi Transfer ARIMA (2,1,0).

Tabel 34. Hasil Peramalan Produksi Karet Model Fungsi Transfer Arima (0,1,1) Tahun 2021 – 2025

```
Series: dataestimasi[, "Produksi"]
Regression with ARIMA(0,1,1) errors

Coefficients:
      ma1      xreg
    0.7602 198568.31
s.e. 0.0907 27696.65

sigma^2 estimated as 1.005e+10: log likelihood=-646.13
AIC=1298.25 AICc=1298.78 BIC=1303.99

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 21551.88 97253.99 65933.4 1.185247 3.409855 0.6468077 0.05230624
Time Series:
Start = 52
End = 56
Frequency = 1
[1] 2972801 2990268 2977935 2986643 2980495
```

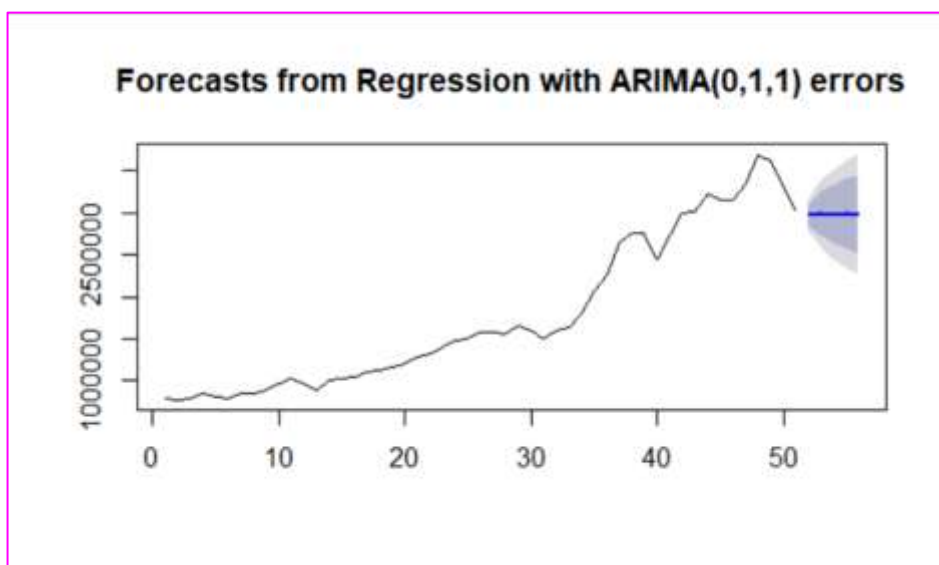
Tabel 35. Hasil Peramalan Produksi Karet Tahun 2021-2025 Model Fungsi Transfer ARIMA (0,1,1)

Tahun	Produksi (Ha)	Pertumbuhan (%)
2020	3,037,348	
2021	2,972,801	(2.13)
2022	2,990,268	0.59
2023	2,977,935	(0.41)
2024	2,986,643	0.29
2025	2,980,495	(0.21)
Rata-rata Pertumbuhan 2021 -2025		(0.37)

Tahun 2020 : Angka Tetap (Ditjenbun)

Tahun 2021 – 2025 : Estimasi Berdasarkan Model Fungsi Transfer

Jika menggunakan model Fungsi Transfer ARIMA (0,1,1) maka MAPE yang dihasilkan 3,41%, MAPE ini lebih baik dibandingkan dengan jika menggunakan model Arima biasa tanpa fungsi transfer. Pada tahun 2020 Angka Tetap (ATAP) untuk produksi karet sebesar 3.037.348 ton (Ditjen Perkebunan, 2021). Hasil peramalan dengan fungsi transfer ARIMA (0,1,1) menunjukkan bahwa selama lima tahun ke depan produksi karet akan turun rata-rata 0,37% per tahun. Tahun 2021 produksi karet diestimasi sebesar 2.972.801 ton atau turun 2,13% dibandingkan tahun 2020. Hasil estimasi tersebut hampir sama dibandingkan Angka Estimasi Ditjenbun untuk produksi karet tahun 2021 adalah sebesar 3.121.475 ton atau lebih rendah 4,76%. Produksi karet diramalkan berfluktuatif hingga tahun 2025, dengan pertumbuhan yang semakin melambat, hingga pada tahun 2025 produksinya diestimasi sebesar 2.980.495 ton. Pada Gambar 17 menunjukkan bahwa hasil peramalan selama lima tahun ke depan (2021 – 2025) produksi karet masih tumbuh negatif meskipun pertumbuhan kecil.



Gambar 17. Plot Peramalan Produksi Karet Nasional Menggunakan Fungsi Transfer ARIMA (0,1,1).

## G. Model VAR (*Vector Auto Regressive*)

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen. Model VAR berlaku pada saat nilai setiap variabel dalam sebuah system tidak hanya bergantung pada lag-nya sendiri, namun juga pada nilai lag variabel lain.

Tahapan yang dilakukan dalam penyusunan model VAR adalah sebagai berikut: persiapan data, pembagian data training dan testing, pemilihan lag dan type, pengajuan asumsi, ramalan data training dan testing, penghitungan MAPE, plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Disamping itu dibahas Interpretasi Impulse Response Function (IRF) dan Variance Decomposition.

Variabel yang digunakan untuk estimasi model VAR adalah luas areal (Areal) dalam satuan hektar, produksi (Produksi) dalam satuan ton, harga karet sheet (HRsheet) dalam

satuan Rp/kg, harga karet dunia (HKaretworld) dalam satuan US\$/kg, volume ekspor karet (volekkaret) dalam satuan ton, dan volume impor (volimkaret) dalam satuan ton. Data produksi karet, produksi, volume ekspor karet, volume impor karet, dan harga karet sheet diperoleh dari publikasi Ditjen Perkebunan, sementara data variabel harga karet dunia diperoleh dari World Bank. Series masing-masing variabel berbeda karena keterbatasan ketersediaan data. Series data produksi karet adalah dari tahun 1971-2020, series data harga sheet adalah dari tahun 1971-2020, series data harga karet dunia adalah dari tahun 1971-2020, sementara series data volume ekspor dan impor karet adalah dari tahun 1971-2020. Format data yang digunakan bisa dalam bentuk excell (CSV).

### **Pembagian Data Training dan Testing**

Series data yang digunakan adalah series tahun 1971 – 2020 akan dibagi menjadi 2 set data yakni set data training (tahun 1971-2014) atau 44 titik dan set data testing (2015-2020) atau 6 titik.

### **Pemilihan Lag (p) dan Type**

Dalam permodelan VAR karet ini digunakan enam variabel, yaitu produksi karet (areal), produksi karet (Produksi), harga sheet (HRsheet), harga karet dunia (HKaretworld), volume eskpor karet (volekkaret), dan volume impor (volimkaret). Selain komposisi variabel tersebut, komponen konstanta dan trend juga perlu dilakukan pengujian untuk mengetahui signifikan atau tidak keberadaannya. Jika keduanya signifikan, maka komponen tersebut harus dimasukkan ke dalam model VAR dengan type “both”. Jika hanya konstanta yang signifikan, maka trend perlu dikeluarkan dari model VAR dengan model VAR type “const”. Jika hanya trend yang signifikan maka konstanta dikeluarkan dari model menggunakan model VAR type “trend”, dan jika keduanya tidak signifikan, maka type yang digunakan model VAR adalah “none”.

Keberadaan konstanta dan trend dapat dideteksi dari plot data awal, namun terkadang hal tersebut sulit dilakukan. Untuk mengantisipasi hal tersebut, sebaiknya dilakukan uji coba/trial and error menggunakan model VAR dari lag  $p=1$  sampai dengan lag  $p=5$  dengan type “both” untuk mengetahui signifikan atau tidak keberadaannya. Untuk panjang lag maksimum bisa dilakukan *trial error* sampai tidak memungkinkan untuk dilakukan permodelan.

Untuk data karet ini, setelah dilakukan running model VAR dengan lag  $p=1$  type “both”, diperoleh informasi komponen konstanta dan trend tidak signifikan, jumlah variabel yang signifikan ada 2 variabel dari total 6 variabel dalam system atau jumlah yg jumlah variabel yang signifikan (33,3%). Selanjutnya dilakukan uji VAR(2) type both, model ini menghasilkan konstanta dan trend yang signifikan, dan ada 4 variabel yang signifikan dari total 12 variabel dalam model (33,3%). Untuk model VAR(3) type both, model ini menghasilkan trend dan konstanta yang signifikan, dan ada 1 variabel yang signifikan dari total 18 variabel dalam model (5,55%). Untuk model VAR(4) type both, model ini menghasilkan konstanta dan trend yang tidak signifikan, tetapi ada 9 variabel yang signifikan dari total 24 variabel dalam model (37,5%). Selanjutnya model VAR(1) model VAR(2) type

constant, model ini menghasilkan konstanta yang tidak signifikan, dan masing-masing hanya ada 2 variabel yang signifikan. Oleh karena untuk type = constant, ternyata konstanta tidak signifikan maka dihentikan pada VAR(3) type=constant (Tabel 33).

Pemilihan lag p ditentukan dengan melihat banyaknya variabel yang signifikan dalam lag tersebut sekaligus memastikan harus ada variabel/peubah yang signifikan pada lag terpilih dimaksud. Setelah dilakukan run model ternyata ada 2 kandidat model VAR terbaik yaitu VAR (2) type “both” dan VAR(4) type “both“. Untuk menguji dua model VAR yang terbaik maka dilakukan pengujian MAPE untuk data training dan MAPE data testing.

Tabel 36. Hasil Pengujian Model VAR pada Beberapa Tingkat Lag p dan Type

Lag (p)	Type	Signifikansi Type	Jumlah Variabel Signifikan	Jumlah Total Variabel	Adj-R <sup>2</sup>
p=1	both	Const dan Trend Tidak Signifikan	2	6	98,49%
p=2	both	Const dan Trend Signifikan	4	12	98,70%
p=3	both	Const dan Trend Signifikan	1	18	98,39%
p=4	both	Const dan Trend Tidak Signifikan	9	24	99,02%
p=1	Constant	Const Tidak Signifikan	2	6	98,47%
p=2	Constant	Const Tidak Signifikan	2	12	98,48%
p=3	Constant	Const Tidak Signifikan	1	18	98,16%

Hasil pengujian besaran MAPE untuk 3 kandidat model VAR terbaik seperti pada Tabel 37, menunjukkan bahwa untuk data training yang memiliki MAPE terkecil adalah VAR(4) type “both”, sedangkan untuk MAPE data testing yang paling kecil adalah VAR(2) type “both”. Oleh karena tujuan peramalan adalah untuk menghasilkan angka yang akurat maka model terbaik yang terpilih adalah model VAR(2) type both, karena menghasilkan MAPE data testing terkecil, meskipun demikian perlu juga dicoba untuk VAR(2) type = both karena menghasilkan MAPE yang cukup kecil baik untuk data testing maupun training.

Tabel 37. Perbandingan MAPE Pengujian Model VAR

Lag (p)	Type	MAPE Training	MAPE Testing
p=2	both	1,10	6,02
p=4	both	0,64	6,46

## Model VAR(2) Type=both

Untuk model VAR kandidat terbaik adalah Model VAR (p=2) type=both. Model VAR(1) type=both bukan termasuk kandidat terbaik karena komponen constanta dan trend tidak signifikan. Hasil model VAR(2) type both, untuk mengestimasi produksi (t) karet menunjukkan bahwa koefisien constanta menunjukkan hasil yang signifikan pada tingkat kepercayaan 90% dan koefisien trend menunjukkan signifikan pada tingkat 95%. Untuk variabel lain yang signifikan untuk mengestimasi produksi (t) antara lain produksi lag1 (signifikan 99%), harga karet dunia lag 2 (95%), dan volume ekspor karet lag 2 (95%), volume impor karet lag 2 (90%).

Model VAR (2) type both ini menghasilkan nilai Adjusted R Square = 98,70%, artinya keragaman produksi dipengaruhi oleh variabel-variabel penjelasnya sebesar 98,70%. Nilai F hitung = 245,6, sehingga nilai p-value untuk model produksi ini sangat kecil atau jauh lebih kecil dari 0,05, sehingga model layak untuk digunakan.

Tabel 35. Output Model VAR(2) type=both

```

VAR Estimation Results:
=====
Endogenous variables: Produksi, Areal, Hsheet, HKaretworld, volekkaret, volimkaret
Deterministic variables: both
Sample size: 43
Log Likelihood: -2385.864
Roots of the characteristic polynomial:
0.9666 0.8631 0.8631 0.8247 0.8247 0.6564 0.6564 0.6025 0.6025 0.3774 0.3774 0.3645
Call:
VAR(y = karet[1:45, c(2, 3, 4, 6, 7, 8)], p = 2, type = "both")

Estimation results for equation Produksi:
=====
Produksi = Produksi.l1 + Areal.l1 + Hsheet.l1 + HKaretworld.l1 + volekkaret.l1 + Vo
limkaret.l1 + Produksi.l2 + Areal.l2 + Hsheet.l2 + HKaretworld.l2 + volekkaret.l2 +
volimkaret.l2 + const + trend

      Estimate Std. Error t value Pr(>|t|)
Produksi.l1      8.247e-01  1.970e-01  4.187  0.00024 ***
Areal.l1         8.585e-02  2.165e-01  0.396  0.69465
Hsheet.l1        3.383e-01  2.416e+00  0.140  0.88963
HKaretworld.l1  -4.340e+04  2.805e+04  -1.547  0.13261
Volekkaret.l1    3.023e-02  1.696e-01  0.178  0.85975
Volimkaret.l1   -3.644e+00  2.162e+00  -1.686  0.10259
Produksi.l2      8.042e-02  2.002e-01  0.402  0.69079
Areal.l2         3.450e-02  1.702e-01  0.203  0.84081
Hsheet.l2       -2.707e-01  2.735e+00  -0.099  0.92184
HKaretworld.l2  7.633e+04  2.851e+04  2.677  0.01209 *
Volekkaret.l2   -3.033e-01  1.438e-01  -2.110  0.04365 *
Volimkaret.l2   -4.128e+00  2.107e+00  -1.959  0.05976 .
const           2.676e+05  1.388e+05  1.929  0.06357 .
trend           1.032e+04  4.205e+03  2.453  0.02041 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 53460 on 29 degrees of freedom
Multiple R-Squared:  0.991,    Adjusted R-squared:  0.987
F-statistic: 245.6 on 13 and 29 DF,  p-value: < 2.2e-16

```

## Pengujian Asumsi VAR(2) type 'both'

Selanjutnya dilakukan pengujian asumsi non autokorelasi, normalitas, dan homoskedastisitas pada sisaan model VAR terbaik. Untuk data karet akan dilakukan pengujian sisaan pada dua model terbaik VAR (2) type 'both'.

Pemeriksaan autokorelasi residual model menggunakan fungsi "serial.test" yang di dalamnya dilakukan pengujian Portmanteau-and Breusch-Godfrey test. Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi. Pengujian Jarque-Bera tests untuk menguji kenormalan, hasil pengujian menunjukkan Nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal. Pemeriksaan heteroskedastisitas model menggunakan fungsi "arch.test" yang di dalamnya dilakukan pengujian ARCH-LM tests. Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.

Tabel 36. Ouput Pengujian Asumsi VAR(2) type=both

```
Portmanteau Test (asymptotic)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 545.13, df = 504, p-value = 0.0998

$JB

      JB-Test (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 129.65, df = 12, p-value < 2.2e-16

$Skewness

      Skewness only (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 22.191, df = 6, p-value = 0.001118

$Kurtosis

      Kurtosis only (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 107.46, df = 6, p-value < 2.2e-16

      ARCH (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 777, df = 2646, p-value = 1
```

## Ramalan Data Training, Testing, Penghitungan MAPE, dan Plot

Selanjutnya dilakukan peramalan data, baik untuk data training maupun untuk data testing sekaligus dilakukan penghitungan MAPE. MAPE (Mean Absolute Percentage Error) merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan data hasil peramalan. Untuk menguji suatu model lebih baik dengan model yang lain, maka dilakukan pengujian model dengan membandingkan Nilai MAPE baik untuk data training maupun data testing. Data Testing hasil ramalan produksi dengan VAR(2) type=both, menghasilkan MAPE =6,02%. Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data aktual dengan data hasil ramalan pada data testing adalah 6,02%. Data training hasil ramalan produksi dengan nilai p=2 type=both menghasilkan MAPE =1,10%. Model VAR ini menunjukkan ketika menggunakan data training sangat baik, terlihat dari MAPE yang kecil yaitu hanya sebesar 1,10%,



namun ketika digunakan untuk melakukan estimasi maka MAPE melonjak menjadi 6,02%, artinya kemampuan dalam meramalkan tidak sebaik data training.

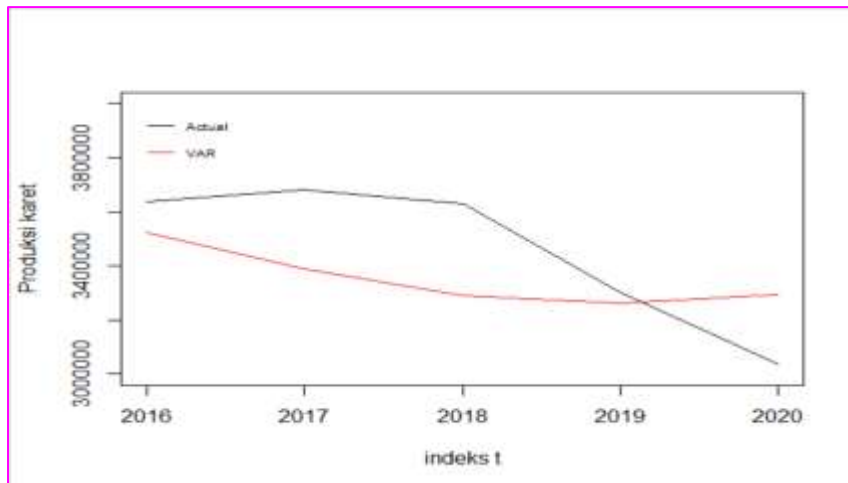
Tabel 37. Pengujian Nilai MAPE untuk Model VAR(2) type=both

MAPE TESTING						
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
1.220	3.138	7.908	6.020	8.458	9.378	1

MAPE TRAINING						
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
0.05785	0.44677	0.93614	1.09535	1.56978	3.21314	

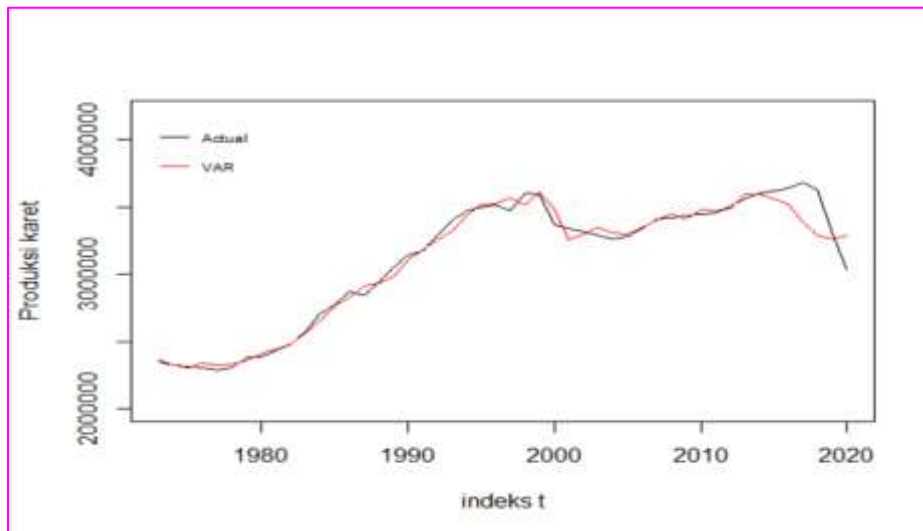
Dari grafik di bawah dapat dilihat bahwa pergerakan ramalan pada data testing mengikuti pergerakan data asli/actual, yang berbeda hanya tahun 2020 dimana nilai actual turun tetapi angka estimasi naik. Sehingga mungkin model VAR (2) type “both” kemampuan dalam meramalkan perlu dipertimbangkan untuk digunakan. Hasil peramalan dengan model VAR (2) ini cenderung lebih rendah dari data aktual. Pada tahun 2016 data actual sedikit lebih tinggi dari data estimasi dengan VAR (2), pada tahun 2017 dan 2018 beda antara data aktual dan data estimasi dengan VAR(2) ini semakin agak lebar, tetapi tahun 2019 antara data aktual dan hasil estimasi hamper berimpit, jadi sangat baik. Untuk tahun 2020 perbedaan antara data estimasi dan data aktual semakin lebar lagi, tetapi hasil estimasi lebih tinggi dari data aktual. Untuk data testing ini rata-rata penyimpangannya adalah sebesar 6,02%. Dari segi besaran MAPE sebenarnya model ini sudah cukup baik karena MAPE masih dibawah 10%.



Gambar 18. Plot Ramalan dan Aktual Data Testing VAR (2) Type “both”

Jika plot antara data testing dan data training digabungkan maka bentuk plotnya seperti Gambar 18. Untuk data tahun 1971 - 2015 atau data training plot sangat baik, karena antara data aktual dan estimasi dengan model VAR (2) hampir selalu berimpit plotnya, sehingga MAPE akan kecil. MAPE hasil pengujian untuk data training adalah sebesar 1,10%, suatu nilai yang cukup kecil karena rata-rata penyimpangan hanya 1,10%, artinya model cukup akurat. Namun plot tahun 2016 – 2020 menunjukkan data aktual dengan data estimasi dengan Model VAR (2) ini mulai

agak renggang, dimana hasil estimasi rata-rata lebih rendah dari data aktual. Hasil MAPE data testing ini, menunjukkan nilai yang lebih besar dari MAPE training yaitu sebesar 6,02%.



Gambar 19. Plot Ramalan dan Aktual Data Training dan Testing VAR (2) Type “both”

### Model VAR(4) Type=both

#### Ramalan Data Training, Testing, Penghitungan MAPE, dan Plot

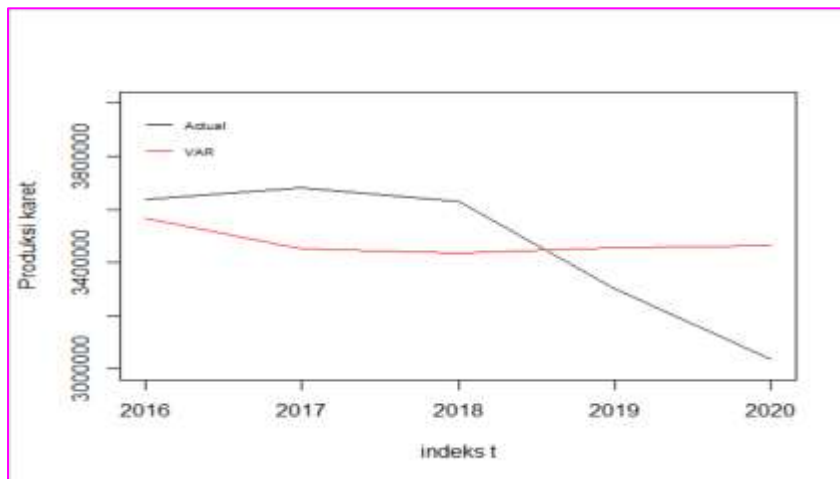
Selanjutnya dilakukan peramalan data, baik untuk data training maupun untuk data testing sekaligus dilakukan penghitungan MAPE. MAPE (Mean Absolute Percentage Error) merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan data hasil peramalan. Untuk menguji suatu model lebih baik dengan model yang lain, maka dilakukan pengujian model dengan membandingkan Nilai MAPE baik untuk data training maupun data testing. Data Testing hasil ramalan produksi dengan VAR(4) type=both, menghasilkan MAPE =6,45%. Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data aktual dengan data hasil ramalan pada data testing adalah 6,45%. Data training hasil ramalan produksi dengan nilai p=4 type=both menghasilkan MAPE =0,64%. Model VAR ini menunjukkan ketika menggunakan data training sangat baik, terlihat dari MAPE yang kecil yaitu hanya sebesar 0,64%, namun ketika digunakan untuk melakukan estimasi maka MAPE meningkat menjadi 6,45%. Jika dibandingkan dengan model VAR(2) type=bpth, maka model VAR(4) type=both menghasilkan MAPE yang lebih kecil untuk data training, tetapi untuk data testing sepertinya model VAR(2) both ini lebih baik.

Tabel 39. Pengujian Nilai MAPE untuk Model VAR(4) type=both

MAPE TESTING						
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	NA's
2.007	4.686	5.346	6.457	6.214	14.034	1

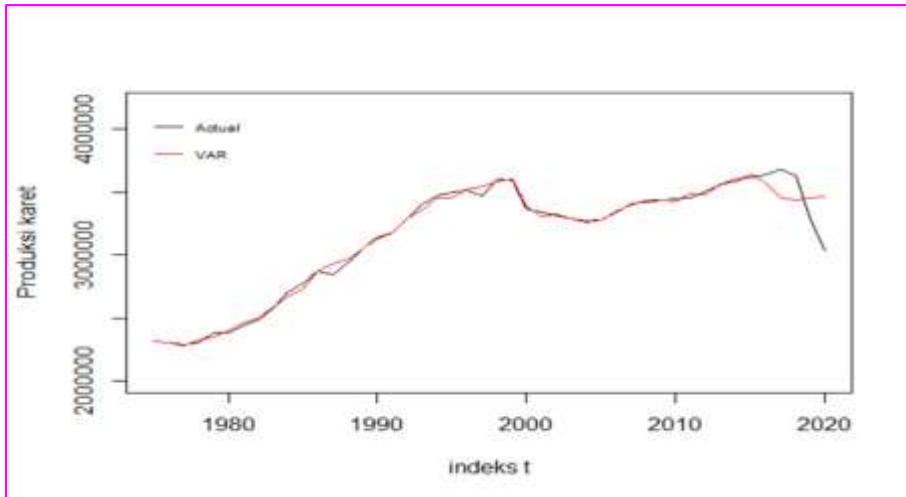
MAPE TRAINING					
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.01293	0.19380	0.48838	0.64171	0.86553	2.97966

Dari grafik di bawah dapat dilihat bahwa pergerakan ramalan pada data testing mengikuti pergerakan data asli/actual, data actual bisa lebih tinggi atau bisa juga lebih rendah dari nilai ramalan, maka sempat terjadi perpotongan antara data actual dan data ramalan (Gambar 20). Oleh karena menghasilkan MAPE yang cukup kecil untuk data testing maka model VAR (4) type “both” untuk peramalan perlu dipertimbangkan. Hasil peramalan dengan model VAR (2) ini bisa lebih rendah atau lebih tinggi dari data aktual. Pada tahun 2016, 2017 dan 2018 data aktual sedikit lebih tinggi dari data estimasi dengan VAR (4) both, pada tahun 2019 dan 2020 data aktual sedikit lebih rendah dari data estimasi.



Gambar 20. Plot Ramalan dan Aktual Data Testing VAR (4) Type “both”

Jika plot antara data testing dan data training digabungkan maka bentuk plotnya seperti Gambar 21. Untuk data tahun 1971 - 2015 atau data training plot sangat baik, karena antara data actual dan estimasi dengan model VAR (4) hampir selalu berimpit plotnya, sehingga MAPE akan kecil. MAPE hasil pengujian untuk data training adalah sebesar 0,64%, suatu nilai yang cukup kecil karena rata-rata penyimpangan hanya 0,64%, artinya model cukup akurat. Pada saat plot tahun 2016 – 2020 menunjukkan data aktual dengan data estimasi dengan model VAR (4) ini mulai agak renggang, dimana hasil estimasi dapat lebih rendah atau lebih tinggi dari data aktual. Hasil MAPE data testing ini, menunjukkan nilai yang lebih besar dari MAPE training yaitu sebesar 6,45%.



Gambar 21. Plot Ramalan dan Aktual Data Training dan Testing VAR (4) Type “both”

### Pengepasan Model untuk Seluruh Data dan Peramalannya

Ada 2 Model VAR tentative terbaik dari hasil perbandingan nilai MAPE yaitu model VAR (2) type “both” dan VAR(4) type “both, sehingga dapat dilanjutkan pengepasan model pada keseluruhan data. Setelah dilakukan pengepasan model, maka dilakukan estimasi produksi karet untuk 5 tahun kedepan yaitu tahun 2021 -2025.

Peramalan dengan model VAR (2) type “both” menggunakan keseluruhan data menunjukkan produksi karet tahun 2021 diestimasi sebesar 2,83 juta ton, tetapi turun pada tahun 2022 menjadi 2,74 juta ton, terus turun meningkat sehingga pada tahun 2025 produksi karet diperkirakan mencapai 2,53 juta ton. Hasil run model untuk VAR(4) type “both” menghasilkan angka estimasi yang cenderung terus turun sangat tajam, maka tahun 2021 produksi karet diperkirakan 2,82 juta ton, tahun 2022 turun sedikit menjadi 2,21 juta ton, tahun 2023 turun menjadi 1,57 juta ton, tahun 2024 kembali turun tajam menjadi 984 ribu ton, dan akhirnya tahun 2025 kembali turun menjadi 200 ribu ton saja. Hasil peramalan menggunakan Model VAR(4) type=both, kurang realistis terutama untuk tahun 2022 – 2025, sehingga dipertimbangkan untuk tidak menggunakan model ini.

Tabel 40. Hasil Peramalan Produksi Karet Tahun 2021-2025, Model VAR(2) Type=both

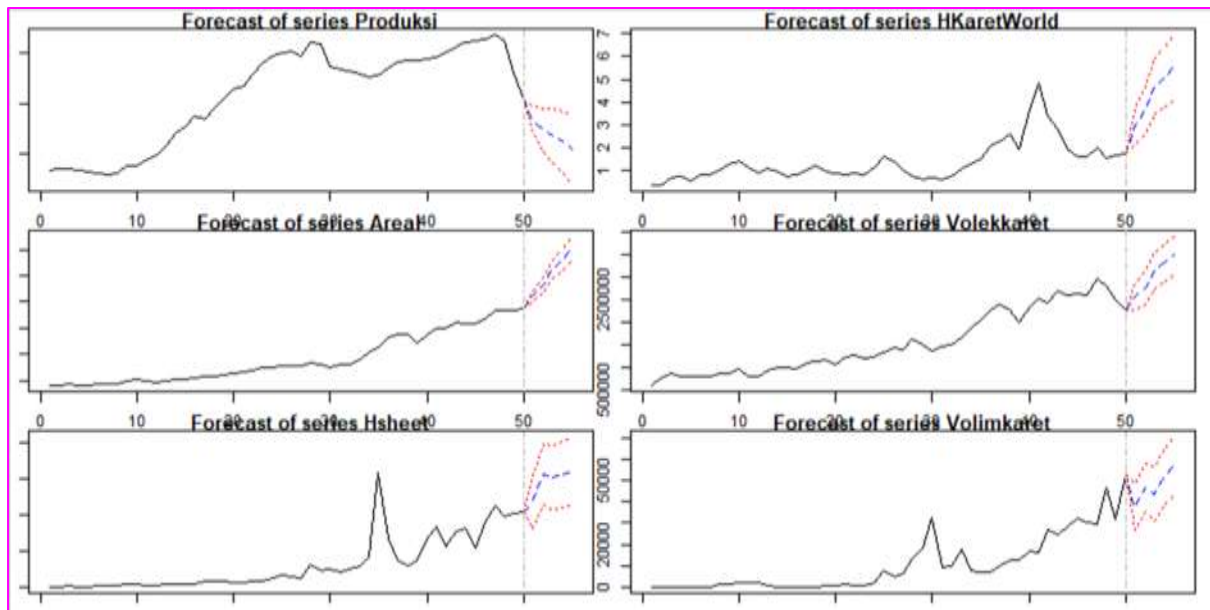
[1] 2835469 2736057 2675123 2614292 2537725

Tabel 41. Hasil Peramalan Produksi Karet Tahun 2021-2025, Model VAR(4) Type=both

[1] 2822290.0 2215947.4 1572909.7 984603.8 200438.7

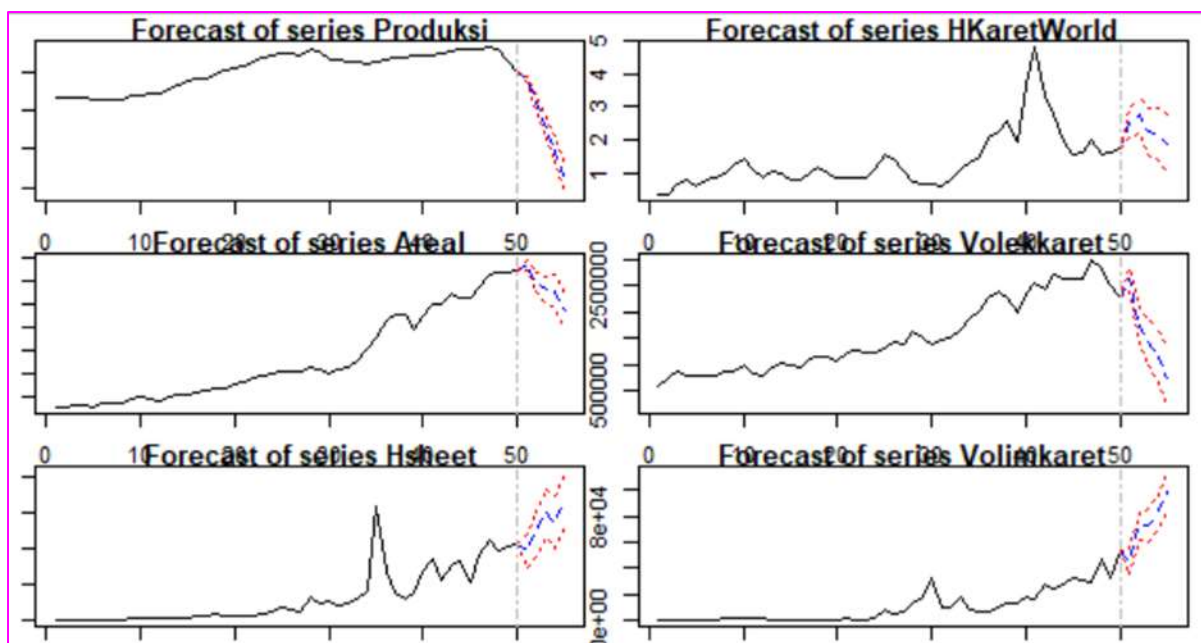
Keunggulan model VAR ini pada saat dilakukan peramalan maka semua variabel dalam system akan ikut diramalkan. Hasil peramalan bukan hanya target variabel, tetapi semua variabel yang masuk dalam sistem. Untuk Model VAR (2) both, hasil peramalan 5 tahun ke depan untuk

variabel produksi menunjukkan produksi akan terus menurun secara linier. Sebaliknya untuk cenderung terus naik dalam 5 tahun ke depan, hal ini menunjukkan bahwa ada dugaan produktivitas karet cenderung turun. Sebaliknya harga karet sheet dalam negeri cenderung akan naik tetapi fluktuatif, sementara harga karet dunia diperkirakan akan perlahan terus meningkat kemudian sedikit menurun kembali. Volume ekspor karet cenderung meningkat pada tahun 2021-2025, sebaliknya volume impor karet berfluktuasi tetapi cenderung naik.



Gambar 22. Hasil Estimasi untuk Semua Variabel dalam Sistem Model VAR(2) Both

Untuk Model VAR (4) both, hasil peramalan 5 tahun ke depan untuk variabel produksi menunjukkan produksi akan cenderung menurun secara tajam. Demikian juga untuk luas areal akan turun secara tajam dalam 5 tahun ke depan. Sebaliknya harga karet sheet dalam negeri diperkirakan berfluktuasi tetapi ada kecenderungan meningkat, sementara harga karet dunia diperkirakan juga fluktuatif, ada peningkatan 2 tahun ke depan tetapi kemudian turun kembali. Volume ekspor karet cenderung turun karena harga karet dunia yang masih turun, sebaliknya volume impor karet berfluktuasi tetapi cenderung meningkat tajam.



Gambar 23. Hasil Estimasi untuk Semua Variabel dalam Sistem Model VAR(4) Both

Untuk model VAR (2) type “both”, pergerakan hasil ramalan produksi karet cenderung mengalami penurunan dengan rata-rata penurunan produksi tahun 2021 -2025 sebesar 3,52% per tahun. Untuk model VAR(4) type “both” pergerakan hasil ramalan produksi karet menurun tajam dengan rata-rata penurunan sebesar 34,84% per tahun. Dari data historisnya pertumbuhan produksi karet setiap tahun paling tinggi 16%, dan penurunan paling tinggi 13%. Pertumbuhan di atas 4% akan sulit dicapai mengingat kondisi saat ini tidak ada penambahan dari ekstensifikasi yang cukup signifikan. Pada saat ini ada tend penurunan produksi karet nasional. Rata-rata pertumbuhan produksi karet dengan Model VAR (2) both, sepertinya lebih realistis dibandingkan model VAR(4) both.

Jika berdasarkan nilai MAPE, maka model terbaik yang terpilih untuk peramalan data produksi karet adalah model VAR (2) type “both” karena memberikan nilai MAPE testing yang lebih kecil sekaligus memberikan plot ramalan yang lebih lambat penurunannya mengikuti pola harga karet dunia. Model VAR(4) both meskipun menghasilkan MAPE data training dan testing yang cukup baik, tetapi hasil peramalan menunjukkan lonjakan penurunan produksi yang kurang realistis.

Untuk model VAR(2) both, jika dibandingkan dengan tahun 2020 maka produksi karet hasil estimasi tahun 2021 akan turun sebesar 6,65%, sehingga menjadi sekitar 2,83 juta ton. Pada tahun 2022 produksi karet masih turun sebesar 3,51% menjadi 2,73 juta ton. Hal ini didukung dari realitas produksi karet yang hanya sedikit saja mengalami peningkatan, karena harga karet dunia yang masih cenderung turun. Penurunan produksi karet hasil estimasi selama 5 tahun kedepan berkisar antara 2,23% sampai 6,65% per tahun. Penurunan ini bisa saja terjadi karena beberapa tahun belakangan ini harga karet alam dunia cenderung turun dan stabil di harga rendah. Beberapa petani karet telah mengkonversi lahannya ke tanaman yang lebih menguntungkan seperti kelapa sawit atau tanaman perkebunan lainnya.

Tabel 43. Hasil Estimasi Produksi Karet Nasional dengan Model VAR(2) dan VAR(4) Both

Tahun	Model VAR(2) Both	Pertumbuhan (%)	Model VAR(4) Both	Pertumbuhan (%)
2020	3,037,348		3,037,348	
2021	2,835,469	(6.65)	2,822,290	(7.08)
2022	2,736,057	(3.51)	2,215,947	(21.48)
2023	2,675,123	(2.23)	1,572,909	(29.02)
2024	2,614,292	(2.27)	984,603	(37.40)
2025	2,537,725	(2.93)	204,438	(79.24)
<b>Rata-rata Pertumbuhan</b>		<b>(3.52)</b>		<b>(34.84)</b>

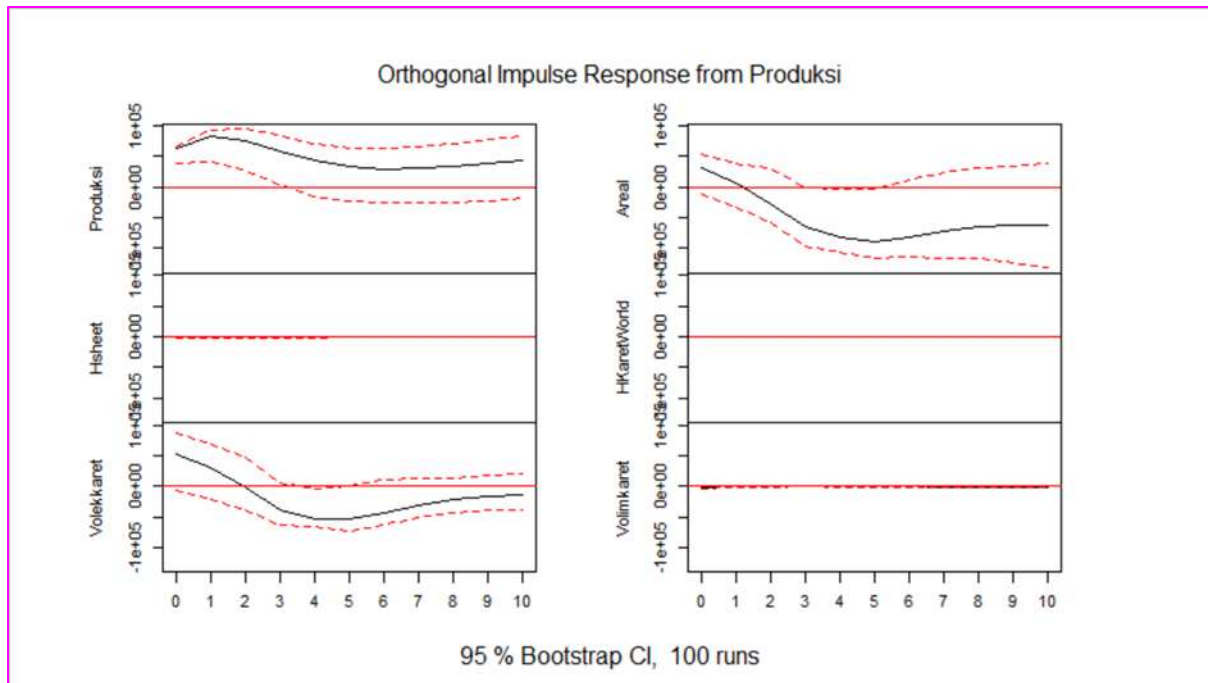
Keterangan : Tahun 2020 Angka Tetap Ditjenbun  
Tahun 2021 – 2025 : Estimasi Berdasarkan Model

### Interpretasi Impulse Response Function (IRF) dan Variance Decomposition

Setelah diperoleh model terbaik, untuk permodelan VAR dapat dilakukan interpretasi lebih mendalam terkait Impulse Response Function (IRF) dan Variance Decomposition. Berikut interpretasi ketiga hal di atas untuk model terbaik VAR (2) type “both”. **Impulse Response Function** akan menjelaskan bagaimana perubahan atau gejolak yang terjadi pada suatu variabel/peubah di tahun tertentu akan berdampak pada variabel/peubah lain di tahun tertentu dan tahun-tahun setelahnya.

Dari grafik Impulse Response Function produksi model terbaik VAR (2) type “both”, dapat dilihat bahwa jika terjadi perubahan pada produksi di tahun tertentu maka akan berdampak pada produksi itu sendiri sampai 3 tahun ke depan. Dampak tersebut akan hilang setelah 3 tahun. Sedangkan dampak perubahan produksi tidak berdampak pada harga sheet, harga karet dunia, dan volume impor karet, namun masih sedikit berdampak pada volume ekspor kedepan dan luas areal. Perubahan produksi berdampak pada volume ekspor dan luas areal sampai dengan satu tahun dan akan hilang dampaknya setelah melewati tahun pertama tersebut, sedangkan pada produksi dampaknya terjadi sampai dengan 3 tahun ke depan.

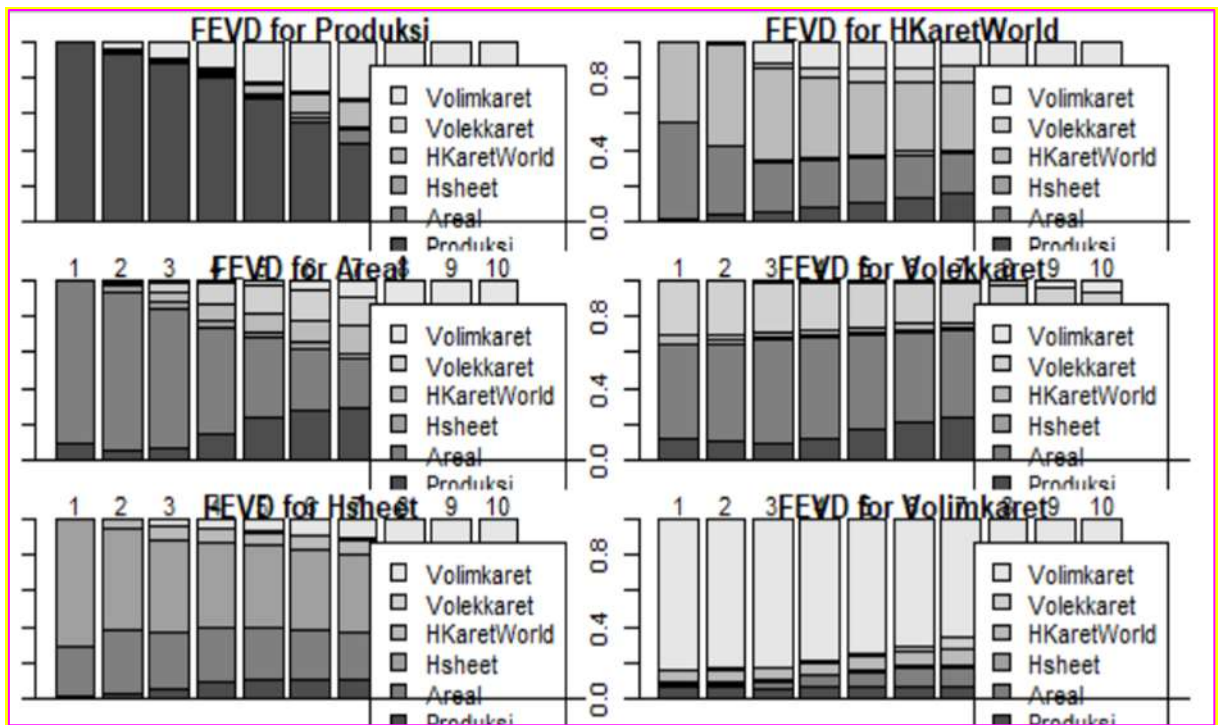




Gambar 24. Impuls Respon Beberapa Variabel Terhadap Produksi Model VAR(2) Both

**Variance Decomposition** atau dekomposisi keragaman digunakan untuk melihat variabel apa saja yang mempengaruhi komposisi keragaman suatu variabel/peubah. Dari grafik dekomposisi keragaman model terbaik VAR (2) type “both” di atas dapat dilihat bahwa komposisi produksi pada tahun pertama dipengaruhi sepenuhnya oleh produksi itu sendiri. Pada tahun kedua, komposisi produksi 95% dipengaruhi oleh produksi itu sendiri, sekitar 3% dipengaruhi oleh luas areal, dan sekitar 2% dipengaruhi oleh volume ekspor. Pada tahun ketiga, komposisi produksi dipengaruhi sekitar 90% oleh produksi itu sendiri, sekitar 6% dipengaruhi oleh luas areal, dan 2% dipengaruhi oleh harga karet dunia dan sekitar 2% oleh volume ekspor. Semakin bertambahnya tahun, pengaruh produksi karet terhadap keragaman produksi karet sendiri semakin berkurang diikuti makin tingginya pengaruh volume ekspor, volume impor, harga karet dunia, harga karet sheet local dan produksi.





Gambar 25. Dekomposisi Keragaman Beberapa Variabel Untuk Model VAR(2) Both

## H. Pemilihan Model Terbaik Estimasi Produksi Karet Nasional

Salah satu dasar penentuan model terbaik dapat dilakukan dengan membandingkan nilai MAPE untuk data testing dan training yaitu dengan memilih nilai MAPE yang paling kecil, terutama untuk data testing. Selain MAPE yang terkecil, pola pergerakan ramalan juga harus diperhatikan. Pilihlah plot yang paling berhimpit/bersesuaian dengan data asli/aktual atau dengan kata lain performa hasil ramalan seiring dengan data historisnya.

Berdasarkan data historis yang ada produksi karet nasional berfluktuasi, produksi tahun 2017 sebesar 3,68 juta ton atau naik 9,6%. Pada tahun 2018 produksi karet nasional turun sebesar 1,36%, sehingga produksi karet tahun 2018 menjadi sebesar 3,63 juta ton. Pada tahun 2019 produksi karet nasional kembali turun tajam sebesar 9,06%, kemudian pada tahun 2020 kembali turun sebesar 8,00%. Rata-rata penurunan produksi karet nasional selama 5 tahun terakhir atau tahun 2016 – 2020 sebesar -2,20% per tahun.

Tabel 40. Produksi Karet Nasional Tahun 2016 – 2020

Tahun	Produksi (Ha)	Pertumbuhan (%)
2016	3,357,951	
2017	3,680,428	9.60
2018	3,630,357	(1.36)
2019	3,301,405	(9.06)
2020	3,037,348	(8.00)
Rata-rata Pertumbuhan		(2.20)

Sumber : Direktorat Jenderal Perkebunan

Untuk menyusun angka estimasi produksi karet telah dilakukan uji coba dengan 3 (tiga) model. Model yang pertama adalah model time series atau ARIMA, model terbaik untuk ARIMA adalah pada orde ARIMA (2,1,0) dan ARIMA (1,2,3). Untuk model estimasi produksi karet nasional dengan ARIMA (2,1,0) menghasilkan MAPE untuk data training sebesar 4,74% dan MAPE untuk data testing sebesar 7,19%. Hal ini menunjukkan bahwa kemampuan untuk melakukan estimasi dengan model ARIMA ini rata-rata akan mengalami kesalahan sekitar 7,19% lebih tinggi atau 7,19% lebih rendah. Hasil estimasi dengan model ARIMA (2,1,0) pertumbuhan 5 tahun kedepan relatif lambat, yaitu hanya 0,24%/tahun. Hal ini mendekati dengan data 5 tahun ke belakang (2016 – 2020), dimana pertumbuhan produksi rata-rata turun 2,2% per tahun.

Model ARIMA (1,2,3) patut dipertimbangkan juga sebagai model tentative terbaik karena menghasilkan MAPE untuk data training sebesar 4,15% dan MAPE untuk data testing sebesar 7,30%. Model ARIMA (1,2,3) menghasilkan MAPE data testing yang sedikit lebih besar dari ARIMA (2,1,0) artinya kemampuan untuk mengestimasi lebih baik ARIMA (2,1,0). Hasil estimasi dengan model ARIMA(1,2,3) pertumbuhan 5 tahun kedepan relatif lebih cepat, yaitu mencapai 2,09%/tahun. Pertumbuhan ini lebih tinggi dibandingkan dengan data 5 tahun ke belakang (2016 – 2020), dimana pertumbuhan produksi mencapai – 2,20% per tahun.

Metode estimasi yang kedua adalah dengan model fungsi transfer, untuk melakukan estimasi produksi karet dengan variabel bebas adalah harga karet dunia. Untuk model fungsi transfer ARIMA (2,1,0) ini menghasilkan MAPE data training 3,67%, sementara untuk MAPE data testing sebesar 4,26%. Sementara untuk model Fungsi Transfer ARIMA (0,1,1) menghasilkan MAPE data training yang lebih kecil yaitu sebesar 3,40% dan MAPE testing yang lebih besar yaitu 4,43%. Model fungsi transfer ini menghasilkan MAPE yang lebih kecil dibandingkan model ARIMA meskipun perbedaan tidak terlalu signifikan, sehingga model fungsi transfer lebih akurat dalam melakukan estimasi. Hasil estimasi juga menunjukkan angka yang lebih realistis, dengan angka estimasi tahun 2021 sebesar 2,919 juta ton, atau turun 3,89%. SEmentara jika menggunakan model Fungsi Transfer Arima (0,1,1) hasil estimasi tahun 2021 produksi karet mencapai 2,972 jut ton atau turun 2,13% dari tahun 2020. Disamping itu untuk estimasi 5 tahun kedepan jika menggunakan Fungsi Transfer Arima (2,1,0) angka pertumbuhan rata-rata sebesar 1,06%/tahun, jika menggunakan Fungsi Transfer Arima (0,1,1) angka pertumbuhan rata-rata sebesar 0,07%/tahun, sementara angka pertumbuhan 5 tahun sebelumnya sebesar -2,20%/tahun, jadi model Fungsi Transfer Arima (2,1,0) pertumbuhan lebih mendekati data historisnya.

Tabel 41. Perbandingan Hasil Estimasi dan MAPE Model Arima, Fungsi Transfer dan VAR

PRODUKSI		Model ARIMA				Fungsi Transfer				Model VAR			
		ARIMA (2,1,0)		ARIMA (1,2,1)		Arima (2,1,0) xreg=hrkworld		Arima (0,1,1) xreg=hrkworld		VAR (2) type=both		VAR (4) type=both	
			(%)		(%)		(%)		(%)		(%)		(%)
	MAPE Training	4.74		4.15		3.67		3.40		1.10		0.64	
	MAPE Testing	7.19		7.30		4.26		4.43		6.02		6.46	
ATAP	2016	3,357,951		3,357,951		3,357,951		3,357,951		3,357,951		3,357,951	
	2017	3,680,428	9.60	3,680,428	9.60	3,680,428	9.60	3,680,428	9.60	3,680,428	9.60	3,680,428	
	2018	3,630,357	(1.36)	3,630,357	(1.36)	3,630,357	(1.36)	3,630,357	(1.36)	3,630,357	(1.36)	3,630,357	
	2019	3,301,405	(9.06)	3,301,405	(9.06)	3,301,405	(9.06)	3,301,405	(9.06)	3,301,405	(9.06)	3,301,405	
	2020	3,037,348	(8.00)	3,037,348	(8.00)	3,037,348	(8.00)	3,037,348	(8.00)	3,037,348	(8.00)	3,037,348	
Angka Estimasi (AESTI)	2021	2,974,904	(2.06)	3,045,340	0.26	2,919,246	(3.89)	2,972,801	(2.13)	2,835,469	(6.65)	2,822,290	(7.08)
	2022	2,986,303	0.38	3,135,344	2.96	2,962,994	1.50	2,990,268	0.59	2,736,057	(3.51)	2,215,947	(21.48)
	2023	2,999,765	0.45	3,202,055	2.13	3,005,103	1.42	2,577,935	(0.41)	2,675,123	(2.23)	1,572,909	(29.02)
	2024	3,003,702	0.13	3,257,738	1.74	3,045,418	1.34	2,986,643	0.29	2,614,292	(2.27)	984,603	(37.40)
	2025	3,003,432	(0.01)	3,308,200	1.55	3,044,920	(0.02)	2,980,495	(0.21)	2,537,725	(2.93)	204,438	(79.24)
Rata-rata Pertumbuhan ATAP 2016 - 2020			(2.20)		(2.20)		(2.20)		(2.20)		(2.20)		-2.20
	AESTI 2021 - 2025		0.24		2.09		1.06		0.07		(2.73)		-41.79

Untuk model estimasi yang terakhir adalah dengan model VAR (Vector Auto Regressive). Untuk model VAR ini menggunakan 6 variabel yaitu produksi, luas areal, harga karet sheet, harga karet dunia, volume ekspor dan volume impor karet. Model yang terbaik untuk Model VAR ada dua yaitu adalah nilai  $p=2$  dan  $type="both"$ ,  $p=2$  artinya menggunakan variabel bebas sampai lag-2 dan ada konstanta dan trend. Estimasi produksi karet dengan menggunakan model VAR(2) both ini menghasilkan ketelitian yang cukup tinggi yaitu MAPE untuk data training 1,10% dan MAPE untuk data testing 6,02%. MAPE untuk data testing ini model VAR lebih tinggi dibandingkan dengan model Fungsi Transfer, tetapi lebih rendah dibandingkan Model Arima. Jika dibandingkan angka pertumbuhan produksi karet model VAR(2) both antara hasil estimasi 5 tahun kedepan dengan rata-rata pertumbuhan -2,73% per tahun, mendekati dengan angka pertumbuhan 5 tahun terakhir yaitu sebesar -2,20% per tahun. Angka hasil estimasi untuk produksi karet nasional tahun 2021 sebesar 2,83 juta ton, sementara untuk angka tetap tahun 2020 sebesar 3,03 juta ton atau turun sebesar 6,65%.

Model VAR yang kedua yang perlu dipertimbangkan adalah VAR(4) both. Model VAR(4) both menghasilkan MAPE data training sebesar 0,64%, dan MAPE data testing 6,46%. Angka pertumbuhan rata-rata selama 5 tahun ke depan (2021 -2025) untuk model ini sebesar -41,72%/tahun atau terlampaui tinggi penurunannya dibandingkan dengan data historisnya (2016 – 2020) dengan rata-rata pertumbuhan -2,20%/tahun. Kelemahan dari model VAR(4) both ini hasil estimasi turun terlampaui dalam. Estimasi ini diduga under estimasi karena dari data historisnya penurunan produksi karet hanya sekitar 2,20% saja per tahun.

Berdasarkan Tabel 41 diatas, untuk data training dan data testing yang paling baik adalah yang memiliki MAPE terkecil. Untuk data training yang paling kecil adalah model VAR(4) both, sedangkan untuk data testing yang paling kecil adalah Fungsi Transfer ARIMA (2,1,0) dengan factor input harga karet dunia, dengan MAPE 4,26%. Oleh karena tujuan penyusunan model menghasilkan angka estimasi dengan kesalahan yang paling kecil, maka model yang paling kecil MAPE Testing adalah Model Fungsi Transfer ARIMA (2,1,0) dengan faktor input harga karet dunia merupakan model yang terbaik untuk menyusun angka Estimasi Produksi Karet Nasional. Disamping faktor MAPE, hasil estimasi 5 tahun kedepan (2021 – 2025) menunjukkan pertumbuhan rata-rata produksi karet nasional 1,06% per tahun, cukup mendekati dari data aktual 5 tahun kebelakang yaitu pertumbuhan produksi karet tahun 2016 – 2020 sebesar -2,20% per

tahun. Berdasarkan MAPE testing terkecil dan angka pertumbuhan produksi maka model Fungsi Transfer menjadi model terbaik untuk meramalkan produksi karet nasional.

Berdasarkan hasil kajian ini model terbaik berdasarkan MAPE testing terkecil yaitu Model Fungsi Transfer ARIMA (2,1,0). Namun demikian ada pertimbangan lain untuk memilih hasil estimasi yang paling tepat, misalnya intervensi program intensifikasi dan ekstensifikasi. Berdasarkan pertimbangan program yang telah dilakukan maka ada optimisme bahwa hasil peramalan dengan model ARIMA(1,2,3) menjadi model estimasi terpilih dengan angka estimasi tahun 2021 sebesar 3,045 juta ton dan estimasi tahun 2022 sebesar 3,135 juta ton.

## KESIMPULAN

Untuk meningkatkan akurasi dalam penyusunan angka estimasi, maka dilakukan pengembangan metode estimasi produksi karet nasional. Metode estimasi data perkebunan selama ini menggunakan model *Single Smoothing Exponential (SSE)* atau menggunakan *Double Smoothing Exponential (DSE)*. Meskipun dua metode tersebut dapat menghasilkan angka estimasi yang cukup baik, namun masih perlu melakukan pengembangan model alternatif yang diharapkan lebih akurat.

Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1970 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan. Untuk estimasi produksi karet alternatif model pertama adalah Model ARIMA. Model ARIMA terbaik adalah ARIMA (2,1,0), menghasilkan MAPE untuk data training 4,74%, dan MAPE data testing 7,19%. Model ARIMA (1,2,3) juga menghasilkan MAPE yang cukup baik, yaitu MAPE training 4,15% dan MAPE testing 7,30%. Untuk model yang kedua dengan menggunakan Fungsi Transfer ARIMA (2,1,0) dengan variabel input harga karet dunia, menghasilkan MAPE untuk data training sebesar 3,67% dan MAPE data testing 4,26%. Untuk model Fungsi Transfer lainnya adalah ARIMA (0,1,1) menghasilkan MAPE data training sebesar 3,40% dan MAPE data testing 4,43%. Untuk model yang keempat model VAR(2) type 'both' ada pengaruh konstanta dan trend, menghasilkan MAPE data training 1,10% dan data MAPE data testing 6,02%. Model tentatif VAR adalah VAR(4) type 'both', menghasilkan MAPE training 0,64% dan MAPE testing 6,46%.

Berdasarkan perbandingan besarnya MAPE baik data testing maupun data training dan hasil estimasi produksi 5 tahun kedepan, maka model terbaik yang terpilih adalah model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet dunia karena menghasilkan akurasi yang paling tinggi sehingga MAPE rata-rata data testing sebesar 4,26%. Hasil estimasi produksi karet nasional untuk model Fungsi Transfer ARIMA(2,1,0) dengan faktor input harga karet dunia untuk tahun 2021 sebesar 2.919.246 ton dan tahun 2022 sebesar 2.962.994 ton, tahun 2023 sebesar 3.005.103 ton, tahun 2024 sebesar 3.045.418 ton, dan tahun 2025 sebesar 3.044.920 ton. Laju pertumbuhan estimasi produksi karet nasional selama 5 tahun kedepan (2021 – 2025) rata-rata 1,06% per tahun. Namun jika mempertimbangkan program yang telah dijalankan oleh direktorat teknis terkait untuk meningkatkan areal karet maka ada optimisme estimasi produksi tahun 2021 sebesar 3.045.340 ton dan tahun 2022 sebesar 3.135.344 ton, yaitu hasil pemodelan dengan ARIMA (1,2,3).

## DAFTAR PUSTAKA

- Anonim. 2013. Pedoman Pelaksanaan Pengelolaan Data Komoditas Perkebunan (PDKP). Dirjen Perkebunan – Kementerian Pertanian.
- Enders, W. 2010. Applied Econometric Time Series. USA: University of Alabama. Wiley, Third Edition.
- Fitriani, D.R, Darsyah, M.Y., & Wasono, R. 2013. Peramalan Fungsi Transfer pada Harga Emas Pasar Komoditi. Seminar Nasional Pendidikan Sains dan Teknologi, Fakultas MIPA, Universitas Muhammadiyah Semarang.
- Guha, B and Bandyopadhyay, G. 2016. Gold Price Forecasting Using ARIMA Model. Journal of Advanced Management Science Vol. 4, No. 2, March 2016
- Gujarati, D.N. and D.C. Porter, 2010. Dasar-dasar Ekonometrika. Jakarta: Penerbit Salemba Empat, Buku 2, Edisi 5.
- Kementerian Pertanian. 2019. Statistik Perkebunan Indonesia 2018-2020 (Karet). Jakarta: Direktorat Jenderal Perkebunan.
- M. Firdaus 2020. Aplikasi Ekonometrika dengan E-Views, Stata dan R. Institut Pertanian Bogor (IPB) Press.
- Montgomery DC, Johnson LA & Gardiner JS. 1990. Forecasting and Time Series Analysis. Singapore:Mc-Graw Hill.
- Myers R. 1994. Classical And Modern Regression with Applications. Boston: PWS – KENT Publishing Company.
- Ryan TP. 1997. Modern Regression Methods. New York,USA: John Wiley & Sons, INC.



# **KAJIAN MODEL PERAMALAN PRODUKSI KOPI DI INDONESIA: PENDEKATAN MODEL ARIMA, FUNGSI TRANSFER DAN VAR**

Oleh: Roch Widaningsih

## **ABSTRAK**

Kopi merupakan tanaman perkebunan unggulan di Indonesia. Tahun 2020, nilai ekspor kopi mencapai 821,92 juta USD, menempati urutan ke lima setelah kelapa sawit, karet, kajeo dan kelapa. Angka tetap (ATAP) data produksi kopi oleh Badan Pusat Statistik (BPS) adalah data tahunan dengan lag  $n-1$  tahun. Guna merumuskan kebijakan produksi, pasokan dan distribusi komoditas kopi dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Tahun 2021, Pusdatin mengkaji ulang tiga metode yaitu metode ARIMA, fungsi transfer dan VAR (*Vector Auto Regression*) dengan menggunakan *analisis R Studio* guna melakukan pemodelan data kopi di Indonesia. Metode yang dipilih guna melakukan peramalan data produksi kopi adalah metode yang terbaik dikaji dari nilai MAPE (*Mean Absolute Percentage Error*) terkecil dan kerealistisan hasil peramalan. Berdasarkan dari hasil nilai MAPE disimpulkan ARIMA (4,1,3) adalah yang terbaik dengan MAPE data training 4,58%.

## **PENDAHULUAN**

Sebagai produsen kopi ketiga terbesar di dunia, Indonesia menempatkan kopi sebagai salah satu komoditas unggulan perkebunan. Tahun 2020, nilai ekspor kopi menempati urutan ke kelima komoditas terbesar di Indonesia setelah kelapa sawit, karet, kakao dan kelapa. Nilai ekspor kopi mencapai 0.82 Milyard USD.

Peluang industri kopi di Indonesia sangat bagus. Apalagi dengan semakin dikenalnya kopi Indonesia di Eropa dan Amerika, terutama kopi khusus (*specialty coffee*) seperti kopi Gayo, kopi Mandailing, kopi Lampung, kopi Bajawa dan lainnya. Menteri Pertanian. Andi Amran Sulaiman mempromosikan komoditas kopi Indonesia kepada dunia melalui forum World Coffee Producers Forum (WCPF) yang dihelat di Kolombia, Selasa (11/7/2017).

Fokus pemerintah saat ini adalah meningkatkan produksi komoditas kopi dalam negeri melalui penyediaan bibit berkualitas tinggi, pemupukan tepat waktu, manajemen air dan program peremajaan untuk mengganti tanaman kopi yang sudah tua. Pengembangan kopi ke depan juga akan lebih memperhatikan aspek kearifan lokal, sehingga dapat dihasilkan jenis-jenis kopi specialty yang bernilai tinggi dari berbagai daerah.

Ditjen Perkebunan dalam renstra menempatkan komoditas kopi menjadi salah satu komoditas yang menjadi sasaran pokok sub agenda prioritas peningkatan agroindustri yaitu

peningkatan produksi komoditas andalan dan prospektif ekspor serta mendorong perkembangan agroindustri di pedesaan, selain komoditas kelapa sawit, kakao, teh dan kelapa (Direktorat Jenderal Perkebunan, 2015).

ATAP data produksi kopi oleh Badan Pusat Statistik (BPS) adalah data tahunan dengan lag  $n-1$  tahun. Guna merumuskan kebijakan produksi, pasokan dan distribusi komoditas kopi sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Metode ramalan produksi kopi dapat digunakan dengan menggunakan beberapa metode statistik yang telah dikembangkan oleh para ahli.

Makalah ini akan dikaji metode ARIMA, Regresi, Fungsi Transfer dan VAR dalam melakukan pemodelan dan peramalan produksi kopi di Indonesia.

Tujuan dari disusunnya kegiatan ini adalah:

- n. Melakukan analisis dan peramalan data produksi kopi menggunakan model ARIMA, Regresi, Fungsi Transfer dan VAR di Indonesia
- o. Membandingkan metode tersebut dalam memperoleh ramalan data produksi komoditas kopi.
- p. Menentukan metode terbaik dalam meramal data produksi komoditas kopi di Indonesia.

## **BAHAN DAN METODE**

Data yang digunakan dalam melakukan pemodelan produksi kopi adalah data series produksi kopi tahun 1980-2019. Peubah yang diasumsikan mempengaruhi besaran produksi kopi untuk Fungsi Transfer adalah volume impor dan untuk pemodelan VAR adalah luas areal, volume ekspor, volume impor, harga dunia, dan harga urea.

Peramalan data produksi kopi, dilakukan uji coba beberapa metode yakni mengkaji metode ARIMA, Regresi, Fungsi Transfer dan VAR (Vector Autoregression), menggunakan software *Analisis R Studio*.

Metode yang digunakan dalam melakukan peramalan data produksi kopi di Indonesia adalah :

- o. ARIMA (*Autoregressive Integrated Moving Average*)

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk modelnya adalah



$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \theta_2 \dots \theta_p \phi_1 \phi_2 \dots \phi_q$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

#### p. Regresi

Regresi adalah suatu teknik analisis statistik yang digunakan untuk menyelidiki hubungan atau pengaruh antara dua atau lebih variabel.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Dimana :

$y$  = variabel respon/dependen/terikat

$x_1, x_2, \dots, x_k$  = variabel prediktor/independen/bebas

$\beta_0$  = konstanta/intercept

$\beta_1, \beta_2, \dots, \beta_k$  = koefisien regresi variabel bebas

$\varepsilon$  = error/residu/sisaan

#### q. Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut.

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t$$

Dimana:

- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$

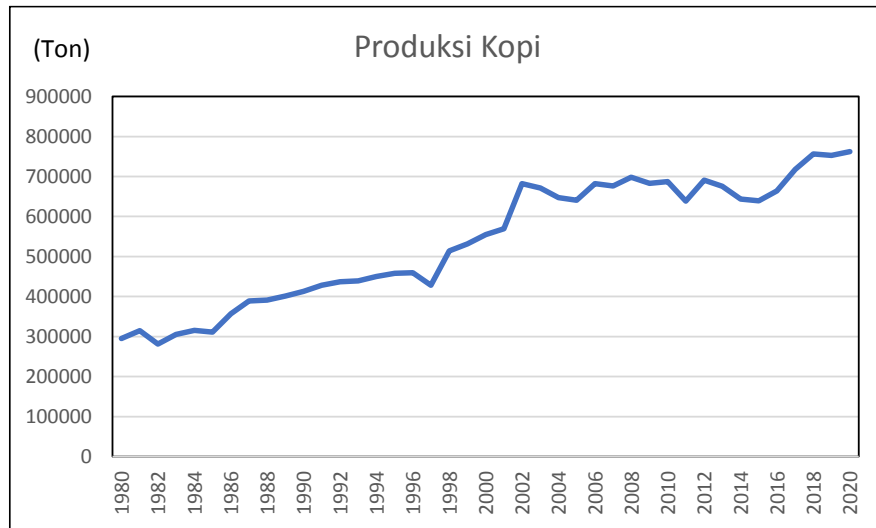
- $p \rightarrow$  ordo AR bagi noise  $N_t$

r. VAR (*Vector* )

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen., karena pada kenyataannya suatu variabel dapat bertindak sebagai endogen atau dependen dari suatu variabel yang lainnya (Gujarati, 2010).

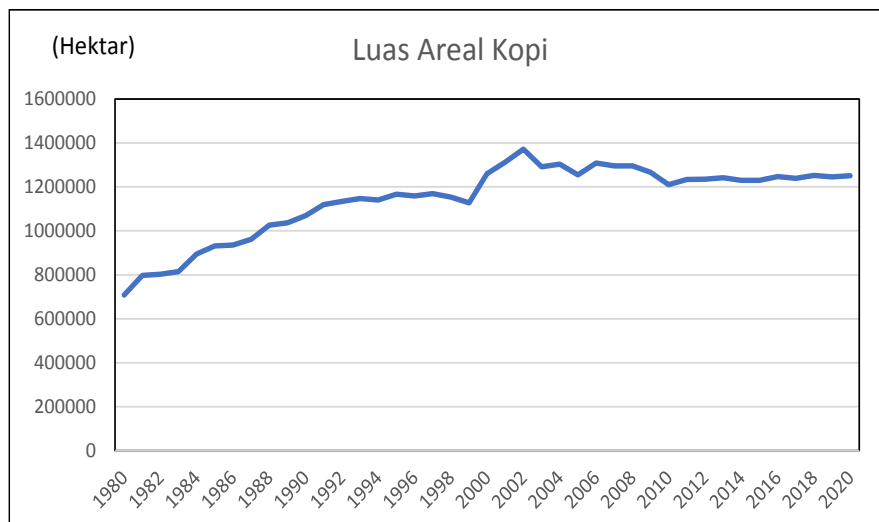
## HASIL DAN PEMBAHASAN

### i. Eksplorasi Data



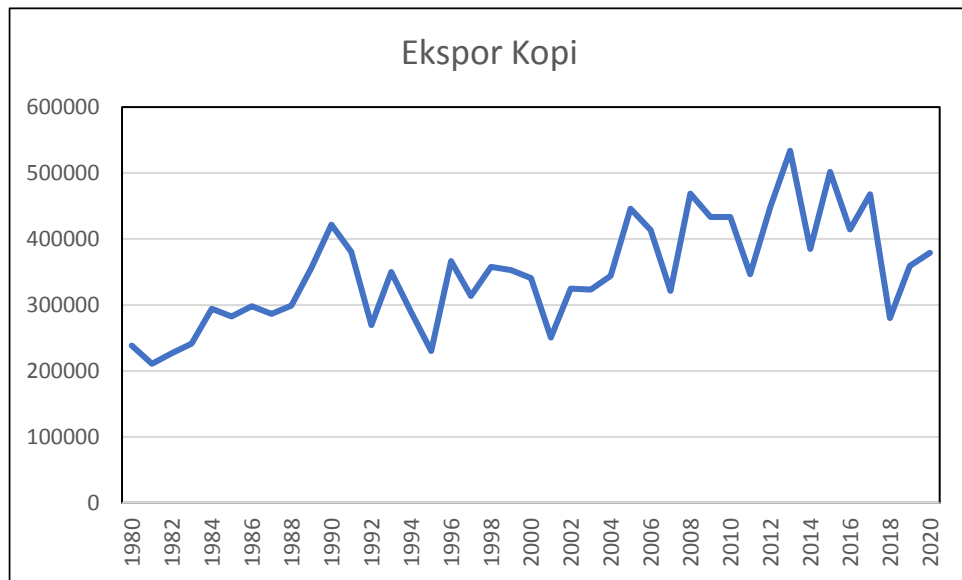
**Gambar 1.** Perkembangan Produksi Kopi di Indonesia, 1980-2020

Produksi kopi periode 1980-2020 terus meningkat dengan rata-rata pertumbuhan sebesar 2,58% (Gambar 1.)



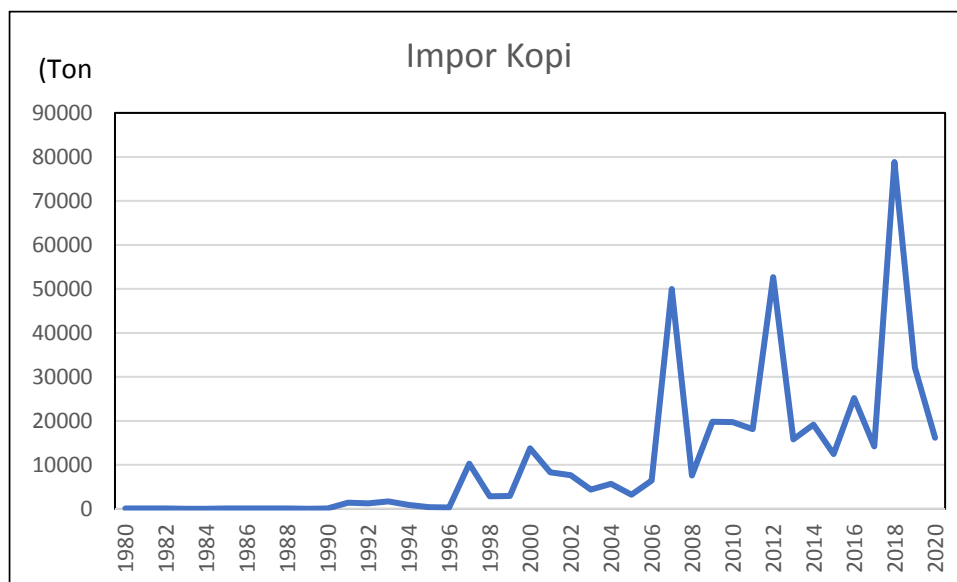
**Gambar 2.** Perkembangan Luas Areal Kopi di Indonesia, Tahun 1980-2020

Luas areal kopi periode 1980-2020 sedikit meningkat dengan rata-rata pertumbuhan sebesar 1,50% (Gambar 2.)



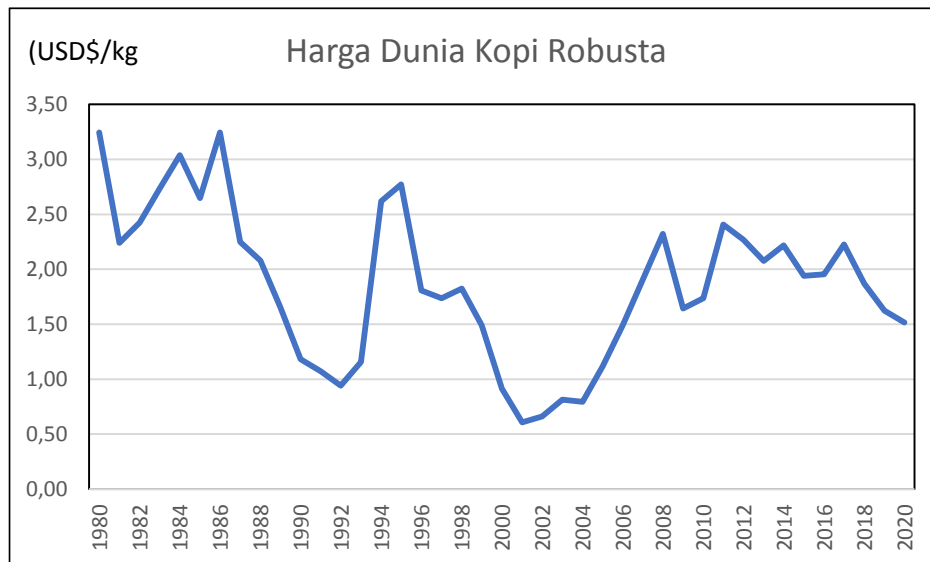
**Gambar 3.** Perkembangan Volume Ekspor Kopi di Indonesia, Tahun 1980-2020

Volume ekspor kopi di Indonesia periode 1980-2020 berfluktuasi, cenderung meningkat dengan rata-rata pertumbuhan sebesar 3,47% (Gambar 3.)



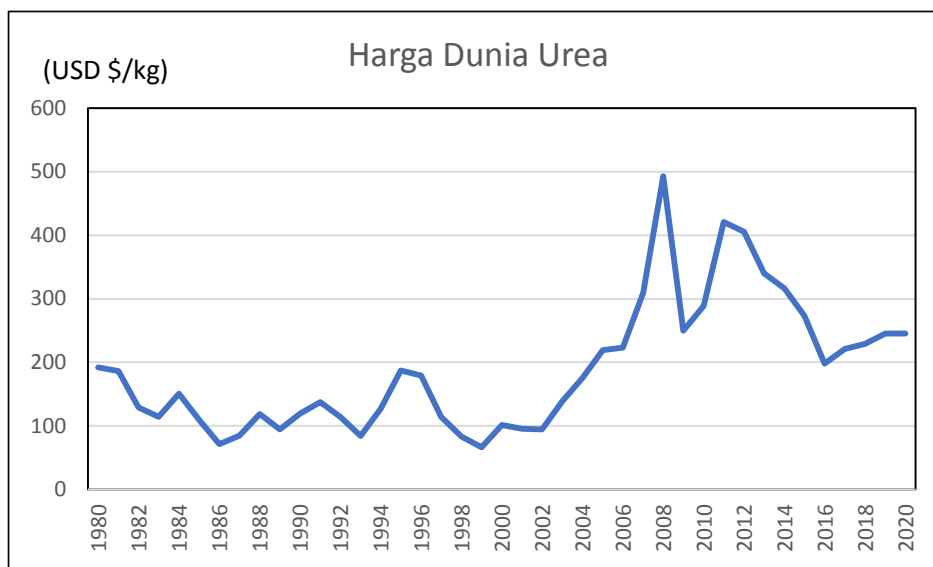
**Gambar 4.** Perkembangan Volume Impor Kopi di Indonesia, Tahun 1980-2020

Volume impor kopi di Indonesia periode 1980-2020 berfluktuasi, cenderung meningkat sangat tinggi dengan rata-rata pertumbuhan sebesar 156,46% (Gambar 4.)



**Gambar 5.** Perkembangan Harga Dunia Kopi Robusta di Indonesia, 1980-2020

Harga dunia kopi periode 1980-2020 berfluktuasi cenderung meningkat dengan rata-rata pertumbuhan sebesar 1,51% (Gambar 5.)



**Gambar 6.** Perkembangan Harga Urea di Indonesia, Tahun 1980-2020

Perkembangan harga urea periode 1980-2019 berfluktuasi cenderung meningkat dengan rata-rata pertumbuhan sebesar 4,54% (Gambar 6.)

## j. ARIMA

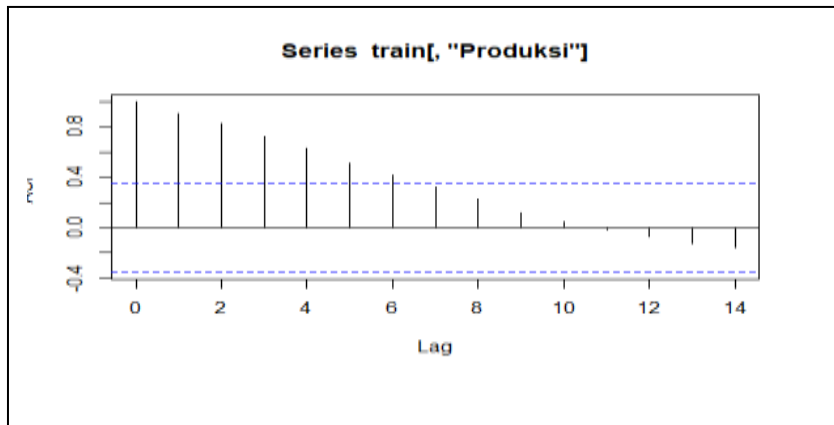
Dalam melakukan pemodelan produksi kopi menggunakan model Autoregressive Integrated Average (ARIMA), data yang digunakan adalah periode 1980-2020. Periode data tersebut kemudian dipisahkan menjadi data set training dan testing. Panjang series data pada data set training adalah 1980-2013, dataset testing adalah periode 2014-2020. Dataset training digunakan untuk melakukan penyusunan model, sementara dataset testing digunakan untuk validasi model.

Uji kestasioneran data seperti yang disyaratkan apabila melakukan pemodelan ARIMA dilakukan secara visual menggunakan hasil plot data maupun uji formal statistik. Gambar 1 menunjukkan produksi terus meningkat dengan tahun tertentu sedikit berfluktuasi dengan peningkatan yang berbeda setiap tahunnya, sehingga dapat disimpulkan bahwa data kurang stasioner. Hasil uji statistikpun menunjukkan bahwa data kurang stasioner karena uji Augmented Dickey-Fuller (ADF) yang mengindikasikan bahwa data produksi kopi tidak stasioner, terlihat dari nilai  $p$  yang lebih besar dari taraf nyatanya (0.05) (Tabel 1).

**Tabel 1.** Hasil Uji Augmented Dickey-Fuller Produksi Kopi

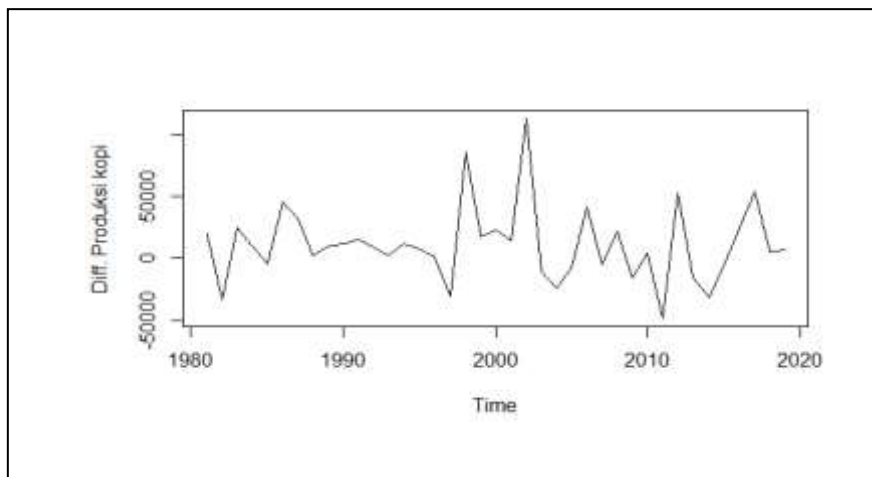
```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
  
Test regression trend  
  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)  
  
Residuals:  
  Min      1Q  Median      3Q      Max  
-53286 -11483  -2807  10653 105982  
  
Coefficients:  
            Estimate Std. Error t value Pr(>|t|)  
(Intercept)  8.161e+04  3.444e+04   2.369  0.0235 *  
z.lag.1      -2.487e-01  1.209e-01  -2.057  0.0472 *  
tt           3.041e+03  1.565e+03   1.943  0.0601 .  
z.diff.lag  -1.031e-02  1.695e-01  -0.061  0.9518  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
  
Residual standard error: 30550 on 35 degrees of freedom  
Multiple R-squared:  0.1247, Adjusted R-squared:  0.04964  
F-statistic: 1.662 on 3 and 35 DF, p-value: 0.1931  
  
value of test-statistic is: -2.0572 3.4947 2.1232  
  
Critical values for test statistics:  
      1pct  5pct 10pct  
tau3 -4.15 -3.50 -3.18  
phi2  7.02  5.13  4.31  
phi3  9.31  6.73  5.61
```

- Plot ACF untuk pemeriksaan autokorelasi



**Gambar 7.** Plot ACF Data Produksi Kopi

- Plot ACF mengindikasikan data produksi kopi tidak stasioner (memperkuat hasil Uji ADF sebelumnya)
- Untuk menjadi stasioner maka data dideferencing, dan setelah differencing satu kali data sudah stasioner seperti ditunjukkan pada Tabel 2.
- Data Produksi Kopi Setelah *Differencing*



**Gambar 8.** Plot Data Produksi Kopi seteah Deferencing

**Tabel 2.** Hasil Uji Augmented Dickey-Fuller Differencing Produksi Kopi

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-47311  -5917   7625  22470 110935

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1      -0.91213    0.23111  -3.947 0.000352 ***
z.diff.lag  -0.06829    0.16345  -0.418 0.678550
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 33790 on 36 degrees of freedom
Multiple R-squared:  0.4987, Adjusted R-squared:  0.4708
F-statistic: 17.9 on 2 and 36 DF, p-value: 4.003e-06

Value of test-statistic is: -3.9467

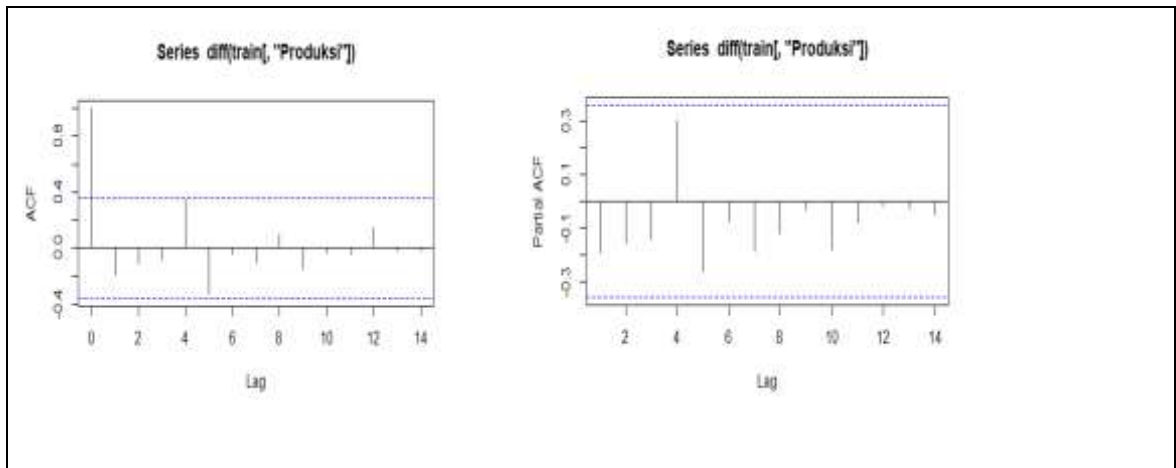
Critical values for test statistics:
      1pct  5pct 10pct
taul -2.62 -1.95 -1.61
```

- P-value hasil uji < taraf uji
- Statistik uji < critical value
- Disimpulkan: data STASIONER

Setelah mendapatkan data produksi kopi yang stasioner, maka langkah selanjutnya adalah menentukan ordo AR dan MA menggunakan fungsi auto arima yang terdapat pada RStudio. Dari hasil pendugaan auto arima pada Tabel 3, maka model ARIMA yang terpilih adalah ARIMA (0,1,0) yang berarti model ARIMA yang optimal untuk mengestimasi data yang diolah adalah ARIMA dengan ordo *Auto Regressive* (AR) = 0, ordo *differencing* = 1, dan ordo *Moving Average* (MA) = 0.



➤ **Pendugaan Model ARIMA**



**Gambar 9.** Plot ACF PACF hasil *differencing*

Plot ACF PACF hasil *differencing* tidak menunjukkan perilaku yang khas sesuai model ARIMA tertentu.

**Tabel 3.** Hasil Fungsi AUTOARIMA

```

Series: train[, "Produksi"]
ARIMA(0,1,0) with drift

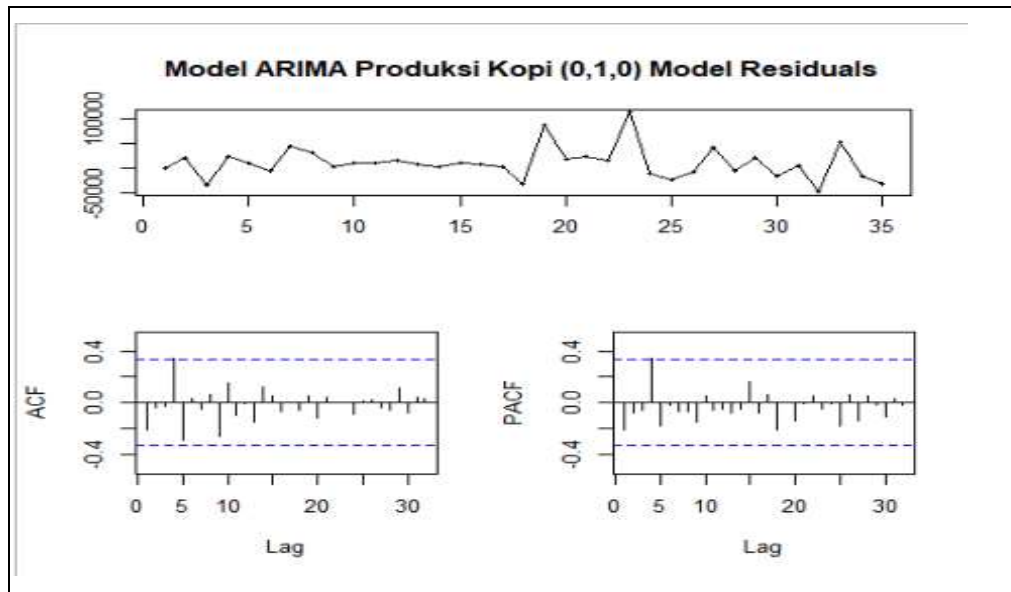
Coefficients:
      drift
11542.667
s.e.      5448.433

sigma^2 estimated as 1.01e+09: log likelihood=-388.42
AIC=780.84  AICC=781.24  BIC=783.83

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 8.336182 30835.15 20990.96 -0.1570285 4.124206 0.8991456 -0.2487313
    
```

Dengan fungsi autoarima diperoleh model ARIMA dengan AIC terkecil adalah ARIMA(0,1,0)

➤ **Pemeriksaan Plot Sisaan**



**Gambar 10.** Plot ACF PACF Sisaan ARIMA (0,1,0)

Plot ACF PACF sisaan tidak menunjukkan perilaku yang khas sesuai model ARIMA tertentu.

➤ **Peramalan data testing**

**Tabel 4.** Hasil Peramalan Data Testing

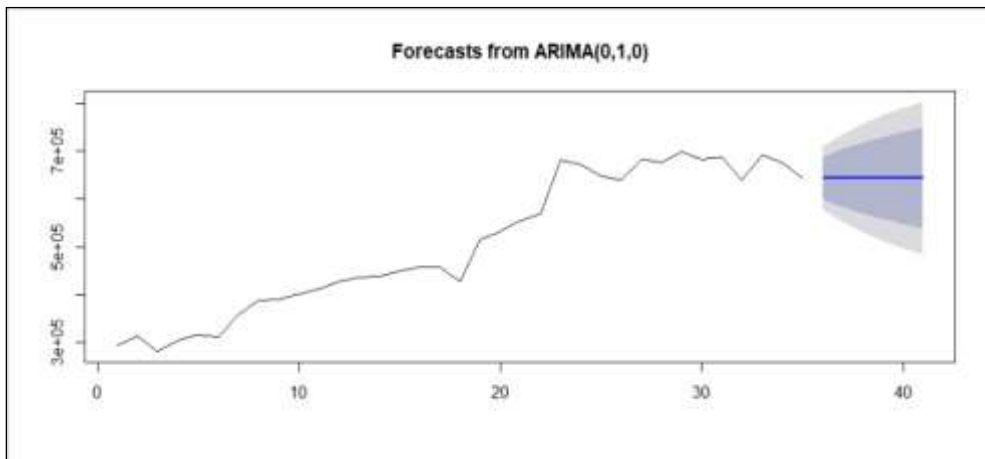
```

Time Series:
Start = 36
End = 41
Frequency = 1
[1] 643857 643857 643857 643857 643857 643857
> accuracy(ramalan_arima,test[, "Produksi"])

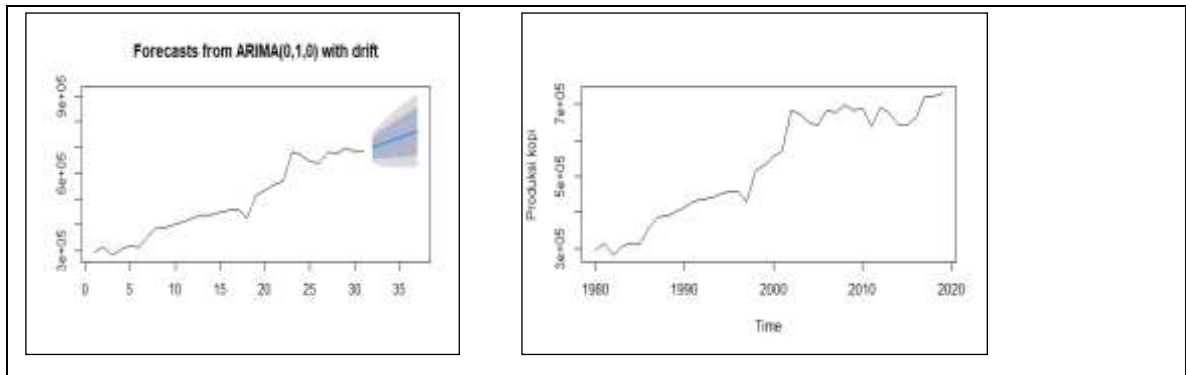
```

	ME	RMSE	MAE	MPE	MAPE	MASE
ACF1						
Training set	9976.542	32841.46	22934.83	2.023504	4.579476	0.9717857
Test set	71498.000	85977.97	72998.67	9.576159	9.810874	3.0930713
NA						

➤ **Plot Ramalan ARIMA**



**Gambar 12.** Ramalan model ARIMA (0,1,0)



**Gambar 11.** Perbandingan plot ramalan model ARIMA (0,1,0) dan plot data asli produksi kopi

➤ **Pengepasan Model untuk Seluruh Data**

**Tabel 5.** Hasil Pengepasan Model Seluruh Data ARIMA (0,1,0)

Series: kopi[, "Produksi"]							
ARIMA(0,1,0)							
sigma^2 estimated as 1.071e+09: log likelihood=-472.6							
AIC=947.21 AICC=947.31 BIC=948.89							
Training set error measures:							
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	11407.37	32331.17	22861.61	2.127007	4.366231	0.9759169	-0.1377965

➤ **Peramalan ARIMA (0,1,0) untuk seluruh Data**

**Tabel 6.** Hasil Ramalan Produksi Kopi ARIMA (0,1,0)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	762380	720431.2	804328.8	698224.9	826535.1
2022	762380	703055.4	821704.6	671650.9	853109.1
2023	762380	689722.6	835037.4	651260.0	873500.0
2024	762380	678482.4	846277.6	634069.7	890690.3
2025	762380	668579.6	856180.4	618924.7	905835.3

Karena hasil ramalan dari tahun 2020-2025 produksi kopi sama maka dicari ARIMA ordo lain, melalui fungsi armaselect (diff) dan diujicobakan ARIMA (1,1,0), ARIMA (2,1,0), ARIMA (4,1,0), ARIMA (5,1,0), dan ARIMA (4,1,3).

➤ **Peramalan ARIMA (1,1,0) untuk seluruh Data**

**Tabel 7.** Hasil Pengepasan Model Seluruh Data

```

Series: kopi[, "Produksi"]
ARIMA(1,1,0)

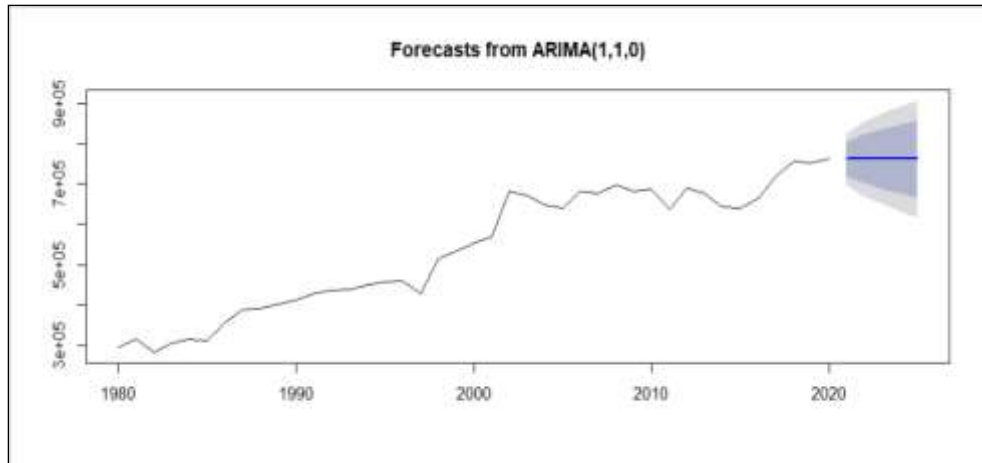
Coefficients:
      ar1
      0.0040
s.e.    0.1571

      sigma^2 estimated as 1.099e+09:  log likelihood=-472.6
      AIC=949.21   AICC=949.53   BIC=952.58

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 11362.89 32330.91 22859.94 2.118597 4.365097 0.9758455 -0.1413685
    
```

**Tabel 8.** Hasil Ramalan Produksi Kopi ARIMA (1,1,0)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	762419.3	719936.5	804902.2	697447.4	827391.2
2022	762419.5	702219.8	822619.1	670352.1	854486.9
2023	762419.5	688641.0	836197.9	649585.1	875253.9
2024	762419.5	677199.0	847639.9	632086.1	892752.9
2025	762419.5	667121.1	857717.9	616673.2	908165.8



**Gambar 12.** Ramalan model ARIMA (1,1,0)

➤ **Peramalan ARIMA (2,1,0) untuk seluruh Data**

**Tabel 9.** Hasil Pengepasan Model Seluruh Data ARIMA (2,1,0)

```

Series: kopi[, "Produksi"]
ARIMA(2,1,0)

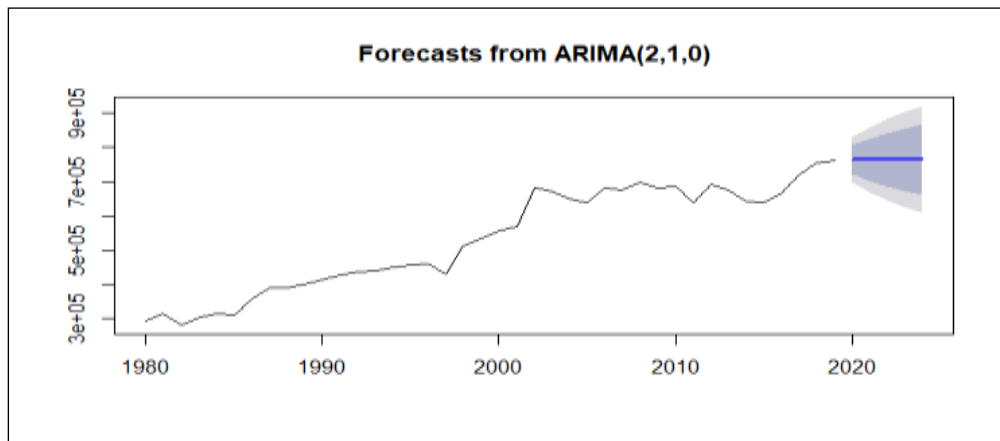
Coefficients:
      ar1      ar2
      0.0026  0.0674
s.e.  0.1564  0.1568

sigma^2 estimated as 1.122e+09:  log likelihood=-472.51
AIC=951.02  AICc=951.69  BIC=956.09

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
ACF1
Training set 10621.23 32252.95 22673.01 1.981222 4.328464 0.9678659 -0.122
4003
  
```

**Tabel 10.** Hasil Ramalan Produksi Kopi ARIMA (2,1,0)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	762167.3	719232.9	805101.7	696504.8	27829.8
2022	762831.6	702033.8	823629.5	669849.3	55813.9
2023	762819.0	686616.6	839021.5	646277.5	79360.6
2024	762863.8	673877.0	851850.5	626770.3	98957.2
2025	762863.0	662621.3	863104.8	609556.5	16169.6



**Gambar 13.** Ramalan model ARIMA (2,1,0)

➤ **Peramalan ARIMA (4,1,0) untuk seluruh Data**

**Tabel 11.** Hasil Pengepasan Model Seluruh Data ARIMA (4,1,0)

```

Series: kopi[, "Produksi"]
ARIMA(4,1,0)
Coefficients:
      ar1      ar2      ar3      ar4
      -0.0001  0.0409  0.0219  0.3871
s.e.      0.1428  0.1402  0.1449  0.1470

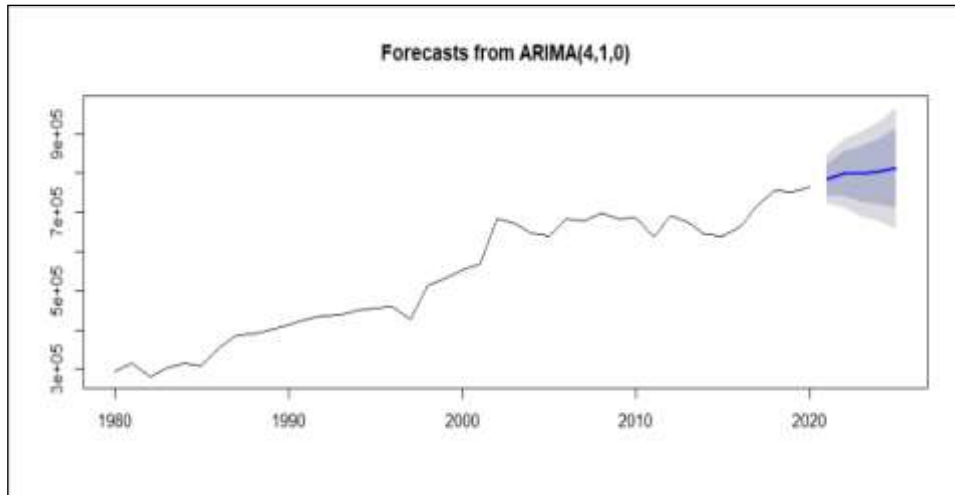
sigma^2 estimated as 999344958:  log likelihood=-469.43
AIC=948.87  AICC=950.63  BIC=957.31

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 7204.956 29622.18 21043.78 1.352242 4.060417 0.8983173 -0.02199532

```

**Tabel 12.** Hasil Ramalan Produksi Kopi ARIMA (4,1,0)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	784007.6	743494.7	824520.6	722048.4	845966.8
2022	799076.0	741784.7	856367.3	711456.5	886695.5
2023	798804.5	727669.4	869939.6	690012.8	907596.3
2024	803714.6	720565.9	886863.3	676549.6	930879.6
2025	812405.3	710564.9	914245.7	656653.8	968156.8



**Gambar 14.** Ramalan model ARIMA (4,1,0)

➤ **Peramalan ARIMA (3,1,0) untuk seluruh Data**

**Tabel 13.** Hasil Pengepasan Model Seluruh Data ARIMA (3,1,0)

```

Series: kopi[, "Produksi"]
ARIMA(3,1,0)

Coefficients:
      ar1      ar2      ar3
 0.0017  0.0667  0.0169
s.e.  0.1566  0.1569  0.1589

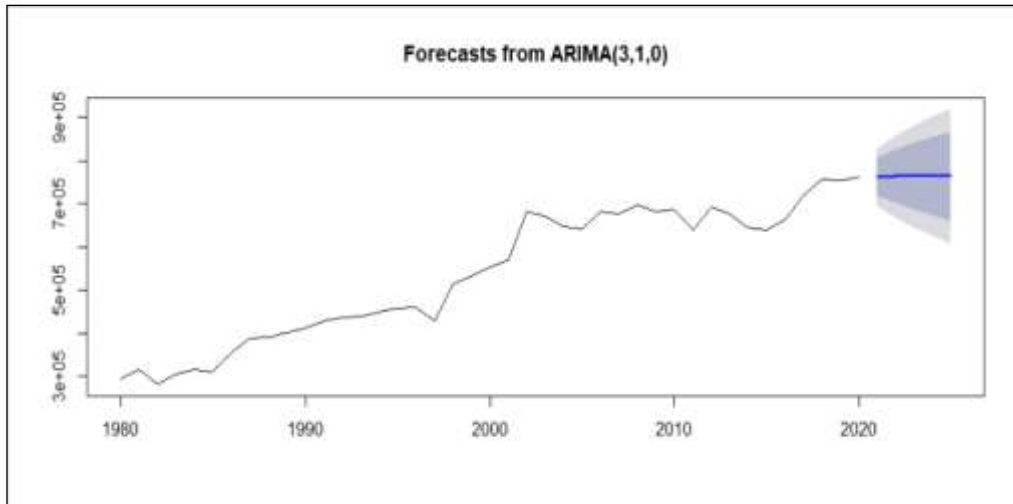
sigma^2 estimated as 1.152e+09:  log likelihood=-472.5
AIC=953.01  AICC=954.15  BIC=959.77

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 10464.76 32248.11 22661.53 1.952212 4.316781 0.9673757 -0.1245128

```

**Tabel 14.** Hasil Ramalan Produksi Kopi ARIMA (3,1,0)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	762805.0	719300.7	806309.2	696271.0	829339.0
2022	763403.8	701826.0	824981.7	669228.6	857579.1
2023	763600.0	686449.1	840751.0	645607.8	881592.3
2024	763647.5	673189.9	854105.0	625304.6	901990.4
2025	763670.8	661536.5	865805.0	607469.9	919871.6



**Gambar 15.** Ramalan model ARIMA (3,1,0)

➤ **Peramalan ARIMA (5,1,0) untuk seluruh Data**

**Tabel 15.** Hasil Pengepasan Model Seluruh Data ARIMA (5,1,0)

```

Series: kopi[, "Produksi"]
ARIMA(5,1,0)

Coefficients:
      ar1      ar2      ar3      ar4      ar5
    0.0457  0.0408  0.0251  0.3825 -0.1131
s.e.  0.1560  0.1392  0.1438  0.1463  0.1609

sigma^2 estimated as 1.014e+09:  log likelihood=-469.19
AIC=950.38  AICC=952.92  BIC=960.51

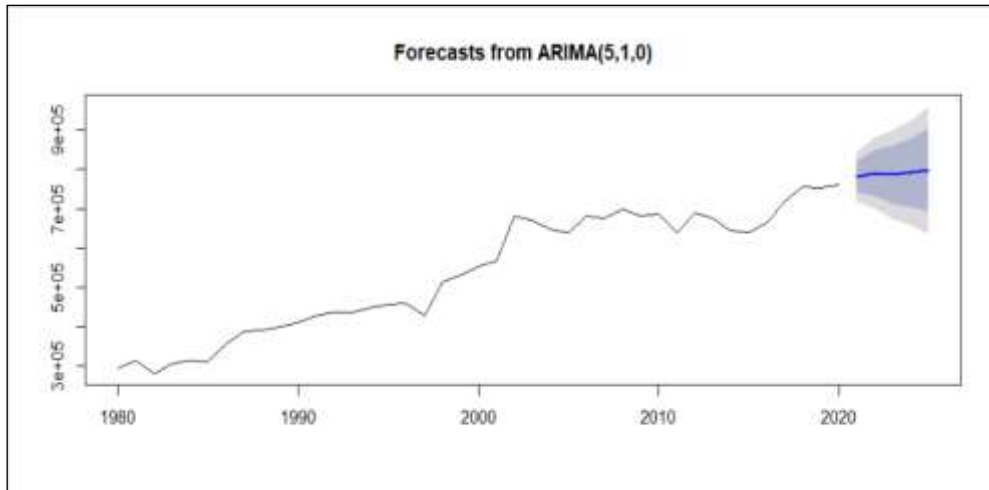
Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 7674.055 29424.44 21083.52 1.443707 4.065969 0.9000139 -0.06392642

```

**Tabel 16.** Hasil Ramalan Produksi Kopi ARIMA 5,1,0)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	781558.2	740744.8	822371.5	719139.5	843976.8
2022	791198.0	732145.3	850250.8	700884.7	881511.4
2023	787004.2	713105.5	860903.0	673985.8	900022.6
2024	791861.5	705022.8	878700.2	659053.2	924669.8
2025	798373.5	692026.7	904720.4	635730.0	961017.0





**Gambar 16.** Ramalan model ARIMA (5,1,0)

➤ **Peramalan ARIMA (4,1,3) untuk seluruh Data**

**Tabel 17.** Hasil Pengepasan Model Seluruh Data ARIMA (4,1,3)

```

Series: kopi[, "Produksi"]
ARIMA(4,1,3)

Coefficients:
      ar1      ar2      ar3      ar4      ma1      ma2      ma3
s.e.  0.3933  0.4639  0.3719  0.1616  0.4265  0.5295  0.4158

sigma^2 estimated as 1.06e+09:  log likelihood=-468.93
AIC=953.87  AICc=958.51  BIC=967.38

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 7080.402 29212.38 20132.06 1.332242 3.903393 0.859398 -0.06995996

```

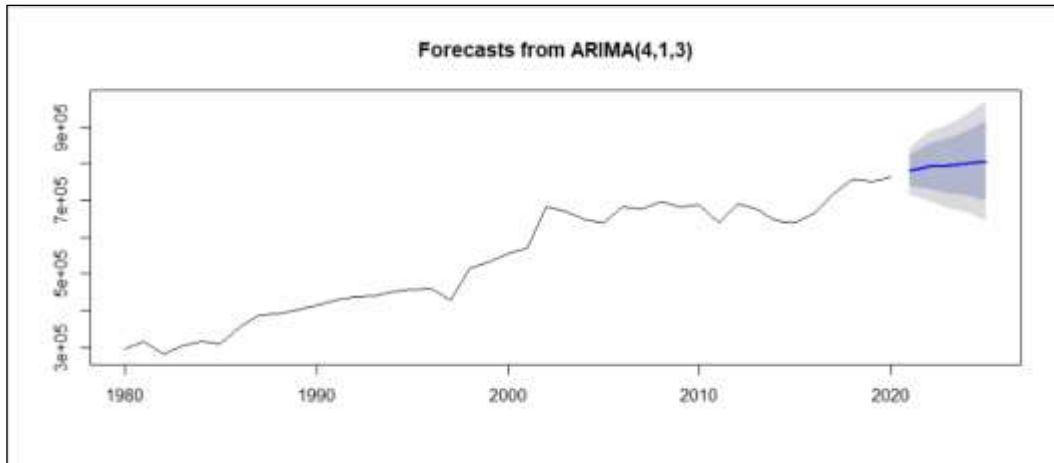
```

> ramalan_arima2 = forecast(model.arima, 5)
> ramalan_arima2

```

**Tabel 18.** Hasil Ramalan Produksi Kopi ARIMA (4,1,3)

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	781052.2	739323.2	822781.2	717233.2	844871.2
2022	793192.6	732183.3	854201.9	699887.0	886498.2
2023	793786.6	718683.7	868889.4	678926.6	908646.5
2024	801371.1	713381.0	889361.1	666801.9	935940.2
2025	807618.5	700443.7	914793.3	643708.7	971528.2



**Gambar 17.** Ramalan model ARIMA (4,1,0)

**e. Fungsi Transfer**

Peubah yang diasumsikan dapat mempengaruhi gejala produksi kopi adalah volume impor kopi, dan peubah ini akan digunakan untuk melakukan analisis menggunakan model Fungsi Transfer. Dari pola data volume impor dan uji Augmented Dickey-Fuller (Tabel 19) terlihat belum stasioner oleh karenanya sebelum melakukan pemodelan, series data volume impor ini dilakukan differencing terlebih dahulu (Tabel 20).

**Tabel 19.** Data Volume Impor Kopi yang belum Stasioner

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression trend

Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-12417  -5045  -1548    1941   32475

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -8.755e+03  4.351e+03  -2.012  0.053569 .
z.lag.1      -1.131e+00  2.768e-01  -4.085  0.000318 ***
tt           1.005e+03  3.024e+02   3.323  0.002419 **
z.diff.lag   -4.925e-03  1.841e-01  -0.027  0.978846
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9981 on 29 degrees of freedom
Multiple R-squared:  0.5709, Adjusted R-squared:  0.5265
F-statistic: 12.86 on 3 and 29 DF, p-value: 1.604e-05

Value of test-statistic is: -4.0849 5.6455 8.3491

Critical values for test statistics:
      1pct   5pct  10pct
tau3  -4.15  -3.50  -3.18
phi2   7.02   5.13   4.31
phi3   9.31   6.73   5.61
```

**Tabel 20.** Data Volume Impor Kopi yang sudah stasioner

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression none

Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)

Residuals:
    Min       1Q   Median       3Q      Max
-11457  -1187    -11     1071   45034

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
z.lag.1      -2.3636     0.3266  -7.238 4.67e-08 ***
z.diff.lag    0.5140     0.1919   2.679  0.0119 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 11100 on 30 degrees of freedom
Multiple R-squared:  0.8251, Adjusted R-squared:  0.8134
F-statistic: 70.76 on 2 and 30 DF, p-value: 4.385e-12

Value of test-statistic is: -7.238

Critical values for test statistics:
      1pct  5pct 10pct
tau1 -2.62 -1.95 -1.61
```

Pencarian ordo model ARIMA input untuk volume impor dilakukan dengan fungsi auto arima. Dari hasil fungsi auto arima, diperoleh ordo ARIMA terbaik untuk volume impor adalah ARIMA(0,1,2).

**Tabel 21.** Hasil Fungsi Auto ARIMA Volume Impor Kopi

```
Series: train.h[, "Impor"]
ARIMA(0,1,2) with drift

Coefficients:
      ma1      ma2      drift
   -1.1265  0.3277  835.5537
s.e.   0.2240  0.2500  355.2434

sigma^2 estimated as 99583421: log likelihood=-360.5
AIC=729.01  AICc=730.39  BIC=735.11

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -164.1619 9391.616 5298.176 -1835.858 1853.92 0.8091253 0.0232345
```

Model ARIMA (0,1,2) menghasilkan mape sangat tinggi (MAPE=1853,92) oleh sebab itu dicari model tentarif model input dengan fungsi armaselect.(Tabel.22).

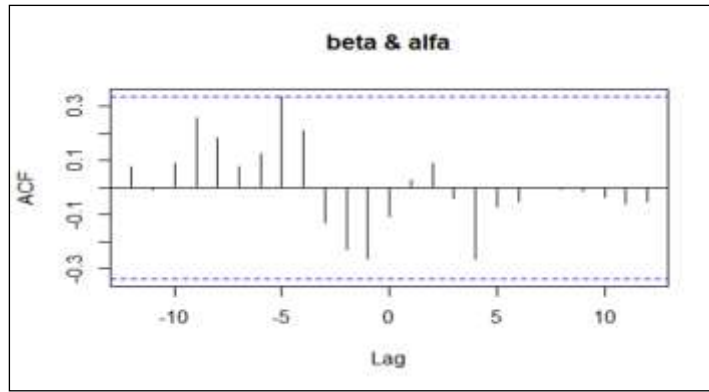
**Tabel 22.** Hasil Fungsi ARIMA Select Volume Impor Kopi

	p	q	sbc
[1,]	5	5	629.5004
[2,]	4	0	631.4366
[3,]	4	5	633.7132
[4,]	5	0	635.8928
[5,]	2	0	638.2362
[6,]	1	0	641.8160
[7,]	3	0	642.3428
[8,]	4	1	649.3488
[9,]	0	0	650.5462
[10,]	5	1	653.3187

Dari hasil pencarian ordo arima optimum volume impor kopi menggunakan fungsi auto arima, diputuskan model arima input volume impor kopi adalah ARIMA (2,1,0). Tahap selanjutnya adalah melakukan *prewhitening* dan analisis korelasi silang antara residual model arima volume impor kopi dan produksi kopi menggunakan ARIMA (2,1,0). Dari hasil plot ccf (*cross correlation function*) antara alfa (residual impor kopi) dan beta (residual produksi kopi), diperoleh hasil bahwa tidak terdapat *lag* dari pengaruh yang diberikan oleh impor kopi terhadap produksi kopi atau  $lag = 0$ . Hal tersebut dapat disimpulkan dari tidak adanya *bar* atau garis yang melewati ambang batas pada nilai *lag* positif (Gambar 15).

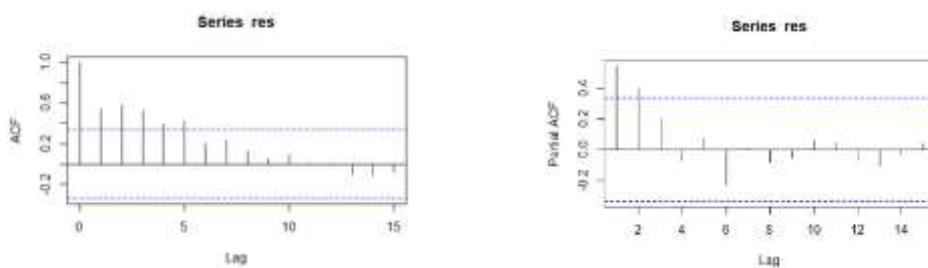
*Prewhitening* → alfa dan beta. Alfa adalah nilai residual dari model arima variabel input (vol impor kopi), sedangkan Beta adalah nilai residual dari model arima output (produksi kopi)

- pengaruh dari input volume impor kepada produksi kopi Indonesia terjadi saat yang sama, →  $b$  dan  $s = 0$ ,  $r$  atau jeda pengaruh variabel output = 0
- fungsi transfer alfa beta → residual → pendugaan ARIMA residual



**Gambar 15.** Plot ACF Alfa (Residual Impor Kopi) dan Beta (Residual Produksi Kopi)

Karena tidak adanya *bar* atau garis yang melewati ambang batas pada nilai *lag* positif, maka model ARIMA untuk pengepasan model fungsi transfer adalah ARIMA(0,0,0). Kemudian residual model ARIMA(0,0,0) tersebut diteliti stasioneritasnya. Dari plot ACF dan PCF residual, diperoleh hasil bahwa residual model cukup stasioner (Gambar 13) .



**Gambar 16.** Plot ACF dan PACF Residual Model ARIMA(0,0,0)

Langkah selanjutnya adalah pencarian model ARIMA terbaik untuk residual tersebut atau dapat disebut dengan *noise*. Hasil dari fungsi auto arima menunjukkan bahwa model ARIMA(0,1,1) merupakan model terbaik (Tabel 23). Namun karena ordo ARIMA(0,1,1) dinilai kurang baik untuk peramalan karena akan menghasilkan peramalan yang konstan, maka diperlukan referensi ordo ARIMA lain menggunakan fungsi *arma select* sebagaimana tampak pada Tabel 24.

**Tabel 23.** Hasil Fungsi Auto ARIMA Volume Impor Kopi

```

ARIMA(0,1,1)
Coefficients:
      ma1
      -0.6354
s.e.    0.1141

sigma^2 estimated as 7.196e+09:  log likelihood=-433.84
AIC=871.68  AICc=872.07  BIC=874.73

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 15845.81 82370.02 55037.49 35.7852 109.9809 0.9048914 -0.1695355
    
```

**Tabel 24.** Hasil Pencarian Ordo ARIMA Optimum *Noise* menggunakan Fungsi Armaselect (diff)

	p	q	sbic
[1,]	2	0	778.7308
[2,]	1	0	778.8688
[3,]	3	0	783.2589
[4,]	4	0	785.7555
[5,]	0	0	786.7743
[6,]	5	0	789.8602
[7,]	0	1	795.8918
[8,]	1	1	798.3766
[9,]	0	2	800.0231
[10,]	2	1	800.9551

Setelah dilakukan uji signifikansi model ARIMA dan ARIMA-xreg pada pilihan ordo arima di atas, diperoleh ordo ARIMA *noise* yang signifikan adalah ARIMA(1,1,1) sebagaimana tampak pada Tabel 25. berikut.

**Tabel 25.** Hasil Uji Signifikansi ARIMA-Xreg (1,1,1)

```

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1 -0.81900    0.15736  -5.2046 1.944e-07 ***
ma1 -0.47426    0.16976  -2.7937 0.005212 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Oleh karena telah didapatkan model input serta *noise* yang terbaik untuk peramalan, maka langkah selanjutnya adalah melakukan peramalan. Peramalan dilakukan menggunakan dua jenis data input, yakni nilai aktual dan data ramalan. Dari hasil peramalan menggunakan nilai aktual

sebagai input, diperoleh nilai MAPE sebesar 2,99% (Tabel 26 sementara peramalan menggunakan data ramalan sebagai input memiliki MAPE sebesar 3,08% (Tabel 27).

**Tabel 26.** Pengukuran Tingkat Kesalahan Model Fungsi Transfer dengan Nilai Aktual sebagai Input

```

Series: train.h."Produksi"
ARIMA(1,1,1)
Coefficients:
      ma1
      -0.6354
s.e.    0.1141

sigma^2 estimated as 7.196e+09:  log likelihood=-433.84
AIC=871.68  AICc=872.07  BIC=874.73

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 15845.81 82370.02 55037.49 35.7852 109.9809 0.9048914 -0.1695355
    
```

**Tabel 27.** Pengukuran Tingkat Kesalahan Model Fungsi Transfer dengan Data Ramalan sebagai Input

```

Series: train.h[, "Produksi"]
ARIMA(1,1,1)
Coefficients:
      ar1      ma2
      -0.0766  0.0682
s.e.    0.1714  0.1723

sigma^2 estimated as 1.166e+09:  log likelihood=-402.13
AIC=810.25  AICc=811.05  BIC=814.83

Training set error measures:
              ME      RMSE      MAE      MPE      MAPE      MASE
ACF1
Training set 10039.98 32648.38 22537.36 2.041018 4.524264 0.9549445 -0.125
3401
    
```

Setelah diperoleh model fungsi transfer yang optimal, maka model tersebut digunakan untuk melakukan peramalan produksi kopi 5 tahun ke depan (2020-2024). Hasil peramalan produksi kopi Indonesia menggunakan fungsi transfer dengan volume impor kopi sebagai peubah ditampilkan pada Tabel 28 berikut.

**Tabel 28.** Hasil Peramalan Produksi Kopi menggunakan Fungsi Transfer

No.	Tahun	Produksi Kopi (Ton)
1.	2021	760245,3
2.	2022	756753,4
3.	2023	756735,1
4.	2024	758216,1
5.	2025	757098,1

**f. Model Vector Auto Regression (VAR) untuk Peramalan Produksi Kopi**

Pemodelan dengan menggunakan model VAR dilakukan dengan memasukkan peubah produksi kopi, luas areal kopi, volume ekspor kopi, volume impor kopi, harga kopi dunia, dan harga pupuk urea dunia. Hal ini diasumsikan bahwa apabila terjadi gejolak pada variabel-variabel tersebut, akan mempengaruhi gejolak produksi kopi ataupun sebaliknya.

Pada tahap pertama, perlu dilakukan pemeriksaan apakah terdapat pengaruh trend dalam model. Hasil pengujian signifikansi trend dilakukan pada  $lag=1$  dan  $lag=2$ . Hasil pengujian trend pada  $lag=1$  menunjukkan bahwa *trend* berpengaruh signifikan pada persamaan produksi kopi (signifikan 5%) dan persamaan harga pupuk urea dunia (signifikan 10%). Sementara hasil pengujian *trend* pada  $lag=2$  menunjukkan bahwa *trend* berpengaruh signifikan hanya pada persamaan harga pupuk urea dunia (signifikan 5%).

Dari hasil *running* model menggunakan  $lag$  1 dan 2, diperoleh informasi komponen *trend* tidak terlalu berpengaruh signifikan sehingga *trend* dikeluarkan dalam model untuk *running* model VAR (p) selanjutnya.

Langkah selanjutnya adalah menentukan  $lag$  (p) yang akan digunakan dalam model VAR. Penentuan  $lag$  (p) dilakukan dengan melakukan pengujian VAR(1) hingga VAR(4) dengan pertimbangan kecukupan series data. Dari hasil *running* menggunakan  $p=1$  s.d  $p=4$ , diperoleh informasi bahwa variabel yang signifikan cukup banyak pada  $p=4$ , sehingga model VAR yang digunakan selanjutnya adalah VAR (4). Hal tersebut diperkuat juga dengan nilai *Adjusted R-squared*: 0,9683 atau dapat dikatakan bahwa 96,83% keragaman produksi kopi dapat dijelaskan



oleh kelima variabel/peubah penjelas termasuk konstanta. Hasil pengujian model VAR(4) dapat dilihat dalam Tabel 29.

Model VAR(5) dan VAR(6) tidak dirun karena keterbatasan series. Jika dirun maka muncul error.

Dugaan model VAR(4) untuk produksi kopi adalah sebagai berikut :

**Tabel 29.** Hasil Estimasi VAR Produksi Kopi dengan Peubah Luas Areal, Volume Ekspor, Volume Impor, Harga Kopi Dunia, dan Harga Pupuk Urea Dunia pada Lag=4, type “const”

```

VAR Estimation Results:
=====
Endogenous variables: Produksi, Luas, Ekspor, Impor, World.Price
Deterministic variables: const
Sample size: 37
Log Likelihood: -1662.075
Roots of the characteristic polynomial:
0.9428 0.9428 0.9229 0.9229 0.9087 0.9047 0.9047 0.8611 0.8611 0.8599 0.8599 0.8499
0.8499 0.7935 0.7935 0.6633 0.6633 0.3203 0.3203 0.2252
Call:
VAR(y = kopi[1:41, c(2, 3, 4, 5, 6)], p = 4, type = "const")

Estimation results for equation Produksi:
=====
Produksi = Produksi.l1 + Luas.l1 + Ekspor.l1 + Impor.l1 + world.Price.l1 + Produksi
.l2 + Luas.l2 + Ekspor.l2 + Impor.l2 + World.Price.l2 + Produksi.l3 + Luas.l3 + Eks
por.l3 + Impor.l3 + World.Price.l3 + Produksi.l4 + Luas.l4 + Ekspor.l4 + Impor.l4 +
World.Price.l4 + const

      Estimate Std. Error t value Pr(>|t|)
Produksi.l1  7.393e-01  2.362e-01  3.130  0.00646 **
Luas.l1      2.255e-01  1.825e-01  1.236  0.23443
Ekspor.l1   -1.595e-01  1.273e-01  -1.254  0.22799
Impor.l1    -5.001e-01  5.721e-01  -0.874  0.39489
World.Price.l1 6.059e+03  1.529e+04  0.396  0.69723
Produksi.l2  1.418e-01  2.972e-01  0.477  0.63978
Luas.l2     8.336e-02  1.982e-01  0.421  0.67970
Ekspor.l2   -1.108e-02  1.261e-01  -0.088  0.93110
Impor.l2    -4.911e-01  5.755e-01  -0.853  0.40609
World.Price.l2 -1.180e+04  1.825e+04  -0.647  0.52699
Produksi.l3  4.240e-02  2.679e-01  0.158  0.87624
Luas.l3     -3.660e-01  2.059e-01  -1.777  0.09457
Ekspor.l3   -4.283e-03  1.171e-01  -0.037  0.97128
Impor.l3    -8.431e-01  7.274e-01  -1.159  0.26346
World.Price.l3 9.072e+03  1.824e+04  0.497  0.62566
Produksi.l4  2.619e-01  2.392e-01  1.095  0.28980
Luas.l4     1.435e-01  1.739e-01  0.825  0.42145
Ekspor.l4   7.375e-02  1.054e-01  0.699  0.49436
Impor.l4    -8.907e-01  7.009e-01  -1.271  0.22199
World.Price.l4 2.934e+04  1.482e+04  1.980  0.06516
const      -1.789e+05  1.432e+05  -1.249  0.22959
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 28820 on 16 degrees of freedom
Multiple R-Squared: 0.9806, Adjusted R-squared: 0.9564
F-statistic: 40.53 on 20 and 16 DF, p-value: 4.099e-10

```

**Tabel 30.** Covariance Matrix Residuals

	Produksi	Luas	Ekspor	Impor	world.Price
Produksi	8.305e+08	277799515	-306888692	56147305	46.9015
Luas	2.778e+08	1385909289	33460311	919635	-7326.5793
Ekspor	-3.069e+08	33460311	3453347420	-575609000	-9167.3337
Impor	5.615e+07	919635	-575609000	178091757	1196.4646
World.Price	4.690e+01	-7327	-9167	1196	0.2256

Correlation matrix of residuals:					
	Produksi	Luas	Ekspor	Impor	world.Price
Produksi	1.000000	0.258941	-0.18122	0.145997	0.003426
Luas	0.258941	1.000000	0.01529	0.001851	-0.414338
Ekspor	-0.181217	0.015295	1.000000	-0.733982	-0.328430
Impor	0.145997	0.001851	-0.73398	1.000000	0.188755
World.Price	0.003426	-0.414338	-0.32843	0.188755	1.000000

Setelah terpilih *lag* model VAR yang optimal, maka perlu dilakukan pengujian asumsi non-autokorelasi, normalitas, dan homoskedas pada sisaan model VAR (4). Dari hasil pengujian non autokorelasi, diperoleh informasi bahwa model tersebut terbebas dari autokorelasi dengan tingkat kesalahan 5% (Tabel 30). Residual memenuhi asumsi kenormalan dengan tingkat kesalahan 5% melalui *normality test* (*Jarque Bera Test*, pengujian skewness, dan kurtosis) sebagaimana tampak pada Tabel 32, sedangkan asumsi homoskedastisitas atau kehomogenan ragam signifikan terpenuhi dengan *p-value* = 1 (Tabel 33).

**Tabel 31.** Hasil Uji Asumsi Non-Autokorelasi model VAR (2) menggunakan *Serial Test*

Portmanteau Test (asymptotic)
data: Residuals of VAR object varhsheet4
Chi-squared = 295.83, df = 300, p-value = 0.5572

**Tabel 32.** Hasil Uji Asumsi Normalitas Regresi Linier Produksi Kopi menggunakan *Normality Test*

JB-Test (multivariate)
data: Residuals of VAR object varhsheet4
Chi-squared = 3.2338, df = 10, p-value = 0.9754
\$Skewness
Skewness only (multivariate)
data: Residuals of VAR object varhsheet4
Chi-squared = 1.3946, df = 5, p-value = 0.9249
\$kurtosis
kurtosis only (multivariate)
data: Residuals of VAR object varhsheet4
Chi-squared = 1.8392, df = 5, p-value = 0.8709

**Tabel 33.** Hasil Uji Asumsi Homoskedas Regresi Linier Produksi Kopi menggunakan ARCH Test

ARCH (multivariate) data: Residuals of VAR object varhsheet4 Chi-squared = 480, df = 1125, p-value = 1
--

Sebelum melakukan peramalan, perlu dilakukan evaluasi nilai kesalahan model VAR yang terbentuk dengan menghitung nilai MAPE atau nilai rata-rata persentase kesalahan peramalan. Nilai MAPE VAR(4) untuk dataset *training* dan *testing* masing-masing sebesar 4,29% dan 1,41% seperti tersaji pada Tabel 34.

**Tabel 34.** Nilai MAPE Model VAR(4) Produksi Kopi Indonesia

Produksi	
Min. : 0,9256	Min. : 0.3058
1st Qu.: 4,0572	1st Qu.: 1,3791
Median : 6,1150	Median : 2.4655
<b>Mean : 6,7256</b>	<b>Mean : 3,0986</b>
3rd Qu.: 6,7792	3rd Qu.: 4,1367
Max. : 16,8260	Max. : 8,2469
MAPE data testing= <b>6,7256</b>	MAPE data training= <b>3,0986</b>

Setelah diperoleh model VAR yang optimal yakni VAR(4), maka model tersebut digunakan untuk melakukan peramalan produksi kopi 5 tahun ke depan (2021-2025). Hasil peramalan produksi kopi Indonesia menggunakan metode VAR dengan peubah produksi kopi, volume ekspor kopi, volume impor kopi, harga kopi dunia, dan harga pupuk urea dunia ditampilkan pada Tabel 35. berikut.

**Tabel 35.** Hasil Peramalan Produksi Kopi menggunakan VAR(4)

No.	Tahun	Produksi Kopi (Ton)
1.	2021	745940,0
2.	2022	685471,7
3.	2023	708978,5
4.	2024	760094,9
5.	2025	765635,4

## KESIMPULAN DAN SARAN

### Kesimpulan

- Hasil analisis dari empat model dalam meramalkan produksi kopi adalah sebagai berikut:
  - Model ARIMA terbaik adalah ARIMA (4,1,3) dengan MAPE data training 4,58% dan MAPE data testing 9,81%
  - Model Fungsi Transfer terbaik ARIMA (2,1,0) Xreg=volume impor. Model Fungsi Transfer tersebut memiliki MAPE data training 4,55% dan MAPE data testing 9,78%
  - Model VAR terbaik dengan memasukkan peubah luas areal kopi, volume ekspor kopi, volume impor kopi, harga kopi dunia, dan harga urea dunia adalah VAR(4) type: const. MAPE yang dihasilkan adalah MAPE data training 4,55% dan MAPE data testing 9,78%
- Dari keempat model yang dicobakan, model terbaik adalah model ARIMA (4,1,3) dengan melihat kerealistisan remalan dan MAPE relative kecil, yaitu MAPE data training sebesar 4,58% dan MAPE data testing sebesar 9,81%.
- Hasil ramalan produksi kopi model ARIMA (4,1,3) untuk 2021 sampai 2025 adalah 781.052 Ton, 793.193 Ton, 793.787 Ton, 801.371 Ton, 807.619 Ton.

## Saran

- Perlu dilakukan kajian mendalam dengan metode peramalan lainnya.
- Untuk model fungsi transfer dan VAR perlu diujicobakan dengan menggunakan variabel lain yang sekiranya lebih berpengaruh, baik secara teoritis maupun praktis.
- Perlu dikaji metode peramalan untuk data non parametrik, sehingga tidak terhalang oleh tidak terpenuhinya berbagai asumsi.

## DAFTAR PUSTAKA

- Direktorat Jenderal Perkebunan. 2019. Statistik Perkebunan Indonesia 2018-2020 Jakarta: Direktorat Jenderal Perkebunan, Kementerian Pertanian. Statistik Perkebunan Indonesia 2018-2020.
- Draper, N. R, dan Smith, H. 1992. Analisis Regresi Terapan, Jakarta : PT Gramedia Pustaka Utama.
- Gujarati, Damodar. N dan Porter, Dawn. C. 2009. Basic Econometrics. Boston: Douglas Reiner.
- Malian, Husni & Sudi Mardianto dan Mewa Ariani. 2004. Faktor-faktor yang Mempengaruhi Produksi, Konsumsi dan Harga Beras serta Inflasi Bahan Makanan. Jurnal Agro Ekonomi, Volume 22 No. 2: 119-146
- Siagian, Viktor dan Muchamad Yusron. 2015. Analisis Faktor-Faktor yang Mempengaruhi Luas Lahan Garapan Usahatani Padi Sawah di Provinsi Banten. Lampung: Politeknik Negeri Lampung
- Thamrin, Syahrini. 2014. Faktor-faktor yang Mempengaruhi Produksi Usahatani Kopi Arabika di Kabupaten Enrekang Sulawesi Selatan. Pangkajene Kepulauan: Politeknik Pertanian Negeri Pangkajene Kepulauan.



# ESTIMASI PRODUKSI GULA INDONESIA, PENDEKATAN METODE ARIMA, VAR DAN FUNGSI TRANSFER

Efi Respati - *Statisticians* - [efi@pertanian.go.id](mailto:efi@pertanian.go.id)  
*Center for Agricultural Data and Information System - Ministry of Agriculture*  
*Jln. Harsono RM 3 Gd D Lt 4, Ragunan-Jakarta Selatan, Indonesia*

## ABSTRAK

Angka estimasi dapat digunakan sebagai sistem peringatan dini terhadap kondisi yang akan datang serta sebagai acuan menyusun kebijakan pengembangan komoditas yang lebih tepat. Estimasi produksi gula dilakukan dengan menggunakan pendekatan 3 metode yakni ARIMA, *Vector Autoregression* (VAR), dan Fungsi Transfer dengan menggunakan R-Studio sebagai alat pengolah datanya. Model estimasi terpilih didasarkan atas keunggulan nilai statistik dan kerealistisan hasil estimasinya adalah model Fungsi Transfer dengan nilai MAPE untuk data series *training* sebesar 9,24% dan MAPE untuk data series *testing* sebesar 4,00. Model tersebut menghasilkan nilai estimasi produksi gula tahun 2021-2025 yang cukup realistis.

*Kata kunci: gula, produksi, arima, fungsi transfer, var, MAPE*

## ABSTRACT

*Estimation figure for certain commodity has important role to be an early warning system. That figure can be used to determine the right commodity development policy. Estimation of sugar production is carried out using 3 approach methods namely ARIMA, Vector Autoregression (VAR), and Transfer Function using R-Studio as a data processing tool. The estimation of the model is based on the superiority of the statistical value and the reality of the estimation result is the Transfer Function model with a MAPE for training data series of 9.24% and for testing data series of 4.00%. This model produces an estimated value of sugar production in 2021-2025 which is quite realistic.*

*Keywords: sugar, production, arima, transfer function, var, MAPE.*

## PENDAHULUAN

Gula merupakan salah satu dari sembilan bahan pangan pokok penduduk Indonesia, sehingga ketersediaannya harus selalu tercukupi. Produksi gula Indonesia saat ini masih belum mencukupi kebutuhannya sehingga dari tahun ke tahun masih harus ada tambahan pasokan dari impor. Produksi gula tahun 2020 sebesar 2,13 juta ton, sementara kebutuhan untuk konsumsi langsung sebesar 2,80 juta ton, sehingga masih mengalami deficit sebesar 670 ribu ton. Apabila diperhitungkan dengan kebutuhan untuk konsumsi dan industri mencapai 5,8 juta ton, sehingga deficit mencapai 3,67 juta ton. Industri gula Indonesia saat ini masih banyak menghadapi kendala, diantaranya: menurunnya lahan perkebunan tebu, kondisi pertanaman tebu rakyat yang sudah menurun produktivitasnya, kualitas tebu rakyat yang sangat beragam, kinerja pabrik gula yang mulai menurun, dll.

Kementerian Pertanian menyusun peta jalan pencapaian swasembada gula konsumsi, dengan target terwujudnya peningkatan produksi sebesar 676 ribu ton gula pada periode tahun 2020-2023. Guna mencapai tujuan tersebut, program peningkatan produksi gula terus diupayakan oleh Kementerian Pertanian, melalui intensifikasi, ekstensifikasi dan pola kemitraan. Implementasi kegiatan pencapaian swasembada gula bekerjasama dengan perusahaan gula baik BUMN maupun swasta. Disamping untuk meningkatkan kembali produksi gula nasional, upaya swasembada gula juga dalam rangka meningkatkan pendapatan petani tebu.

Berbagai program dan kebijakan pergulaan nasional perlu didukung oleh ketersediaan data yang lengkap, akurat dan terkini bahkan estimasi beberapa tahun ke depan. Ketersediaan data diperlukan guna menjadi dasar perumusan kebijakan serta evaluasi pembangunan komoditas tebu/gula yang tepat. Data statistik komoditas gula yang tersedia saat ini adalah data Angka Tetap t-1 tahun, Angka Sementara untuk tahun berjalan serta Angka Estimasi untuk t+1 tahun. Angka estimasi diperoleh dengan model univariate atau hanya mempertimbangkan perilaku series data tersebut. Menurut pendapat ahli, beberapa metode statistik univariate dan multivariate mempunyai keunggulan dalam melakukan pemodelan dan peramalan data statistik, diantaranya model *Autoregressive Integrated Moving Average (ARIMA)*, *Vector Auto Regression (VAR)* dan *Fungsi Transfer (FT)*. Makalah ini membahas metode estimasi produksi gula dengan menggunakan ketiga model tersebut. Metode yang paling handal digunakan untuk melakukan estimasi produksi gula 5 (lima) tahun kedepan.

## METODOLOGI

### 4. Data dan Alat Pengolah Data

Data sekunder yang digunakan dalam makalah ini terdiri dari 49 series yakni dari periode tahun 1972 – 2020 seperti tersaji pada Tabel 1. Data series dibagi menjadi series *data training* (1972 – 2014) yang digunakan untuk mencari model terbaik dan series *data testing* (2015-2020) untuk validasi model.

**Tabel 1. Series Data yang Digunakan**

No	Data	Series	Sumber
1	Produksi gula	1972 – 2020	Ditjen Perkebunan
2	Ekspor Molases	1972 – 2020	BPS
3	Impor gula	1972 – 2020	BPS
4	Harga produsen jagung	1972 – 2020	BPS
5	Harga gula dunia	1972 – 2020	World Bank



Alat pengolah data yang digunakan adalah software R dan RStudio. R adalah bahasa pemrograman dan perangkat lunak gratis yang dikembangkan oleh Ross Ihaka dan Robert Gentleman pada tahun 1993. R memiliki banyak fungsi dan *package* untuk statistik dan visualisasi data yang lengkap. Sementara, RStudio merupakan software yang digunakan untuk mempermudah menulis dan menggunakan bahasa R. RStudio adalah *integrated development environment* (IDE) untuk R. RStudio mencakup konsol, editor penyorotan sintaks yang mendukung eksekusi kode langsung, serta alat untuk merencanakan, riwayat, debugging, dan manajemen ruang kerja. RStudio tersedia dalam edisi *open source* dan komersial dan dapat dijalankan di desktop (Windows, Mac, dan Linux) atau di browser yang terhubung ke RStudio Server atau RStudio Server Pro (Debian / Ubuntu, Red Hat / CentOS, dan SUSE Linux). Ringkasnya, jika bahasa R adalah mesin, RStudio merupakan *interface*-nya. RStudio memudahkan kita untuk berkomunikasi dengan R sehingga kita bisa menjalankan fungsi-fungsi statistika dan data science.

## 5. Tinjauan Literatur

Metode statistik yang dikembangkan oleh para ahli untuk melakukan peramalan data sangat beragam, baik peubah tunggal maupun peubah ganda, diantaranya metode ARIMA, Fungsi Transfer dan VAR (*Vector Autoregression*).

### b. Autoregressive Integrated Moving Average (ARIMA)

*Autoregressive Integrated Moving Average* (ARIMA) atau biasa disebut juga dengan metode *time series Box Jenkins*, sangat sesuai digunakan untuk melakukan peramalan jangka pendek, sementara untuk peramalan jangka panjang kurang baik ketepatannya. Metode ARIMA merupakan metode yang hanya menggunakan peubah dependen dan mengabaikan peubah independen sewaktu melakukan peramalan.

Metode ARIMA dibagi kedalam tiga kelompok model, yaitu *Autoregressive model* (AR), *Moving Average model* (MA) dan model campuran yang memiliki karakteristik kedua model di atas yaitu *Autoregressive Integrated Moving Average* (ARIMA).

#### 1) Autoregressive Model (AR)

AR adalah suatu model yang menjelaskan pergerakan suatu peubah melalui peubah itu sendiri di masa lalu.

Model *autoregressive* orde ke-*p* dapat ditulis sebagai berikut:

ARIMA (*p*,0,0)

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t \dots\dots\dots(1)$$

dimana:

$Y_t$ = data *time series* sebagai variable dependen pada waktu ke-*t*

$Y_{t-p}$ = data *time series* pada kurun waktu ke (*t-P*)

$\mu$ = suatu konstanta

$\theta_1 \dots \theta_p$ = parameter autoregresive ke-*p*

$\varepsilon_t$ = nilai kesalahan pada waktu ke *t*

#### 2) Moving Average Model (MA)

MA adalah suatu model yang melihat pergerakan peubahnya melalui sisaannya di masa lalu.

Bentuk model MA dengan ordo *q* atau MA (*q*) atau model ARIMA (0,*d*,*g*) dapat ditulis sebagai berikut:

$$Y_t = \mu - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t \dots\dots\dots(2)$$

dimana:

$Y_t$ = data *time series* sebagai variable dependen pada waktu ke-*t*

$\phi_1 \dots \phi_q$ = parameter-parameter moving average

$\varepsilon_{t-q}$  = nilai kesalahan pada waktu ke (t-q)

### 3) Autoregressive Intergrated Moving Average (ARIMA)

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampaunya. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t \dots \dots \dots (3)$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \theta_2 \dots \theta_p \phi_1 \phi_2 \dots \phi_q$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

#### e. Model Fungsi Transfer

Model Fungsi Transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke-t, tetapi juga pada waktu t+1, t+2, ..., t+k. Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara peubah input dan peubah output.

Tujuan pemodelan fungsi transfer adalah untuk menetapkan model yang sederhana, yang menghubungkan deret output ( $Y_t$ ) dengan deret input ( $X_t$ ) dan gangguan/noise ( $n_t$ ). Wei (1994) juga menjelaskan bahwa di dalam fungsi transfer terdapat rangkaian output yang mungkin dipengaruhi oleh rangkaian multiple input. Pada kasus single input peubah, dapat menggunakan metode korelasi silang yang dianjurkan oleh Box and Jenkins (1976). Teknik ini juga dapat digunakan ketika terdapat single input peubah yang lebih dari satu selama antar variable input tidak berkorelasi silang. Jika beberapa atau semua peubah input berkorelasi silang maka teknik prewhitening atau metode korelasi silang tidak dapat digunakan secara langsung. Alasan utama bagi perlunya suatu perencanaan atau peramalan adalah adanya tenggang waktu pengambilan keputusan yang dapat berkisar dari beberapa hari atau sampai beberapa tahun. Pada analisis fungsi transfer untuk peramalan deret berkala univariate, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (noise). Deret input mempengaruhi deret output melalui sebuah fungsi transfer yang mendistribusikan pengaruhnya secara dinamis melalui beberapa periode waktu yang akan datang dengan persentase tertentu yang disebut sebagai bobot respons impuls atau bobot fungsi transfer.

Model umum Fungsi Transfer:

$$y_t = \nu(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-s} + \frac{\theta_q(B)}{\varphi_p(B)} \varepsilon_t \quad \dots (4)$$

Dimana:

- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$
- $q \rightarrow$  ordo MA bagi noise  $N_t$

## f. *Vector Autoregression (VAR)*

*Vector Autoregression (VAR)* dikemukakan pertama kali oleh Christopher Sims (1980). Sims mengembangkan model ekonometri dengan mengabaikan pengujian asumsi secara apriori. VAR dikembangkan oleh Sims sebagai kritik atas metode simultan. Jumlah peubah yang besar dan klasifikasi endogen dan eksogen pada metode simultan merupakan dasar dari kritik tersebut. Penggunaan pendekatan struktural atas pemodelan persamaan simultan biasanya menerapkan teori ekonomi di dalam usahanya untuk mendeskripsikan hubungan antar peubah yang ingin diuji. Akan tetapi sering ditemukan bahwa teori ekonomi saja ternyata tidak cukup kaya di dalam menyediakan spesifikasi yang ketat dan tepat atas hubungan dinamis antar peubah (Gujarati, 2010). Model VAR merupakan jalan keluar atas permasalahan ini melalui pendekatan non-struktural atau tidak berdasarkan pada teori ekonomi tertentu (*atheoretical*). Metode VAR memperlakukan seluruh peubah secara simetris tanpa memperlakukan peubah dependen dan independen atau dengan kata lain model ini memperlakukan seluruh peubah sebagai peubah endogen, karena pada kenyataannya suatu peubah dapat bertindak sebagai endogen atau dependen dari suatu peubah yang lainnya (Gujarati, 2010).

Kelebihan dalam penggunaan metode VAR (Gujarati, 2010):

- n. Kemudahan dalam penggunaan, tidak perlu mengkhawatirkan tentang penentuan peubah endogen dan peubah eksogen.
- o. Kemudahan dalam estimasi, metode *Ordinary Least Square (OLS)* dapat diaplikasikan pada tiap persamaan secara terpisah.
- p. *Forecast* atau peramalan yang dihasilkan pada beberapa kasus ditemukan lebih baik daripada yang dihasilkan oleh model persamaan simultan yang kompleks.
- q. Perangkat estimasi yang digunakan adalah *Impulse Respon Function (IRF)* untuk melacak respon dari peubah dependen dalam sistem VAR terhadap *shock* dari *error term* dan *Variance Decomposition* yang memberikan informasi mengenai pentingnya masing-masing *error term* dalam mempengaruhi peubah-peubah dalam VAR.

Di sisi lain, terdapat beberapa kritik terhadap model VAR menyangkut permasalahan berikut (Gujarati, 2010) :

- 16) Model VAR merupakan model yang *atheoretical* atau tidak berdasarkan teori, hal ini tidak seperti pada persamaan simultan. Pada persamaan simultan, pemilihan peubah yang akan dimasukkan dalam persamaan memegang peranan penting dalam mengidentifikasi model.
- 17) Pada model VAR penekanannya terletak pada *forecasting* atau peramalan sehingga model ini kurang cocok digunakan dalam menganalisis kebijakan.
- 18) Permasalahan yang besar dalam model VAR adalah pada pemilihan *lag length* atau panjang lag yang tepat. Karena semakin panjang lag, maka akan menambah jumlah parameter yang akan bermasalah pada *degrees of freedom*.
- 19) Peubah yang tergabung pada model VAR harus stasioner. Apabila tidak stasioner, perlu dilakukan transformasi bentuk data, misalnya melalui *first difference*.
- 20) Sering ditemui kesulitan dalam menginterpretasi tiap koefisien pada estimasi model VAR, sehingga sebagian besar peneliti melakukan interpretasi pada estimasi fungsi *impulse respon*.

### **Estimasi Model VAR**

Enders (2004) menjelaskan, ketika peneliti tidak memiliki kepastian untuk menentukan bahwa suatu peubah adalah eksogen, maka suatu perluasan analisis fungsi perpindahan alami akan memperlakukan masing-

masing peubah secara simetris. Sebagai contoh, pada kasus-kasus peubah yang membiarkan alur waktu atau *time path*  $\{s_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{y_t\}$  dan membiarkan *time path*  $\{y_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{s_t\}$ .

Di dalam sistem *bivariate*, hubungan tersebut dapat digambarkan seperti pada persamaan 5 di bawah ini:

$$\begin{aligned} s_t &= b_{10} - b_{12}y_t + \gamma_{11}s_{t-1} + \gamma_{12}y_{t-1} + \varepsilon_{s_t} \\ y_t &= b_{20} - b_{21}s_t + \gamma_{21}s_{t-1} + \gamma_{22}y_{t-1} + \varepsilon_{y_t} \end{aligned} \quad \dots\dots(5)$$

Dengan mengasumsikan bahwa kedua peubah  $s_t$  dan  $y_t$  adalah stasioner:  $\varepsilon_{s_t}$  dan  $\varepsilon_{y_t}$  adalah *disturbances* yang memiliki rata-rata nol dan matriks kovarians terbatas atau bersifat *white noise* dengan standar deviasi yang berurutan  $\sigma_s$  dan  $\sigma_y$ ; serta  $\{\varepsilon_{s_t}\}$  dan  $\{\varepsilon_{y_t}\}$  adalah *disturbances* yang independen dengan rata-rata nol dan kovarian terbatas (*uncorrelated white-noise disturbances*). Kedua persamaan di atas merupakan orde pertama VAR, karena panjang *lag* nya hanya satu. Agar persamaan 5 lebih mudah dipahami dan digunakan sebagai alat analisis maka ditransformasikan dengan menggunakan matriks aljabar, dan hasilnya dapat dituliskan secara bersama seperti pada persamaan di bawah ini:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Atau dengan bentuk lain:

$$Bx_t = \Gamma_0 + \Gamma_1 X_{t-1} + \varepsilon_t \quad \dots\dots\dots(6)$$

Dimana:

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \quad x_t = \begin{bmatrix} s_t \\ y_t \end{bmatrix} \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix}$$

$$\Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Dengan melakukan pengalihan antara persamaan 6 dengan  $B^{-1}$  atau invers matriks B, maka akan dapat ditentukan model VAR dalam bentuk standar, seperti dituliskan pada persamaan di bawah ini:

$$x_t = A_0 + A_1 x_{t-1} + \ell_t \quad \dots\dots\dots(7)$$

$$A_0 = B^{-1} \Gamma_0$$

dimana  $A_1 = B^{-1} \Gamma_1$

$$\ell_t = B^{-1} \varepsilon_t$$

Untuk tujuan notasi, maka  $\{a_{i0}\}$  dapat didefinisikan sebagai elemen ke-i dari vektor  $A_0$ ;  $\{a_{ij}\}$  sebagai elemen dalam baris ke-i dan baris ke-j dari matriks  $A_1$ ; dan  $\{e_{it}\}$  sebagai elemen ke-i dari vektor  $e_t$ . Dengan menggunakan notasi baru yang telah dijelaskan sebelumnya, maka persamaan 7 dapat ditulis menjadi:

$$\begin{aligned} s_t &= a_{10} + a_{11}s_{t-1} + a_{12}y_{t-1} + e_{1t} \\ y_t &= a_{20} + a_{21}s_{t-1} + a_{22}y_{t-1} + e_{2t} \end{aligned} \quad \dots\dots\dots(8)$$

**Fungsi Impulse Response**

Fungsi *impulse response* menggambarkan tingkat laju dari *shock* peubah yang satu terhadap peubah yang lainnya pada suatu rentang periode tertentu. Sehingga dapat dilihat lamanya pengaruh dari *shock* suatu peubah terhadap peubah lain sampai pengaruhnya hilang atau kembali ke titik keseimbangan. Analisis fungsi *impulse respon* dapat dituliskan dalam bentuk *Vector Moving Avarage (VMA)* dari bentuk standar VAR pada persamaan 9.

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix} \dots\dots\dots (9)$$

dimana  $s_t$  dan  $y_t$  memiliki hubungan dengan  $e_{1t}$  dan  $e_{2t}$  secara berurutan. Selanjutnya dengan menggunakan operasi aljabar matriks maka *vector error* dapat ditentukan sebagai berikut:

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1-b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{st-i} \\ \varepsilon_{yt-i} \end{bmatrix} \dots\dots\dots (10)$$

dengan menggabungkan persamaan (9) dan (10) akan didapat:

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \frac{1}{1-b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{st-i} \\ \varepsilon_{yt-i} \end{bmatrix} \dots (11)$$

Persamaan 11 dapat disederhanakan dengan mendefinisikan matriks 2x2  $\Phi_i$  dengan elemen  $\Phi_{jk}$  (i) seperti persamaan berikut :

$$\Phi_i = A_1^i / (1-b_{12}b_{21}) \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \dots\dots\dots (12)$$

sehingga diperoleh bentuk matriks persamaan fungsi *impulse respon*:

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \sum_{i=0}^n \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{st-1} \\ \varepsilon_{yt-1} \end{bmatrix} \dots (13)$$

dimana :

- $\Phi_{ij}(i)$  = efek dari *structural shock* pada  $s$  dan  $y$
- $\Phi_{ij}(0)$  = *impact multipliers*
- $\sum \Phi_{ij}(i)$  = *cumulative multipliers*
- $\sum \Phi_{ij}(i)$  pada saat  $n \rightarrow \infty$  = *long run multipliers*

**Variance Decomposition**

*Variance decomposition* atau disebut juga *forecast error variance decomposition* merupakan perangkat pada model VAR yang akan memisahkan variasi dari sejumlah peubah yang diestimasi menjadi komponen-komponen *shock* atau menjadi peubah *innovation*, dengan asumsi bahwa peubah-peubah *innovation* tidak saling berkorelasi. Kemudian, *variance decomposition* akan memberikan informasi mengenai proporsi dari pergerakan pengaruh *shock* pada sebuah peubah terhadap *shock* peubah yang lain pada periode saat ini dan periode yang akan datang.

Bentuk VMA dari peubah  $x$  pada satu periode ke depan dapat dituliskan sbb.:

$$x_{t+1} = \bar{x} + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t+1-i} \dots\dots\dots (14)$$

*Forecast error* pada satu periode kedepan adalah:

$$E_t x_{t+1} = \bar{x} + \sum_{i=1}^{\infty} \phi_i \varepsilon_{t+1-i} \dots \dots \dots (15)$$

Forecast satu periode ke depan dilambangkan dengan  $\Phi_0 \varepsilon_{t+1}$ . Forecast error pada periode n ke depan adalah:

$$x_{t+n} - E_t x_{t+n} = \bar{x} + \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i} \dots \dots \dots (16)$$

Forecast error pada n periode ke depan untuk peubah s adalah:

$$s_{t+n} - E_t y_{t+n} = \phi_{11}(0) \varepsilon_{st+n} + \phi_{11}(1) \varepsilon_{st+n-1} + \dots + \phi_{11}(n-1) \varepsilon_{st+1} + \phi_{12}(0) \varepsilon_{yt+n} + \phi_{12}(1) \varepsilon_{yt+n-1} + \dots + \phi_{12}(n-1) \varepsilon_{yt+1} \quad (17)$$

Variance dari forecast error  $s_{t+n}$  periode n ke depan adalah  $\sigma_s(n)^2$  dimana:

$$\sigma_s(n)^2 = \sigma_s^2 [\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2] + \sigma_y^2 [\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2] \dots \dots \dots (18)$$

Forecast error variance decomposition adalah proporsi dari  $\sigma_s(n)^2$  terhadap shock s dan shock y. Sehingga forecast error variance decomposition pada shock s adalah:

$$\sigma_s^2 [\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2] / \sigma_s(n)^2 \dots \dots \dots (19)$$

sedangkan forecast error variance decomposition pada shock y adalah:

$$\sigma_y^2 [\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2] / \sigma_y(n)^2 \dots \dots \dots (20)$$

## HASIL PEMBAHASAN

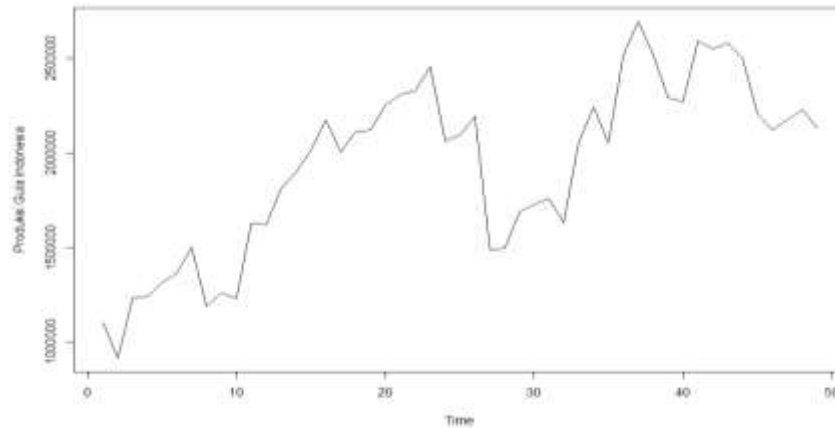
Estimasi produksi gula Indonesia dilakukan guna mendapatkan estimasi tahun 2021 – 2025, menggunakan pendekatan 3 metode, yakni ARIMA, VAR dan Fungsi Transfer. Metode terbaik dipilih berdasarkan signifikansi parameter statistik, nilai MAPE terkecil dan performa hasil estimasi yang realistis.

### 1. Estimasi Metode ARIMA

Syarat analisis menggunakan metode ARIMA adalah kestasioneran data. Hasil uji Augmented Dickey-Fuller data produksi gula series tahun 1972-2020 baru stasioner dengan pembedaan (*differencing*) tingkat 2, seperti ditunjukkan dari nilai statistik uji lebih kecil dari nilai kritis sebagai berikut:

```
Augmented Dickey-Fuller Test Unit Root Test:
Value of test-statistic is: -5.7408 11.0043 16.484

Critical values for test statistics:
      1pct  5pct 10pct
tau3  -4.15 -3.50 -3.18
phi2   7.02  5.13  4.31
phi3   9.31  6.73  5.61
```



Gambar 1. Perkembangan Produksi Gula Indonesia, 1972-2020

Data series produksi gula tahun 1972-2020 dibagi menjadi series *data training* (1972 – 2014) yang digunakan untuk mencari model terbaik dan series *data testing* (2015-2020) untuk validasi model.

Penentuan ordo lag  $p$  (AR) dan  $q$  (MA) dilakukan melalui investigasi plot ACF dan PACF namun belum terlihat kriteria lag AR dan MA-nya. Dengan memanfaatkan *script* pada software R-Studio ditunjukkan 10 model ARIMA *tentative* berdasarkan nilai SBC terkecil seperti tersaji pada Tabel 2.

**Tabel 2. Model ARIMA Tentative Produksi Gula dengan Differencing Tingkat 2**

No.	p	q	sbc
[1,]	5	4	1022.915
[2,]	5	5	1023.268
[3,]	3	0	1025.772
[4,]	2	0	1025.853
[5,]	4	0	1028.636
[6,]	5	2	1029.269
[7,]	5	3	1029.642
[8,]	5	1	1030.408
[9,]	4	1	1031.061
[10,]	5	0	1032.769

**Tabel 3. Signifikansi Lag AR dan MA pada Model ARIMA Produksi Gula**

Model	Signifikansi									
	ar1	ar2	ar3	ar4	ar5	ma1	ma2	ma3	ma4	ma5
(5,2,4)	*	***						*		
(5,2,5)	-	-								
(3,2,0)	***	***	*							
(2,2,0)	***	***								
(4,2,0)	***	***	*	-						
(5,2,2)	-	-	-	-	-	-	-			
(5,2,3)	**	**	-	-	*	-	-	**		
(5,2,1)	-	-	-	-	-	***				
(4,2,1)	-	-	-	-	-	***				
(5,2,4)	*	***	-	-	-	-	-	*	-	

Keterangan : Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Model akhir yang dipilih didasarkan pada signifikansi pada semua komponen AR dan MA, yakni model ARIMA (3,2,0). Hasil uji signifikansi terhadap komponen AR1, AR2, dan AR3 adalah sebagai berikut:

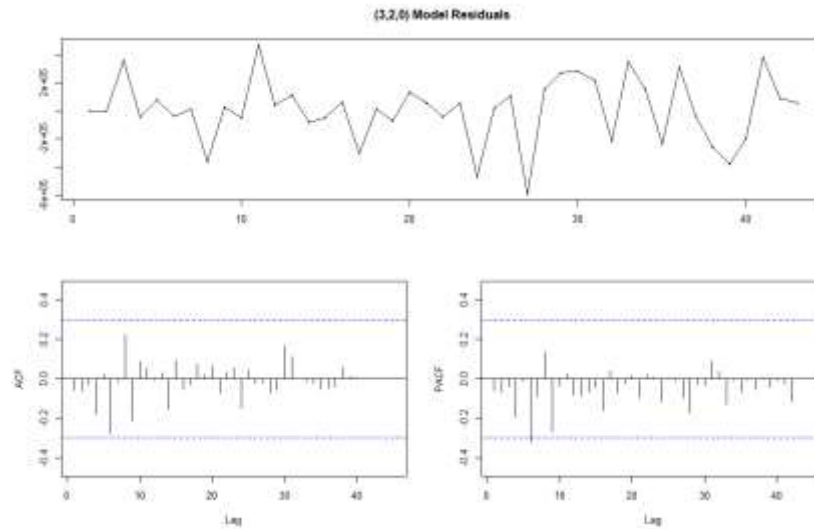
```

call:
arima(x = train[, "Produksi"], order=c(3,2,0))
Coefficients:
      ar1      ar2      ar3
s.e.   -0.8667  -0.6813  -0.2848
       0.1415   0.1578   0.1391

Estimate Std. Error z value Pr(>|z|)
ar1 -0.86669    0.14146 -6.1267 8.973e-10 ***
ar2 -0.68127    0.15785 -4.3160 1.589e-05 ***
ar3 -0.28476    0.13910 -2.0471 0.04065 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
    
```

Pemeriksaan sisaan terhadap model ARIMA (3,2,0) menunjukkan pola terdistribusi normal serta pola ACF dan PACF sisaan yang tidak nyata, seperti tersaji pada Gambar 2. Hasil Uji Ljung-Box yang mengindikasikan autokorelasi sisaan tidak signifikan hingga 30 lag (Tabel 4).





Gambar 2. Plot Sisaan Model ARIMA (3,2,0) Produksi Gula

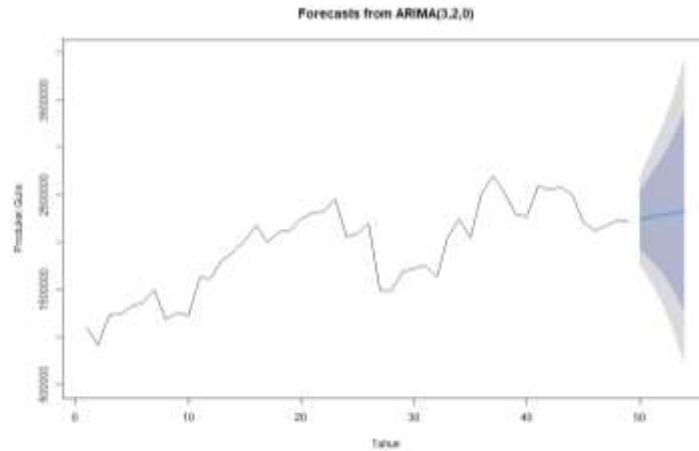
Tabel 4. Hasil Uji Ljung-Box ARIMA (3,2,0)

lags	statistic	df	p-value
5	2.061255	5	0.840606
10	11.90832	10	0.29124
15	14.39452	15	0.495843
20	15.50205	20	0.746996
25	18.83221	25	0.805094
30	24.06578	30	0.76904

Setelah asumsi sisaan telah memenuhi syarat, maka dilakukan peramalan yang menghasilkan nilai MAPE sebesar 9,10% untuk data *training* dan 6,28% untuk data *testing*.

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	4197.06	224452.60	168206.30	0.16	9.10	1.12	-0.05
Test set	-54692.50	155567.40	142465.00	-2.76	6.28	0.95	NA

Ramalan produksi gula tahun 2021–2025 menggunakan metode ARIMA (3,2,0) serta plot hasil ramalannya tersaji pada Tabel 5 dan Gambar 3.



Gambar 3. Hasil Ramalan Produksi Gula dengan Metode ARIMA (3,2,0)

**Tabel 5. Ramalan Produksi Gula Indonesia dengan Metode ARIMA (3,2,0)**

Point	Forecast	Lo 80	Hi 80	Lo 95	Hi 95
50	2234679	1931128	2538231	1770438	2698922
51	2271007	1810567	2731446	1566826	2975188
52	2291881	1675471	2908291	1349164	3234598
53	2304788	1480655	3128920	1044386	3565189
54	2327159	1259737	3394581	694678.5	3959640

Produksi gula Indonesia tahun 2021 berdasarkan model ARIMA (3,2,0) diperkirakan mencapai 2,23 juta ton atau naik 0,36% dibanding tahun sebelumnya. Produksi gula diperkirakan terus mengalami peningkatan hingga tahun 2025 menjadi 2,33 juta ton. Data series yang digunakan untuk penelusuran model ARIMA adalah dari tahun 1972 hingga 2020 sehingga belum melibatkan pabrik gula baru yang sudah mulai giling tebu dan memproduksi gula pada tahun 2021. Oleh karenanya nilai ramalan produksi pada Tabel 5 belum termasuk kontribusi dari 3 (tiga) pabrik gula baru yang mulai berproduksi pada tahun 2021.

## 2. Estimasi Metode Fungsi Transfer

Dengan asumsi bahwa produksi gula Indonesia sangat dipengaruhi oleh harga gula di pasar dunia (Pelelangan Amerika Serikat), maka dilakukan pemodelan Fungsi Transfer dengan menyertakan harga gula dunia sebagai peubah input.

Data deret input harga gula dunia mempunyai tren meningkat dari waktu ke waktu atau tidak stasioner sehingga harus dilakukan pembedaan/*differencing* sebagai syarat untuk melakukan pemodelan ARIMA. Model ARIMA data deret input yang kemudian akan diikutsertakan pada pemodelan Fungsi Transfer. Pembedaan tingkat 1 menggunakan uji Augmented Dickey-Fuller (ADF) telah diperoleh series data harga gula dunia yang stasioner, sebagai berikut:

```
Augmented Dickey-Fuller Test Unit Root Test
Value of test-statistic is: -7.4934 19.402 28.8459
Critical values for test statistics:
  1pct  5pct 10pct
tau3  -4.15 -3.50 -3.18
phi2   7.02  5.13  4.31
phi3   9.31  6.73  5.61
```

Penelusuran model ARIMA untuk deret input Harga Gula Dunia menggunakan script RStudio ditunjukkan 10 model tentative berdasarkan nilai SBC terkecil sebagai berikut:

**Tabel 6. Model Tentative Peubah Input Harga Gula Dunia**

No	p	q	sbc
[1,]	2	1	-239.322
[2,]	3	1	-238.557
[3,]	4	1	-235.347
[4,]	2	2	-233.86
[5,]	3	2	-233.148
[6,]	5	1	-231.753
[7,]	4	2	-230.431
[8,]	2	3	-230.404
[9,]	1	4	-230.045
[10,]	0	4	-229.271

**Tabel 7. Signifikansi Lag AR dan MA pada Model Peubah Input Harga Gula Dunia**

Model	Signifikansi									
	ar1	ar2	ar3	ar4	ar5	ma1	ma2	ma3	ma4	
(2,1,1)	*	**				***				
(3,1,1)	-	*	-			*				
(4,1,0)	*	***	**	.						
(2,1,2)	na	***				na	na			
(3,1,2)	***	***				***	***			
(5,1,1)	***	***	***	***	***	***				
(4,1,2)	-	*	-	-		-	-			
(2,1,3)	***	-				***	**	***		
(1,1,4)	-					-	***	-	-	
(0,1,4)						-	***		*	

Keterangan : Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Berdasarkan model tentative tersebut, diperoleh bahwa model **ARIMA (2,1,1)** merupakan model terbaik dengan komponen AR dan MA yang nyata seperti tersaji pada Tabel 7. Signifikansi model ARIMA (2,1,1) peubah input harga gula dunia sebagai berikut:

```

Series: train[, "Harga_US"] ARIMA(2,1,1)
Coefficients:
      ar1      ar2      ma1
 0.4395 -0.4728 -0.7244
s.e. 0.1709 0.1647 0.1473

sigma^2 estimated as 0.009978: log likelihood=38.12
AIC=-68.23 AICc=-67.15 BIC=-61.28

Training set error measures:
      ME      RMSE      MAE
Training set 0.02520334 0.09513155 0.05329156

MPE      MAPE      MASE      ACF1
3.163427 10.09634 0.7881146 -0.119905

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1 0.43948 0.17095 2.5709 0.010143 *
ar2 -0.47279 0.16470 -2.8705 0.004098 **

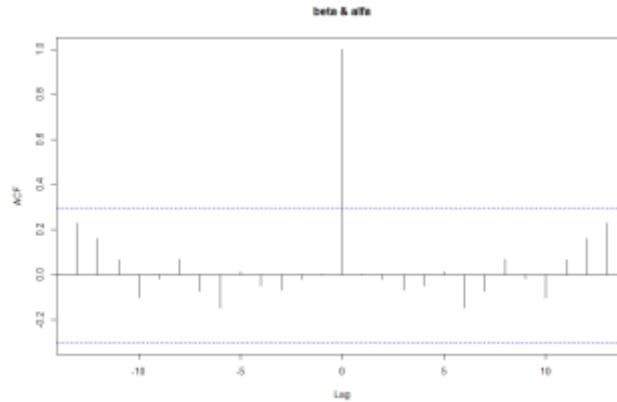
```

```

ma1 -0.72437    0.14729 -4.9181 8.741e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Kedua series data baik deret input (harga gula dunia) maupun deret output (produksi gula), mempunyai tren atau tidak stasioner, sehingga menghasilkan hubungan yang semu. Oleh karenanya, perlu dilakukan analisis korelasi silang terhadap kedua data tersebut dengan hasil sbb.:



Gambar 4. Plot Korelasi Silang Deret Input Harga Gula Dunia dengan Produksi Gula

Plot korelasi antara data series menunjukkan nyata pada lag=0 atau nilai b=0 yang artinya tidak ada jeda pengaruh harga gula dunia terhadap produksi gula Indonesia, dan nilai s=0 atau tidak ada panjang pengaruh harga gula dunia terhadap produksi gula Indonesia. Nilai r=0 atau tidak ada pengaruh produksi gula periode sebelumnya terhadap produksi gula periode ini karena adanya perubahan harga gula dunia. Hal ini mengingatkan data series produksi gula Indonesia merupakan data tahunan.

Selanjutnya, untuk menghasilkan ordo yang paling tepat guna menentukan ordo Fungsi Transfer dilakukan penelusuran model *noise* menggunakan model ARIMA. Pada analisis fungsi transfer untuk peramalan deret berkala univariate, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (*noise*). Model **ARIMA (1,1,3)** merupakan model terbaik untuk data series *noise*, dengan signifikansi masing-masing koefisien adalah sebagai berikut:

```

Series: res
ARIMA(1,1,3)

Coefficients:
      ar1      ma1      ma2      ma3
      -0.7424  1.1907  -0.5288  -0.7769
s.e.      0.1397  0.1735  0.2205  0.1515

sigma^2 estimated as 5.326e+10: loglikelihood=-578.85
AIC=1167.69  AICC=1169.36  BIC=1176.38

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1 -0.74244    0.13972  -5.3138 1.074e-07 ***
ma1  1.19067    0.17352  6.8620 6.792e-12 ***
ma2 -0.52875    0.22045  -2.3985 0.01646 *
ma3 -0.77689    0.15145  -5.1295 2.905e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Diagnosa kelayakan model dilakukan melalui pengepasan model berbasis Fungsi Transfer dengan derajat (r,b,s) = (0,0,0) dan model noise ARIMA (1,1,3) terhadap produksi gula dengan menggunakan data aktual harga gula dunia periode data tahun 2015-2020, diperoleh nilai **MAPE = 9,24%**

```

Regression with ARIMA(1,1,3) errors
Coefficients:
      ar1      ma1      ma2      ma3      xreg
0.0263 -0.3179 0.1746 0.3388 -498329.2
s.e. 0.3946 0.3749 0.2135 0.1793 264826.5

sigma^2 estimated as 4.703e+10: log likelihood=-573.33
AIC=1158.65 AICC=1161.05 BIC=1169.08

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 33341.38 201162.2 160636.5 1.377168 9.23783 1.015139 -0.03037221

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1 2.6334e-02 3.9460e-01 0.0667 0.94679
ma1 -3.1795e-01 3.7490e-01 -0.8481 0.39639
ma2 1.7464e-01 2.1353e-01 0.8179 0.41344
ma3 3.3877e-01 1.7926e-01 1.8898 0.05878
xreg -4.9833e+05 2.6483e+05 -1.8817 0.05987
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

Model selanjutnya adalah dengan menguji model Fungsi Transfer ini untuk melakukan peramalan dengan peubah input menggunakan data aktual yang menghasilkan MAPE data testing sebesar 4,00%, sebagai berikut:

```

Series: test.h[, "Produksi"]
Regression with ARIMA(1,1,3) errors
Coefficients:
      ar1      ma1      ma2      ma3      xreg
0.0263 -0.3179 0.1746 0.3388 -498329.2
s.e. 0.0000 0.0000 0.0000 0.0000 0.0

sigma^2 estimated as 4.703e+10: log likelihood=-66.57
AIC=135.14 AICC=136.48 BIC=134.75

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -34796.45 125974.3 87625.48 -1.629892 3.999649 0.7579692 0.2383744

```

Apabila peubah input harga gula dunia dilakukan peramalan terlebih dahulu, maka model Fungsi Transfer ini menghasilkan MAPE data testing sebesar 4,60% sebagai berikut:

```

Series: test.h[, "Produksi"]
Regression with ARIMA(1,1,3) errors
Coefficients:
      ar1      ma1      ma2      ma3      xreg
0.0263 -0.3179 0.1746 0.3388 -498329.2
s.e. 0.0000 0.0000 0.0000 0.0000 0.0

sigma^2 estimated as 4.703e+10: log likelihood=-67.12
AIC=136.23 AICC=137.56 BIC=135.84

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -39019.71 140461.4 100664.1 -1.827921 4.602187 0.8707543 0.2817625

```

Guna melakukan peramalan produksi gula harus terlebih dahulu dilakukan peramalan variabel input harga gula dunia dengan model ARIMA yang telah diperoleh yakni ARIMA (2,1,1).

Hasil ramalan produksi gula tahun 2021 – 2025 menggunakan model Fungsi Transfer dengan derajat (r,b,s) = (0,0,0) dan model noise ARIMA (1,1,3) adalah sebagai berikut:

Time Series:

Start = 47

End = 51

Frequency = 1

[1] 2182565 2274283 2283864 2282398 2280372

**Tabel 8. Produksi Gula Indonesia Hasil Peramalan Model Fungsi Transfer, 2021 - 2025**

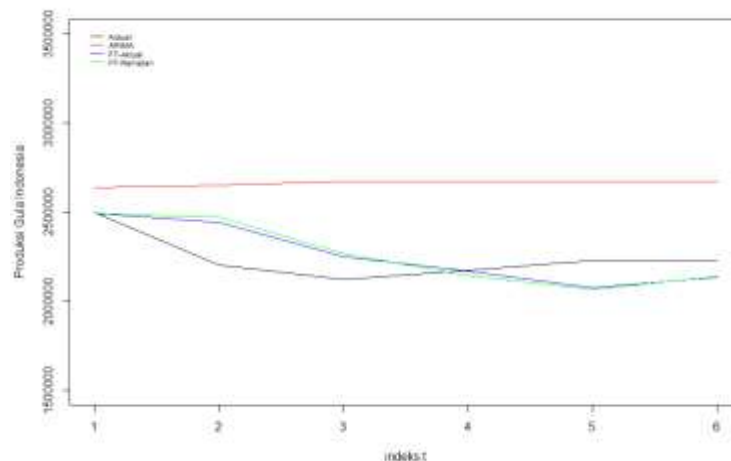
Tahun	Produksi (Ton)	Pertumbuhan (%)
2020	2,130,719	
2021	2,182,565	2.43
2022	2,274,283	4.20
2023	2,283,864	0.42
2024	2,283,398	-0.02
2025	2,280,372	-0.13
Rata-rata Pertumbuhan (%)		1.38

Keterangan : 2020 ATAP

2021-2025 Estimasi Model Fungsi Transfer

Produksi gula Indonesia tahun 2021 diperkirakan mencapai 2,18 juta ton atau naik 2,43% dibandingkan Angka Tetap 2020. Produksi gula Indonesia tahun 2025 diperkirakan mencapai 2,28 juta ton atau naik dengan rata-rata 1,38% per tahun. Data series yang digunakan untuk penelusuran model Fungsi Transfer adalah dari tahun 1972 hingga 2020 sehingga belum melibatkan pabrik gula baru yang sudah mulai giling tebu dan memproduksi gula pada tahun 2021. Oleh karenanya nilai ramalan produksi pada Tabel 8 belum termasuk kontribusi dari 3 (tiga) pabrik gula baru yang mulai berproduksi pada tahun 2021.

Plot hasil ramalan dengan menggunakan metode Fungsi Transfer menggunakan data actual dan hasil ramalan, serta hasil peramalan model ARIMA disajikan pada gambar di bawah ini.



**Gambar 5. Plot Perbandingan Hasil Ramalan Produksi Gula dengan Metode Fungsi Transfer dengan Nilai Aktual dan Metode ARIMA**

### 3. Estimasi Metode *Vector Auto Regression* (VAR)

Model VAR merupakan alat analisis yang sangat berguna dalam memahami adanya hubungan timbal balik (interrelationship) antara peubah ekonomi maupun dalam pembentukan ekonomi yang berstruktur. Ramalan produksi gula menggunakan model VAR akan melibatkan peubah harga gula dunia, ekspor molasses dan impor gula Indonesia, harga konsumen gula dan harga produsen jagung ke dalam sistem persamaan VAR. Model VAR dapat digunakan untuk menjelaskan fenomena semua peubah yang masuk ke dalam sistem, namun dalam sub bab ini lebih dibahas fenomena terhadap peubah produksi gula Indonesia, dan menggunakan hasil model VAR untuk peramalan produksi tahun 2021-2025.

Penelusuran model VAR dilakukan mulai dari lag  $p=1$  hingga  $p=5$  dengan mengikutsertakan tren dan atau konstanta. Hasil penelusuran diperoleh bahwa VAR dengan lag=3 dengan hanya mengikutsertakan konstanta merupakan model terbaik terbukti adanya signifikansi untuk peubah tersebut hingga lag ke-3, seperti tersaji pada Tabel 9.

**Tabel 9. Signifikansi Masing-masing Peubah pada Model VAR  $p=3$  Type=constant**

Signifikansi	Model						
	Produksi	Harga_EU	Harga_US	Ekspor_Molases	Impor_Gula	Hrg_Kons_Gula	Hprod_Jagung
Produksi.11		*			*		
Harga_EU.11		**					
Harga_US.11		*					
Ekspor_Molases.11				**			
Impor_Gula.11				**		***	
Hrg_Kons_Gula.11		*		*			
Hprod_Jagung.11		*				**	**
Produksi.12							
Harga_EU.12							
Harga_US.12							
Ekspor_Molases.12							
Impor_Gula.12						*	
Hrg_Kons_Gula.12							
Hprod_Jagung.12		*				*	
Produksi.13							
Harga_EU.13							
Harga_US.13		*					
Ekspor_Molases.13							
Impor_Gula.13					*		
Hrg_Kons_Gula.13						**	*
Hprod_Jagung.13							
const	*	*	*				

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Model VAR terpilih yakni  $p=3$  type=constant kemudian dilakukan uji normalitas terhadap sisaan, homokedastisitas dan non autokorelasi sudah terpenuhi sebagai berikut:

- Nilai p-value pada uji Portmanteau lebih besar dari 0,05 menunjukkan bahwa sisaan saling bebas atau asumsi mon autokorelasi terpenuhi

```
Portmanteau Test (asymptotic)
data: Residuals of VAR object vargula.c3
Chi-squared = 669.6, df = 637, p-value = 0.1796
```

- Pemeriksaan normalitas menggunakan Uji Jarque-Bera (JP-Test) menghasilkan nilai p-value yang lebih kecil dari 0,05, namun karena data series yang digunakan relatif panjang, maka data series ini dianggap normal.

```
JB-Test (multivariate)
data: Residuals of VAR object vargula.c3
Chi-squared = 82.431, df = 14, p-value = 9.993e-12
```

Skewness only (multivariate)  
 data: Residuals of VAR object vargula.c3  
 Chi-squared = 28.306, df = 7, p-value = 0.0001935

Kurtosis only (multivariate)  
 data: Residuals of VAR object vargula.c3  
 Chi-squared = 54.125, df = 7, p-value = 2.222e-09

- Pemeriksaan heterokedastisitas model VAR dilakukan dengan pengujian ARCH-LM, menghasilkan nilai p-value yang lebih besar dari 0,05 menunjukkan bahwa ragam sisaan model VAR ini sudah homogen atau asumsi heterokedastisitas terpenuhi.

ARCH (multivariate)

data: Residuals of VAR object vargula.c3  
 Chi-squared = 980, df = 3920, p-value = 1

Model VAR p=3 type=constant menghasilkan nilai MAPE sebagai berikut:

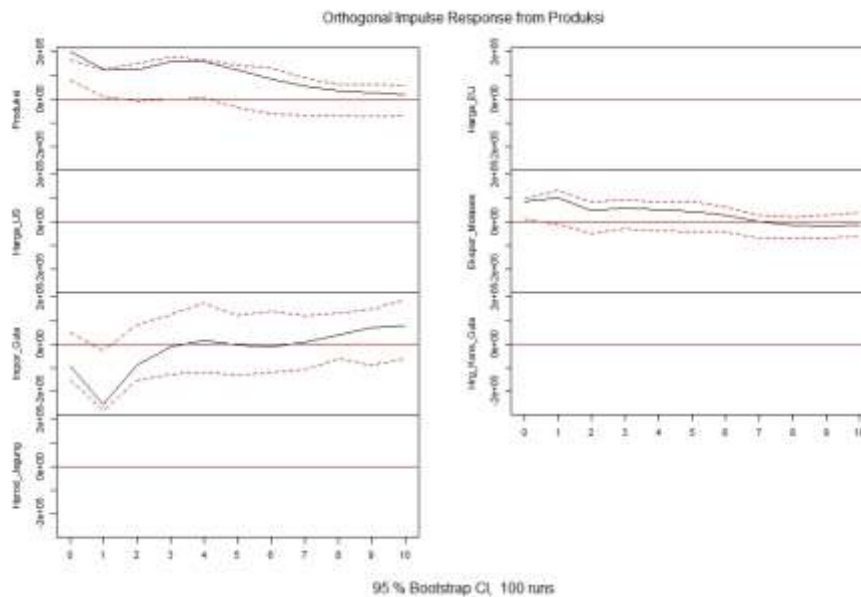
- MAPE series data testing sebesar 6,13%

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.009834	5.423913	7.297927	6.129664	8.316636	9.600008

- MAPE series data training sebesar 5,81%

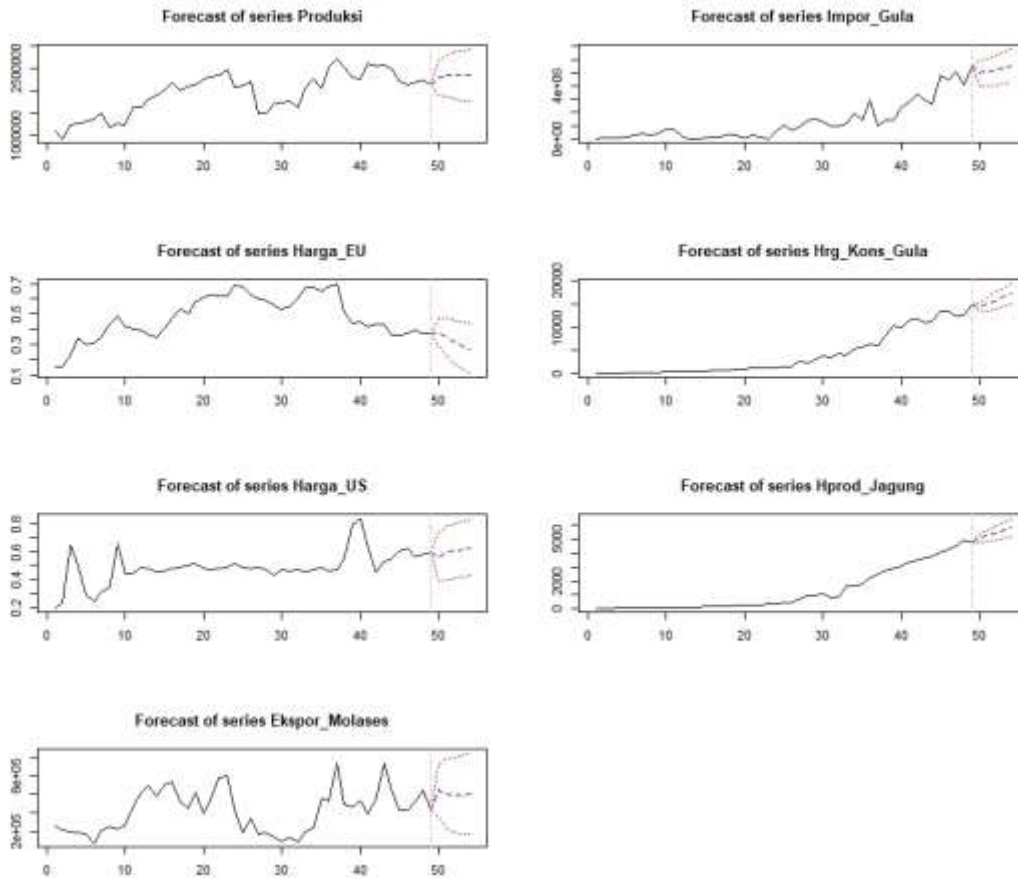
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.4381	1.8653	4.4394	5.8061	7.7665	21.0996

Hasil grafik orthogonal impulse response menunjukkan bahwa apabila ada perubahan produksi gula Indonesia pada tahun tertentu, maka produksi gula akan merespon hingga 1 tahun ke depan, demikian juga ekspor molasses dan impor gula akan terpengaruh hingga 1 tahun ke depan. Perubahan produksi gula Indonesia tidak berdampak pada harga gula dunia maupun harga konsumen gula Indonesia.



Gambar 6. Grafik Orthogonal Response Impuls Dampak Produksi Gula Indonesia





Gambar 7. Grafik Pola Plot Ramalan Model VAR dengan p=2

Model VAR dengan p=3 dengan hanya mempertimbangkan konstanta digunakan untuk melakukan peramalan data produksi gula Indonesia dengan hasil tersaji pada Tabel 10.

**Tabel 10. Hasil Ramalan Produksi Gula dengan Metode VAR p=3 type=constant**

Tahun	Produksi (Ton)	Pertumbuhan (%)
2020	2,130,719	
2021	2,278,141	6.92
2022	2,320,887	1.88
2023	2,323,539	0.11
2024	2,329,235	0.25
2025	2,338,374	0.39
Rata-rata pertumbuhan (%)		1.91

Keterangan : 2020 ATAP

2021-2025 Estimasi Model VAR

Produksi gula Indonesia tahun 2021 berdasarkan hasil model VAR diperkirakan mencapai 2,28 juta ton atau naik 6,92% dibandingkan Angka Tetap 2020. Produksi gula Indonesia tahun 2025 diperkirakan mencapai 2,34 juta ton atau naik dengan rata-rata 1,91% per tahun. Data series yang digunakan untuk penelusuran model VAR adalah dari tahun 1972 hingga 2020 sehingga belum melibatkan pabrik gula baru yang sudah mulai giling tebu dan memproduksi gula pada tahun 2021. Oleh karenanya nilai ramalan produksi pada Tabel 10 belum termasuk kontribusi dari 3 (tiga) pabrik gula baru yang mulai berproduksi pada tahun 2021.

## SIMPULAN DAN SARAN

Berdasarkan atas penelusuran tiga metode yang digunakan untuk memodelkan dan meramalkan data produksi gula Indonesia dapat dilihat perbandingan keterandalannya seperti tersaji pada Tabel 11.

Tabel 11. Keterandalan Tiga Metode Peramalan Produksi Gula

No	Model	Model/ Peubah input	MAPE		ATAP	Estimasi Produksi Gula (Ton)					Rata-rata pertumb. 2020-2025 (%)	
			Training	Testing	2020	2021	2022	2023	2024	2025		
1	ARIMA	(3,2,0)	6,28	9,10	2,130,719	2,234,679	2,271,007	2,291,881	2,304,788	2,327,159		
	Pertumbuhan (%)					4.88	1.63	0.92	0.56	0.97		1.79
2	FUNGSI TRANSFER	Harga Gula Dunia Model: peubah input=ARIMA (2,1,1)	9,24	4,00	2,130,719	2,182,565	2,274,283	2,283,864	2,282,398	2,280,372		
	Pertumbuhan (%)					2.43	4.20	0.42	-0.06	-0.09		1.38
3	VAR p=3, type= constant	harga dunia, impor gula, ekspor molases, harga konsumen gula, harga jagung	5,81	6,21	2,130,719	2,278,141	2,320,887	2,323,539	2,329,235	2,338,374		
	Pertumbuhan (%)					6.92	1.88	0.11	0.25	0.39		1.91

Keterangan: ATAP = Angka Tetap dari Ditjen Perkebunan

Berdasarkan atas keragaan diatas, dapat disimpulkan bahwa Model Fungsi Transfer merupakan model terbaik karena mempunyai nilai MAPE yang terkecil dengan hasil ramalan produksi gula yang cukup realistis. Ramalan produksi gula Indonesia tahun 2021 menggunakan model Fungsi Transfer diperkirakan akan mencapai 2,18 juta ton atau naik 2,43% dibandingkan Angka Tetap 2020. Produksi gula Indonesia diperkirakan akan terus mengalami peningkatan hingga menjadi 2,28 juta ton pada tahun 2025 atau naik dengan rata-rata 1,38% per tahun.

Berdasarkan simpulan diatas, maka beberapa saran yang diajukan adalah sbb.;

- Model Fungsi Transfer sudah mempunyai keterandalan untuk memodelkan data produksi gula Indonesia dan meramalkan beberapa tahun ke depan, namun disarankan untuk dicobakan menggunakan peubah input lainnya apabila ketersediaan data series memenuhi.
- Perlu dikaji metode estimasi yang bisa melibatkan adanya tambahan produksi dari pabrik gula yang baru berkontribusi pada tahun 2021, sementara data series yang digunakan baru sampai tahun 2020.

## DAFTAR PUSTAKA

- Enders, W. 2010. Applied Econometric Time Series. USA: University of Alabama. Wiley, Third Edition.
- Guha, B and Bandyopadhyay, G. 2016. Gold Price Forecasting Using ARIMA Model. Journal of Advanced Management Science Vol. 4, No. 2, March 2016
- Gujarati, D.N. and D.C. Porter, 2010. Dasar-dasar Ekonometrika. Jakarta: Penerbit Salemba Empat, Buku 2, Edisi 5
- Wei, W. W. S. Time Series Analysis: Univariate and Multivariate Methods. California: Addison-Wesley Publishing Company. 1994.



# KAJIAN METODE ESTIMASI PRODUKSI TEH INDONESIA

Diah Indarti - *Statisticians*

*Center for Agricultural Data and Information System - Ministry of Agriculture  
Jln. Harsono RM 3 Gd D Lt 4, Ragunan-Jakarta Selatan, Indonesia*

## ABSTRAK

Komoditas teh adalah salah satu komoditas perkebunan penghasil devisa setelah minyak dan gas, hal ini ditunjang dengan luasan perkebunan teh di Indonesia tergolong cukup luas. Tujuan dari analisis kajian produksi teh Indonesia ini adalah mencari angka estimasi tahun 2022 dengan menggunakan model *Autoregressive Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Autoregression* (VAR) dengan menggunakan software RStudio. Tujuan dari ketiga model ini mencari keakurasian data agar menjadi lebih baik yang ditandai dengan *Mean Absolut Percentage Error* (MAPE) untuk data training maupun data testing semakin kecil. Berdasarkan dari hasil perbandingan MAPE disimpulkan bahwa model yang digunakan adalah model VAR (2)  $p=2$  type "both" dengan MAPE training sebesar 5,79% dan data testing sebesar 6,74% dengan hasil estimasi produksi teh Indonesia tahun 2022 sebesar 146,33 ribu ton.

*Kata kunci: teh, ARIMA, fungsi transfer, var, MAPE*

## ABSTRACT

*Tea commodity is one of the foreign exchange-producing plantation commodities after oil and gas, this is supported by the relatively large area of tea plantations in indonesia. The purpose of this analysis of the study of indonesian tea production is to find the estimated number for 2022 using the autoregressive integrated moving average (arima), transfer function and vector autoregression (var) model using the rstudio software. The purpose of these three models is to look for data accuracy to be better, which is marked by the mean absolute percentage error (mape) for training data and testing data is getting smaller. Based on the results of the mape comparison, it is concluded that the model used is the var (2)  $p=2$  type "both" model with mape training of 5.79% and testing data of 6.74% with the estimated indonesian tea production in 2022 of 146, 33 thousand tons.*

*Keywords: tea, ARIMA, transfer function, var, MAPE*

## PENDAHULUAN

Teh merupakan salah satu minuman yang sangat populer dikonsumsi di banyak negara. Teh sebagai salah satu komoditi hasil perkebunan yang mempunyai peran cukup penting dalam kegiatan perekonomian di Indonesia, yakni sebagai salah satu penghasil devisa negara sesudah minyak dan gas. Hal ini ditunjang dengan perkebunan teh di Indonesia cukup luas dengan jumlah produksi teh yang sangat besar (Setyamidjaja, 2001).

Estimasi produksi teh beberapa tahun ke depan sangat diperlukan sebagai bahan penentu kebijakan maupun *Early Warning System (EWS)* mengingat data Angka Tetap (ATAP) Perkebunan memiliki lag yang cukup jauh dibandingkan data tahun berjalan. Pada kajian metode estimasi produksi teh Indonesia ini menggunakan model *univariate* maupun model *multivariate*. Metode yang digunakan antara lain *Autoregressive Integrated Moving Average (ARIMA)*, Fungsi Transfer dan *Vector Autoregression (VAR)*. Metode *Autoregressive Integrated Moving Average (ARIMA)* menghasilkan angka estimasi produksi teh tanpa ada pengaruh dari variabel lain. Metode Fungsi Transfer menghasilkan angka estimasi produksi teh dengan memasukkan intervensi dari satu variabel pendukung yang dianggap paling berpengaruh terhadap produksi. Metode VAR menghasilkan angka estimasi produksi teh dengan mempertimbangkan pengaruh dari beberapa variabel lain atau lebih dari satu variabel pendukung yang diduga berpengaruh terhadap produksinya.

Saat ini rilis resmi data produksi teh oleh Badan Pusat Statistik (BPS) adalah data tahunan dengan lag  $n-1$  tahun. Guna merumuskan kebijakan produksi, pasokan dan distribusi komoditas teh sangat dituntut ketersediaan data terkini, bahkan ramalan beberapa periode ke depan. Metode ramalan produksi teh dapat digunakan dengan menggunakan beberapa metode statistik yang telah dikembangkan oleh para ahli.

Pada makalah ini akan dikaji metode ARIMA, Fungsi Transfer dan VAR dalam melakukan pemodelan dan peramalan produksi teh Indonesia.

Tujuan dari disusunnya makalah ini adalah:

- a. Melakukan analisis dan peramalan data produksi teh menggunakan model ARIMA, Regresi, Fungsi Transfer dan VAR di Indonesia
- b. Membandingkan ketiga metode tersebut dalam memperoleh ramalan data produksi teh.
- c. Menentukan metode terbaik dalam meramal data produksi teh di Indonesia.

## METODOLOGI

### 1. Sumber Data

Data yang digunakan dalam kajian ini adalah data sekunder yang diperoleh dari Badan Pusat Statistik dan Direktorat Jenderal Perkebunan. Data-data tersebut terdiri dari variabel luas areal teh, produksi teh, luas tanaman menghasilkan teh, volume ekspor teh, volume impor teh dari tahun 1980 hingga tahun 2020. Data tahun 1980 sampai tahun 2014 digunakan untuk membangun model sedangkan data tahun 2015 sampai tahun 2020 digunakan untuk validasi model.

### 2. Software

Software yang digunakan dalam menyusun makalah ini menggunakan software aplikasi R-Studio. Keunggulan dari software R-Studio ini merupakan software yang open source sehingga tidak memerlukan biaya untuk pembelian maupun perpanjangan lisensi. Keunggulan lain dari R adalah mudah dalam melakukan transformasi dan pemrosesan data. Karena R adalah program untuk analisis data, maka kemampuan R dalam transformasi data seperti penyiapan data, import dan export data dalam berbagai format, dan lain-lain.

### 3. Tinjauan Literature

Metode statistik yang digunakan dalam peramalan ini menggunakan peubah tunggal maupun peubah ganda. Metode yang digunakan antara lain ARIMA (*Auto Regressive Integrated Moving Average*), Fungsi Transfer dan VAR (*Vector Autoregression*).

#### **a. *Autoregressive Integrated Moving Average (ARIMA)***

*Autoregressive Integrated Moving average (ARIMA)* atau biasa disebut juga dengan metode time series Box Jenkins, sangat sesuai digunakan untuk melakukan peramalan jangka pendek, sementara untuk peramalan jangka panjang kurang baik ketepatannya. Metode ARIMA merupakan metode yang hanya menggunakan peubah dependen dan mengabaikan peubah independen sewaktu melakukan peramalan.

Metode ARIMA dibagi kedalam tiga kelompok model, yaitu *autoregressive model* (AR), *moving average model* (MA) dan model campuran yang memiliki karakteristik kedua model di atas yaitu *autoregressive integrated moving average* (ARIMA).

**1) Autoregressive Model (AR)**

AR adalah suatu model yang menjelaskan pergerakan suatu peubah melalui peubah itu sendiri di masa lalu.

Model *autoregressive* orde ke-p dapat ditulis sebagai berikut:

ARIMA (p,0,0)

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \epsilon_t \dots\dots\dots(1)$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \dots \theta_p$  = parameter autoregresive ke-p

$\epsilon_t$  = nilai kesalahan pada waktu ke t

**2) Moving Average Model (MA)**

MA adalah suatu model yang melihat pergerakan peubahnya melalui sisaannya di masa lalu.

Bentuk model MA dengan ordo q atau MA (q) atau model ARIMA (0,d,g) dapat ditulis sebagai berikut:

$$Y_t = \mu - \phi_1 \epsilon_{t-1} - \phi_2 \epsilon_{t-2} - \dots - \phi_q \epsilon_{t-q} + \epsilon_t \dots\dots\dots(2)$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$\phi_1 \dots \phi_q$  = parameter-parameter moving average

$\epsilon_{t-q}$  = nilai kesalahan pada waktu ke (t-q)

**3) Autoregressive Intergrated Moving Everage (ARIMA)**

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk modelnya adalah

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \epsilon_{t-1} - \phi_2 \epsilon_{t-2} - \dots - \phi_q \epsilon_{t-q} + \epsilon_t \dots\dots\dots(3)$$

dimana:



$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \theta_q \phi_1 \phi_n$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke-(t-q)

### **b. Metode Fungsi Transfer**

Dalam bidang pertanian, produksi komoditas pertanian sangat dipengaruhi oleh peubah lainnya seperti serangan OPT, penggunaan saprodi, gejolak harga komoditas tersebut atau komoditas lainnya, dan lainnya. Apabila peramalan produksi hanya didasarkan pada besarnya produksi komoditas yang bersangkutan saja tanpa memperhatikan faktor-faktor yang mempengaruhinya, maka informasi untuk pembuatan perencanaan menjadi tidak lengkap, sehingga tujuan peramalan menjadi tidak tercapai secara utuh. Salah satu upaya menganalisis data deret waktu multivariat agar diperoleh hasil yang dapat memberikan informasi yang lengkap dan simultan, adalah dengan menggunakan model Fungsi Transfer.

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai-nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke-t, tetapi juga pada waktu t+1, t+2, ..., t+k. Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara peubah *input* dan peubah *output*.

Tujuan pemodelan fungsi transfer adalah untuk menetapkan model yang sederhana, yang menghubungkan deret output ( $Y_i$ ) dengan deret input ( $X_i$ ) dan gangguan/noise ( $n_i$ ). Wei (1994) juga menjelaskan bahwa di dalam fungsi transfer terdapat rangkaian *output* yang mungkin dipengaruhi oleh rangkaian *multiple input*. Pada kasus *single input* peubah, dapat menggunakan metode korelasi silang yang dianjurkan oleh Box and Jenkins (1976). Teknik ini juga dapat digunakan ketika terdapat *single input* peubah yang lebih dari satu selama antar variable *input* tidak berkorelasi silang. Jika beberapa atau semua peubah input berkorelasi silang maka teknik prewhitening atau metode korelasi silang tidak dapat digunakan secara langsung. Alasan

utama bagi perlunya suatu perencanaan atau peramalan adalah adanya tenggang waktu pengambilan keputusan yang dapat berkisar dari beberapa hari atausampai beberapa tahun. Pada analisis fungsi transfer untuk peramalan deret berkala univariate, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (noise). Deret input mempengaruhi deret output melalui sebuah fungsi transfer yang mendistribusikan pengaruhnya secara dinamis melalui beberapa periode waktu yang akan datang dengan persentase tertentu yang disebut sebagai bobot respons impuls atau bobot fungsi transfer.

Model umum Fungsi Transfer:

$$y_t = v(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\varphi_p(B)} \varepsilon_t \quad \dots\dots\dots(4)$$

Dimana:

- $b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$
- $r \rightarrow$  panjang lag Y periode sebelumnya yang masih mempengaruhi  $Y_t$
- $s \rightarrow$  panjang jeda X periode sebelumnya yang masih mempengaruhi  $Y_t$
- $p \rightarrow$  ordo AR bagi noise  $N_t$
- $q \rightarrow$  ordo MA bagi noise  $N_t$

**c. Vector Autoregression (VAR)**

*Vector Autoregression* (VAR) dikemukakan pertama kali oleh Christopher Sims (1980). Sims mengembangkan model ekonometri dengan mengabaikan pengujian asumsi secara apriori. VAR dikembangkan oleh Sims sebagai kritik atas metode simultan. Jumlah peubah yang besar dan klasifikasi endogen dan eksogen pada metode simultan merupakan dasar dari kritik tersebut. Penggunaan pendekatan struktural atas pemodelan persamaan simultan biasanya menerapkan teori ekonomi di dalam usahanya untuk mendeskripsikan hubungan antar peubah yang ingin diuji. Akan tetapi sering ditemukan bahwa teori ekonomi saja ternyata tidak cukup kaya di dalam menyediakan spesifikasi yang ketat dan tepat atas hubungan dinamis antar peubah (Gujarati, 2010). Model VAR merupakan jalan keluar atas permasalahan ini melalui pendekatan non-struktural atau tidak mendasarkan pada teori ekonomi tertentu (*atheoretical*). Metode VAR memperlakukan seluruh peubah secara simetris tanpa memperlakukan peubah dependen dan independen atau dengan kata lain model ini memperlakukan seluruh peubah sebagai peubah

endogen., karena pada kenyataannya suatu peubah dapat bertindak sebagai endogen atau dependen dari suatu peubah yang lainnya (Gujarati, 2010).

Kelebihan dalam penggunaan metode VAR (Gujarati, 2010):

- a. Kemudahan dalam penggunaan, tidak perlu mengkhawatirkan tentang penentuan peubah endogen dan peubah eksogen.
- b. Kemudahan dalam estimasi, metode *Ordinary Least Square* (OLS) dapat diaplikasikan pada tiap persamaan secara terpisah.
- c. *Forecast* atau peramalan yang dihasilkan pada beberapa kasus ditemukan lebih baik daripada yang dihasilkan oleh model persamaan simultan yang kompleks.
- d. Perangkat estimasi yang digunakan adalah *Impulse Respon Function* (IRF) untuk melacak respon dari peubah dependen dalam sistem VAR terhadap *shock* dari *error term* dan *Variance Decomposition* yang memberikan informasi mengenai pentingnya masing-masing *error term* dalam mempengaruhi peubah-peubah dalam VAR.

Di sisi lain, terdapat beberapa kritik terhadap model VAR menyangkut permasalahan berikut (Gujarati, 2010) :

1. Model VAR merupakan model yang *atheoretic* atau tidak berdasarkan teori, hal ini tidak seperti pada persamaan simultan. Pada persamaan simultan, pemilihan peubah yang akan dimasukkan dalam persamaan memegang peranan penting dalam mengidentifikasi model.
2. Pada model VAR penekanannya terletak pada *forecasting* atau peramalan sehingga model ini kurang cocok digunakan dalam menganalisis kebijakan.
3. Permasalahan yang besar dalam model VAR adalah pada pemilihan *lag length* atau panjang lag yang tepat. Karena semakin panjang lag, maka akan menambah jumlah parameter yang akan bermasalah pada *degrees of freedom*.
4. Peubah yang tergabung pada model VAR harus stasioner. Apabila tidak stasioner, perlu dilakukan transformasi bentuk data, misalnya melalui *first difference*.
5. Sering ditemui kesulitan dalam menginterpretasi tiap koefisien pada estimasi model VAR, sehingga sebagian besar peneliti melakukan interpretasi pada estimasi fungsi *impulse respon*.

### Estimasi Model VAR

Enders (2004) menjelaskan, ketika peneliti tidak memiliki kepastian untuk menentukan bahwa suatu peubah adalah eksogen, maka suatu perluasan analisis fungsi perpindahan alami akan memperlakukan masing-masing peubah secara simetris. Sebagai contoh, pada kasus-kasus peubah yang membiarkan alur waktu atau *time path*  $\{s_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{y_t\}$  dan membiarkan *time path*  $\{y_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{s_t\}$ .

Di dalam sistem *bivariate*, hubungan tersebut dapat digambarkan seperti pada persamaan (1.4) di bawah ini:

$$\begin{aligned} s_t &= b_{10} - b_{12}y_t + \gamma_{11}s_{t-1} + \gamma_{12}y_{t-1} + \varepsilon_{s_t} \\ y_t &= b_{20} - b_{21}s_t + \gamma_{21}s_{t-1} + \gamma_{22}y_{t-1} + \varepsilon_{y_t} \end{aligned} \quad \text{.....(5)}$$

Dengan mengasumsikan bahwa kedua peubah  $s_t$  dan  $y_t$  adalah stasioner:  $\varepsilon_{s_t}$  dan  $\varepsilon_{y_t}$  adalah *disturbances* yang memiliki rata-rata nol dan matriks kovarians terbatas atau bersifat *white noise* dengan standar deviasi yang berurutan  $\sigma_s$  dan  $\sigma_y$ ; serta  $\{\varepsilon_{s_t}\}$  dan  $\{\varepsilon_{y_t}\}$  adalah *disturbances* yang independen dengan rata-rata nol dan kovarian terbatas (*uncorrelated white-noise disturbances*). Kedua persamaan di atas merupakan orde pertama VAR, karena panjang *lag* nya hanya satu. Agar persamaan (4.1) lebih mudah dipahami dan digunakan sebagai alat analisis maka ditransformasikan dengan menggunakan matriks aljabar, dan hasilnya dapat dituliskan secara bersama seperti pada persamaan di bawah ini:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Atau dengan bentuk lain:

$$Bx_t = \Gamma_0 + \Gamma_1 X_{t-1} + \varepsilon_t \quad \text{.....(6)}$$

Dimana:

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \quad x_t = \begin{bmatrix} s_t \\ y_t \end{bmatrix} \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Dengan melakukan pengalihan antara persamaan (4.2) dengan  $B^{-1}$  atau invers matriks B, maka akan dapat ditentukan model VAR dalam bentuk standar, seperti dituliskan pada persamaan di bawah ini:

$$x_t = A_0 + A_1 x_{t-1} + \ell_t \dots \dots \dots (7)$$

$$A_0 = B^{-1} \Gamma_0$$

dimana  $A_1 = B^{-1} \Gamma_1$

$$\ell_t = B^{-1} \varepsilon_t$$

Untuk tujuan notasi, maka  $\{a_{i0}\}$  dapat didefinisikan sebagai elemen ke-i dari vektor  $A_0$ ;  $\{a_{ij}\}$  sebagai elemen dalam baris ke-i dan baris ke-j dari matriks  $A_1$ ; dan  $\{e_{it}\}$  sebagai elemen ke-i dari vektor  $e_t$ . Dengan menggunakan notasi baru yang telah dijelaskan sebelumnya, maka persamaan (4.3) dapat ditulis menjadi:

$$s_t = a_{10} + a_{11}s_{t-1} + a_{12}y_{t-1} + e_{1t}$$

$$y_t = a_{20} + a_{21}s_{t-1} + a_{22}y_{t-1} + e_{2t} \dots \dots \dots (8)$$

**Fungsi Impulse Response**

Fungsi *impulse response* menggambarkan tingkat laju dari *shock* peubah yang satu terhadap peubah yang lainnya pada suatu rentang periode tertentu. Sehingga dapat dilihat lamanya pengaruh dari *shock* suatu peubah terhadap peubah lain sampai pengaruhnya hilang atau kembali ke titik keseimbangan. Analisis fungsi *impulse respon* dapat dituliskan dalam bentuk *Vector Moving Avarage (VMA)* dari bentuk standar VAR pada persamaan (4.4).

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} - \\ s \\ - \\ y \end{bmatrix} + \sum_{i=1}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} e_{1t-i} \\ e_{2t-i} \end{bmatrix} \dots \dots \dots (9)$$

dimana  $s_t$  dan  $y_t$  memiliki hubungan dengan  $e_{1t}$  dan  $e_{2t}$  secara berurutan. Selanjutnya dengan menggunakan operasi aljabar matriks maka *vector error* dapat ditentukan sebagai berikut :

$$\begin{bmatrix} e_{1t} \\ e_{2t} \end{bmatrix} = \frac{1}{1 - b_{12}b_{21}} \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{st-i} \\ \varepsilon_{yt-i} \end{bmatrix} \dots \dots \dots (10)$$

dengan menggabungkan persamaan (4.5) dan (4.6) akan didapat:

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s}_t \\ \bar{y}_t \end{bmatrix} + \frac{1}{1 - b_{12}b_{21}} \sum_{i=0}^{\infty} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}^i \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \begin{bmatrix} \varepsilon_{st-i} \\ \varepsilon_{yt-i} \end{bmatrix} \dots\dots\dots(11)$$

Persamaan (4.7) dapat disederhanakan dengan mendefinisikan matriks 2x2  $\Phi_i$  dengan elemen  $\Phi_{jk}$  (i) seperti persamaan berikut :

$$\Phi_i = A_1^i / (1 - b_{12}b_{21}) \begin{bmatrix} 1 & -b_{12} \\ -b_{21} & 1 \end{bmatrix} \dots\dots\dots(12)$$

sehingga diperoleh bentuk matriks persamaan fungsi *impulse respon*:

$$\begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} \bar{s} \\ \bar{y} \end{bmatrix} + \sum_{i=0}^n \begin{bmatrix} \phi_{11}(i) & \phi_{12}(i) \\ \phi_{21}(i) & \phi_{22}(i) \end{bmatrix} \begin{bmatrix} \varepsilon_{st-1} \\ \varepsilon_{yt-1} \end{bmatrix} \dots\dots\dots(13)$$

dimana :

$\Phi_{ij}(i)$  = efek dari *structural shock* pada s dan y

$\Phi_{ij}(0)$  = *impact multipliers*

$\sum \Phi_{ij}(i)$  = *cumulative multipliers*

$\sum \Phi_{ij}(i)$  pada saat  $n \rightarrow \infty$  = *long run multipliers*

### **Variance Decomposition**

*Variance decomposition* atau disebut juga *forecast error variance decomposition* merupakan perangkat pada model VAR yang akan memisahkan variasi dari sejumlah peubah yang diestimasi menjadi komponen-komponen *shock* atau menjadi peubah *innovation*, dengan asumsi bahwa peubah-peubah *innovation* tidak saling berkorelasi. Kemudian, *variance decomposition* akan memberikan informasi mengenai proporsi dari pergerakan pengaruh *shock* pada sebuah peubah terhadap *shock* peubah yang lain pada periode saat ini dan periode yang akan datang.

Bentuk VMA dari peubah x pada satu periode ke depan dapat dituliskan sbb.:

$$x_{t+1} = \bar{x} + \sum_{i=0}^{\infty} \phi_i \varepsilon_{t+1-i} \dots\dots\dots(14)$$

*Forecast error* pada satu periode kedepan adalah:

$$E_t x_{t+1} = \bar{x} + \sum_{i=1}^{\infty} \phi_i \varepsilon_{t+1-i} \dots\dots\dots(15)$$

*Forecast* satu periode ke depan dilambangkan dengan  $\Phi_0 \varepsilon_{t+1}$ . *Forecast error* pada periode n ke depan adalah:

$$x_{t+n} - E_t x_{t+n} = \bar{x} + \sum_{i=0}^{n-1} \phi_i \varepsilon_{t+n-i} \dots\dots\dots(16)$$

Forecast error pada n periode ke depan untuk peubah s adalah:

$$s_{t+n} - E y_{t+n} = \phi_{11}(0)\varepsilon_{st+n} + \phi_{11}(1)\varepsilon_{st+n-1} + \dots + \phi_{11}(n-1)\varepsilon_{yt+1} \\ + \phi_{12}(0)\varepsilon_{yt+n} + \phi_{12}(1)\varepsilon_{yt+n-1} + \dots + \phi_{12}(n-1)\varepsilon_{yt+1} \dots\dots\dots(17)$$

Variance dari forecast error  $s_{t+n}$  periode n ke depan adalah  $\sigma_s(n)^2$  dimana:

$$\sigma_s(n)^2 = \sigma_s^2 [\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2] + \sigma_y^2 [\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2] \dots\dots\dots(18)$$

Forecast error variance decomposition adalah proporsi dari  $\sigma_s(n)^2$  terhadap shock s dan shock y.

Sehingga forecast error variance decomposition pada shock s adalah:

$$\sigma_s^2 [\phi_{11}(0)^2 + \phi_{11}(1)^2 + \dots + \phi_{11}(n-1)^2] / \sigma_s(n)^2 \dots\dots\dots(19)$$

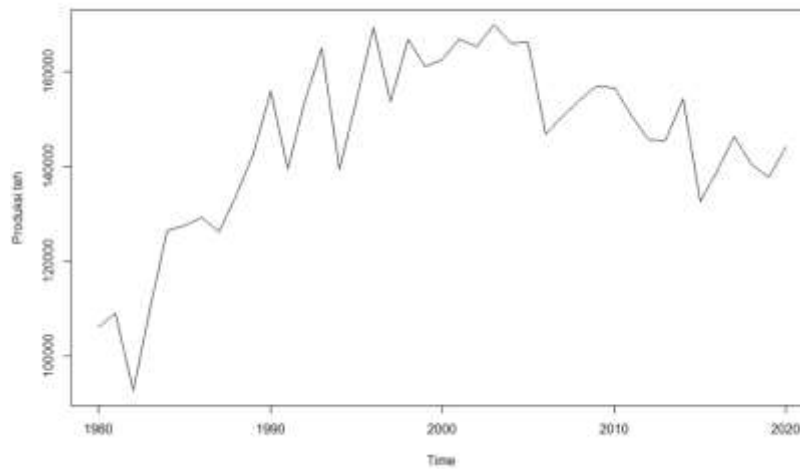
sedangkan forecast error variance decomposition pada shock y adalah:

$$\sigma_y^2 [\phi_{12}(0)^2 + \phi_{12}(1)^2 + \dots + \phi_{12}(n-1)^2] / \sigma_y(n)^2 \dots\dots\dots(20)$$

## HASIL DAN PEMBAHASAN

### A. Model ARIMA

ARIMA sering disebut juga dengan metode Box-Jenkins, maka baik digunakan untuk peramalan jangka pendek sedangkan untuk peramalan jangka panjang untuk ketepatan peramalannya masih kurang baik, biasanya akan cenderung *flat* (mendatar/konstan). Eksplorasi data produksi teh Indonesia dimulai dari tahun 1980 hingga tahun 2020 dalam wujud produksi daun kering. Gambar 1 menunjukkan perkembangan produksi teh tahun 1980 hingga 2020, dimana perkembangan produksi teh di Indonesia berfluktuasi namun cenderung meningkat dengan rata-rata pertumbuhan sebesar 0,77% per tahun. Pada tahun 1980 total produksi teh di Indonesia sebesar 106,17 ribu ton dan mengalami peningkatan di tahun 2020 menjadi 128,01 ribu ton.



Gambar 1. Perkembangan Produksi Teh Tahun 1980-2020

Pada pemodelan *Autoregressive Integrated Average* (ARIMA), data yang digunakan dapat dilakukan pemisahan antara data testing dan data training, pemisahan data ini berfungsi untuk menguji tingkat akurasi dalam melakukan peramalan. Panjang series data produksi teh pada data set training adalah tahun 1980 sampai tahun 2014, sementara dataset testing adalah periode tahun 2015 sampai tahun 2020. Dataset training digunakan untuk melakukan penyusunan model, sementara dataset testing digunakan untuk validasi model.

Uji kestasioneran data pada pemodelan ARIMA dapat dilihat dari hasil plot data dan uji formal statistik. Dari hasil uji formal statistik dapat dilihat dengan uji Augmented Dickey-Fuller, yang menunjukkan bahwa data produksi teh belum stasioner. Hal ini terlihat dari hasil *value of test*



*statistic* -1.7231, nilai kritis pada tingkat kepercayaan 5% sebesar -3,50 artinya nilai ini lebih kecil dari nilai uji statistik maka  $H_0$  tidak ditolak. Untuk mendapat data produksi teh yang stasioner maka dilakukan differencing satu kali, seperti gambar 2.

Tabel 1. Hasil Uji Augmented Dickey Fuller

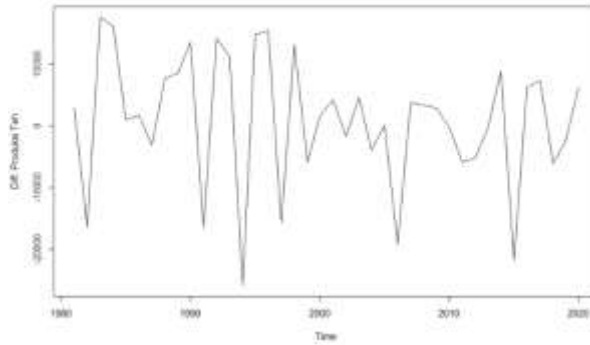
```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####

Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)

Residuals:
    Min     1Q   Median     3Q    Max
-25230.0 -4117.5  946.8  6153.2 19668.3

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 29813.0496 14168.9450  2.104  0.0426 *
z.lag.1      -0.1823    0.1058 -1.723  0.0937 .
tt           -94.2041   165.7896 -0.568  0.5735
z.diff.lag   -0.3034    0.1595 -1.902  0.0654 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9956 on 35 degrees of freedom
Multiple R-squared:  0.2397,    Adjusted R-squared:  0.1745
F-statistic: 3.678 on 3 and 35 DF, p-value: 0.02106
Value of test-statistic is: -1.7231 2.0457 2.7984
Critical values for test statistics:
    1pct 5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```

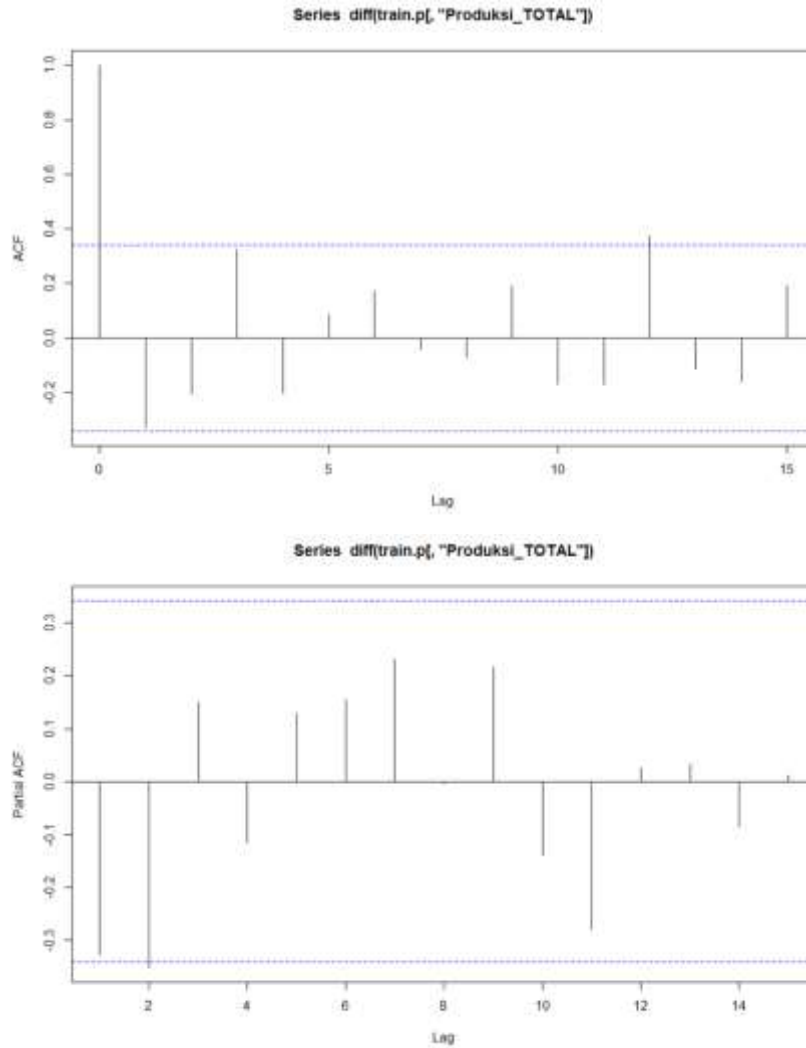


Gambar 2. Perkembangan Produksi Teh Tahun 1980-2020, Setelah Differencing Satu Kali

Dari hasil plot produksi teh yang telah dilakukan differencing satu kali, maka plot tersebut sudah menunjukkan stasioner. Selain itu didukung juga dari hasil uji augmented Dickey Fuller yang menunjukkan Value of test-statistic is: -11.7548 dengan nilai kritis pada tingkat kepercayaan 95% sebesar -1,95 dan tingkat kepercayaan 99% sebesar -2,62 artinya lebih besar dari nilai uji statistik sehingga data produksi teh setelah differencing satu kali sudah stasioner. Stasioneritas data tersebut juga didukung dengan plot ACF dan PACF, dimana pada plot tersebut tidak banyak lag yang keluar dari *confidence interval* (Gambar 3)

Tabel 2. Hasil Uji Augmented Dickey Fuller Setelah Differencing Satu Kali

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
  Min   1Q Median   3Q   Max
-25081 -6861  2875  7132 15550
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
z.lag.1  -2.5075   0.2133 -11.755 1.05e-13 ***
z.diff.lag  0.6415   0.1201   5.344 5.66e-06 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 11010 on 35 degrees of freedom
Multiple R-squared:  0.8766, Adjusted R-squared:  0.8696
F-statistic: 124.3 on 2 and 35 DF, p-value: < 2.2e-16
Value of test-statistic is: -11.7548
Critical values for test statistics:
  1pct 5pct 10pct
tau1 -2.62 -1.95 -1.61
```



Gambar 3. Plot ACF dan PACF Produksi Teh Differencing Satu Kali

Dari hasil plot ACF dan PACF sulit untuk ditentukan orde ARIMA, maka run model dengan menggunakan auto arima disarankan menggunakan ARIMA (0,1,1). Dengan menggunakan ARIMA (0,1,1) menghasilkan MAPE sebesar 5,49% artinya model arima akan menyimpang sekitar -5,49% sampai +4,59% dari data aktual.

Tabel 3. ARIMA dengan Model Auto Arima

Series: train.p[, "Produksi_TOTAL"]
ARIMA(0,1,1)
Coefficients:
ma1
-0.3751
s.e. 0.1419

sigma<sup>2</sup> estimated as 104899113: log likelihood=-351.12  
 AIC=706.25 AICc=706.65 BIC=709.24

Training set error measures:

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	1897.353	9936.225	7702.993	1.14934	5.490948	0.9136643	-0.05934335

Hasil model Auto Arima (0,1,1) menghasilkan nilai estimasi yang flat, maka untuk mendapatkan orde ARIMA terbaik dapat menggunakan *Arima selection*. Pada metode ini akan dikeluarkan beberapa model terbaik. Model terbaik pertama menurut metode *Arima selection* adalah ARIMA (3,1,5) menghasilkan nilai sbc paling kecil sebesar 562,60. Model ARIMA yang digunakan untuk produksi teh Indonesia menggunakan ARIMA (1,1,1) pada urutan ke sembilan dengan nilai sbc sebesar 585,19.

Tabel 4. Arima Selection

	p	q	sbc
[1,]	3	5	562.6014
[2,]	4	5	564.1076
[3,]	5	5	568.3318
[4,]	1	5	576.2201
[5,]	2	5	578.6543
[6,]	0	5	579.1540
[7,]	1	3	582.1004
[8,]	0	3	582.2178
[9,]	1	1	585.1958
[10,]	3	3	586.3579

Untuk menentukan model ARIMA terbaik dari ketiga model tentative tersebut dilakukan pengujian berdasarkan data training dan data testing. Model terbaik adalah model yang menghasilkan MAPE terkecil. Pada analisis ini sudah dilakukan pembagian data training dan data testing. Data training merupakan data produksi teh dari tahun 1980 – 2020, sementara data testing diambil 6 data terakhir, yaitu produksi teh tahun 2015 – 2020. Untuk menguji performa model ARIMA terbaik, dilakukan pengujian dengan data testing. Hasil pengujian data training dan testing terlihat pada tabel 5.

Selanjutnya dilakukan pengujian model ARIMA (1,1,1) apakah koefisien sudah signifikan. Untuk model ARIMA (1,1,1) koefisien ma1 sebesar 0,1680 dan koefisien ini signifikan pada taraf 1%. Sehingga model ARIMA (1,1,1) layak digunakan.

Tabel 5. Uji Koefisien Model ARIMA (1,1,1)

```

Call:
arima(x = train.p[, "Produksi_TOTAL"], order = c(1, 1, 1))
Coefficients:
      ar1    ma1
-0.0245 -0.3580
s.e. 0.3076 0.2597

sigma^2 estimated as 101689286: log likelihood = -351.12, aic = 708.24

z test of coefficients:

      Estimate Std. Error z value Pr(>|z|)
ar1 -0.024529  0.307644 -0.0797  0.9365
ma1 -0.358050  0.259718 -1.3786  0.1680

```

Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan Ljung-Box. Hasil pengujian nilai p-value lag 5 sampai dengan lag 30 tidak ada yang signifikan, sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autokorelasi. Hal ini menunjukkan bahwa ARIMA (1,1,1) layak digunakan.

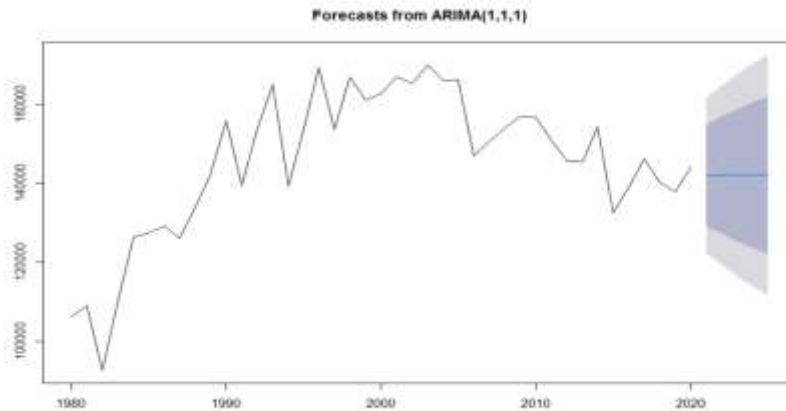
Tabel 6. Uji Sisaan Ljung Box ARIMA (1,1,1)

lags	statistic	df	p-value
5	4.353503	5	0.4997201
10	9.227073	10	0.5106978
15	17.626744	15	0.2827967
20	20.640790	20	0.4185397
25	28.146128	25	0.3011702
30	30.936415	30	0.4185380

Hasil estimasi produksi teh Indonesia dengan menggunakan model ARIMA (1,1,1) menghasilkan angka estimasi untuk 5 tahun ke depan. Hasil estimasi produksi teh tahun 2021 sebesar 141,97 ribu ton. Hasil estimasi produksi teh tahun 2022 sebesar 142,08 ribu ton.

Tabel 7. Hasil Estimasi Produksi Teh ARIMA (1,1,1)

Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
2021	141972.6	129065.4	154879.9	122232.7
2022	142080.1	127228.2	156931.9	119366.2
2023	142074.5	125376.2	158772.8	116536.7
2024	142074.8	123720.9	160428.7	114004.9
2025	142074.8	122202.5	161947.2	111682.7



Gambar 4. Hasil Estimasi Produksi Teh Tahun 2021-2025

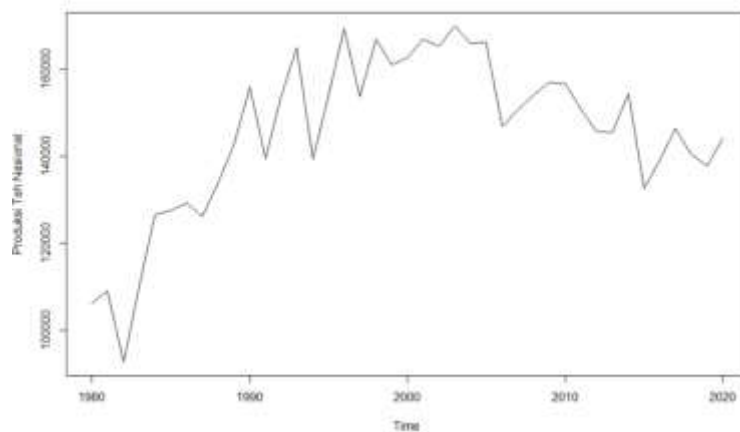
## B. Model Fungsi Transfer

Dalam melakukan estimasi produksi teh dengan menggunakan model fungsi transfer, peubah inputnya adalah luas tanaman menghasilkan teh. Langkah awal dalam mencari model fungsi transfer adalah penyiapan data, yang berupa eksplorasi data dengan menampilkan plot data produksi teh maupun luas tanaman menghasilkan teh. Data yang digunakan bersifat data series yang dimulai dari tahun 1980 hingga tahun 2020. Data produksi dan luas tanaman menghasilkan teh berfluktuasi namun cenderung menurun.

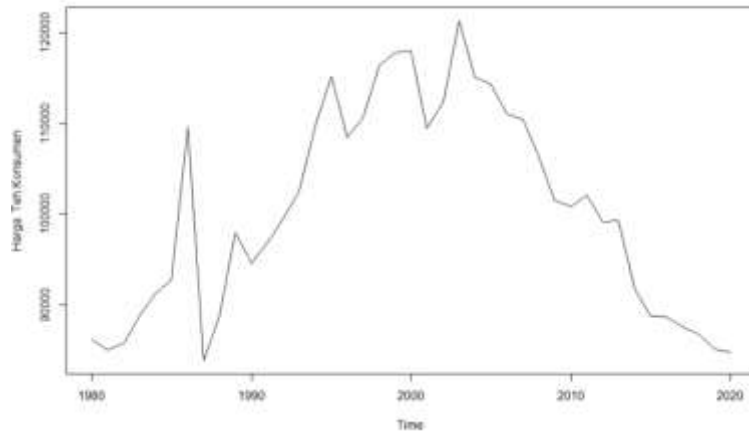
Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (*Autoregressive Integrated Moving Average*). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variabel lain sebagaimana pada model regresi. Model fungsi transfer merupakan suatu model yang menggambarkan nilai prediksi dari suatu time series (deret output atau  $Y_t$ ) berdasarkan nilai-nilai dari deret itu sendiri ( $Y_t$ ) dan berdasarkan pula pada

data time series yang mempunyai hubungan (deret input atau  $X_t$ ) dengan deret output. Model fungsi transfer yang bersifat dinamis berpengaruh tidak hanya pada hubungan linier antara waktu ke- $t$  input  $X_t$  dan waktu ke- $t$  output  $Y_t$ , tetapi juga saat input  $X_t$  dengan saat  $t, t+1, \dots, t+k$  pada output  $Y_t$  (Fitriani et al., 2013). Dalam melakukan estimasi dengan fungsi transfer diperlukan dua variabel, yaitu output  $Y_t$  dan input  $X_t$ . Input  $X_t$  merupakan variabel yang diduga paling berpengaruh terhadap output  $Y_t$ . Pada analisis ini dicontohkan estimasi untuk produksi teh (output  $Y_t$ ) yang diduga dipengaruhi oleh luas tanaman menghasilkan teh.

Eksplorasi data dilakukan dengan menampilkan plot data produksi maupun luas tanaman menghasilkan. Berdasarkan plot data dapat diketahui pola data series 41 titik yang akan digunakan untuk pemodelan yaitu dimulai dari tahun 1980 hingga tahun 2020. Berdasarkan Gambar 5 dan Gambar 6, terlihat bahwa terdapat data produksi teh nasional memiliki tren meningkat dari tahun ke tahun sedangkan luas tanaman menghasilkan cenderung berfluktuasi. Luas tanaman menghasilkan cenderung turun pada sekitar 10 tahun terakhir, akibat persaingan harga dengan produk teh luar negeri. Produksi teh nasional maupun luas tanaman menghasilkan terindikasi tidak stasioner berdasarkan plotnya.



Gambar 5. Produksi Teh Tahun 1980-2020



Gambar 6. Luas Tanaman Menghasilkan Teh Tahun 1980-2020

Dari hasil uji Augmented Dickey Fuller yang telah dilakukan differencing satu kali menghasilkan nilai *test-statistic* yaitu -6.0014 lebih kecil dari *critical values* artinya data luas tanaman menghasilkan sudah stasioner.

Tabel 8. Hasil Uji Augmented Dickey Fuller Luas Tanaman Menghasilkan Setelah Differencing

```
# Augmented Dickey-Fuller Test Unit Root Test #
Test regression none
Call:
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)
Residuals:
  Min    1Q  Median    3Q   Max
-18198.6 -2196.7  931.9  3834.8 18325.1
Coefficients:
      Estimate Std. Error t value Pr(>|t|)
z.lag.1  -1.7284   0.2880  -6.001 1.59e-06 ***
z.diff.lag  0.3070   0.1772   1.732  0.0938 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6573 on 29 degrees of freedom
Multiple R-squared:  0.693,    Adjusted R-squared:  0.6718
F-statistic: 32.73 on 2 and 29 DF, p-value: 3.658e-08

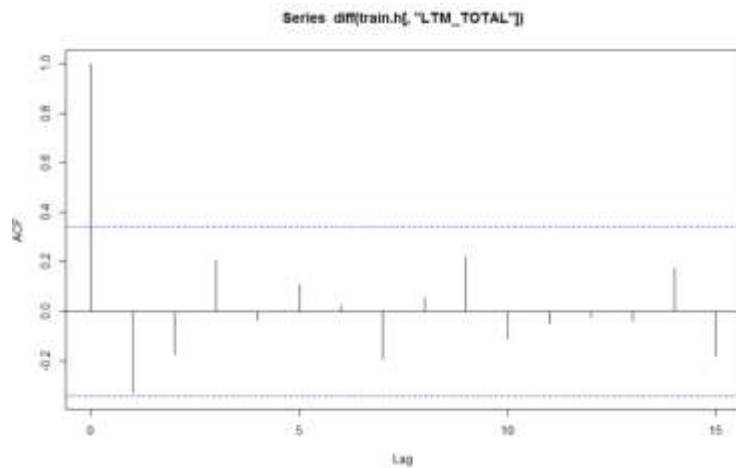
Value of test-statistic is: -6.0014

Critical values for test statistics:
      1pct 5pct 10pct
tau1 -2.62 -1.95 -1.61
```

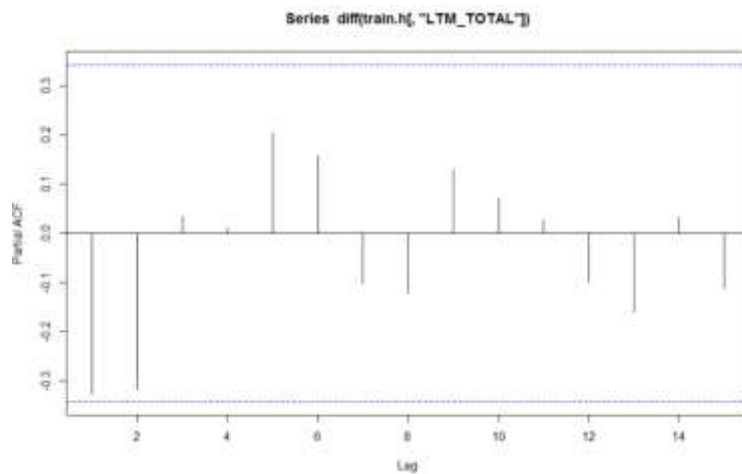


Stasioneritas data dapat juga dilihat dari plot ACF dan PACF. Jika lag pada ACF maupun PACF menunjukkan *tail off* maka data tidak stasioner.

Stasioneritas data luas tanaman menghasilkan teh juga didukung dengan plot ACF dan PACF, dimana pada plot tersebut tidak banyak lag yang keluar dari *confidence interval*.



Gambar 7. Plot ACF Luas Tanaman Menghasilkan Teh Setelah Differencing



Gambar 8. Plot PACF Luas Tanaman Menghasilkan Teh Setelah Differencing

Penelusuran model ARIMA dengan model terbaik dapat menggunakan *script auto arima*, dengan data *training*. Dari hasil auto arima luas tanaman menghasilkan didapat ARIMA (0,1,1) dengan MAPE sebesar 4,30% dengan nilai AIC sebesar 675,33, disajikan pada Tabel 9.

Tabel 9. Model Auto Arima Luas Tanaman Menghasilkan Teh

Series: train.h[, "LTM_TOTAL"]							
ARIMA(0,1,1)							
Coefficients:							
ma1							
-0.4039							
s.e. 0.1417							
sigma^2 estimated as 41069250: log likelihood=-335.66							
AIC=675.33 AICc=675.73 BIC=678.32							
Training set error measures:							
	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	675.788	6217.187	4425.167	0.4604833	4.300691	0.9655485	-0.02837655

Hasil pengolahan luas tanaman menghasilkan teh pada RStudio yang telah dilakukan differencing dua kali akan menampilkan sepuluh model tentatif dengan model arima selection. Model pertama dengan nilai p=5 dan q=1 dan model kesepuluh adalah p=3 dan q=1. Dari kesepuluh model ini dipilih model keenam yaitu ARIMA (1,2,1).

Tabel 10. Arima Selection Data Luas Tanaman Menghasilkan Teh

	p	q	sbc
[1,]	5	1	552.7390
[2,]	5	5	554.6447
[3,]	0	1	556.0142
[4,]	5	2	558.3862
[5,]	2	1	558.4723
[6,]	1	1	559.4557
[7,]	0	3	560.4417
[8,]	0	2	560.6025
[9,]	5	3	560.7322
[10,]	3	1	562.0332

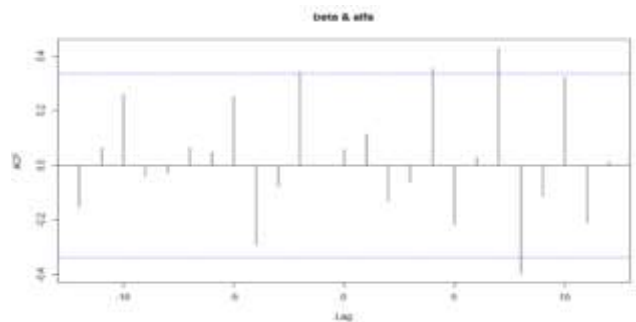
Hasil pengujian koefisien model ARIMA (1,2,1) menghasilkan nilai AIC sebesar 663,24 ditemukan bahwa ar1 tidak signifikan sedangkan ma1 signifikan pada taraf 1%.

Tabel 11. Pengujian Koefisien Model Arima (1,2,1) Luas Tanaman Menghasilkan Teh

```
Call:
arima(x = train.h[, "LTM_TOTAL"], order = c(1, 2, 1))
Coefficients:
    ar1    ma1
 -0.3144 -0.9373
s.e. 0.1708 0.1115

sigma^2 estimated as 44732058: log likelihood = -328.62, aic = 663.24
z test of coefficients:
  Estimate Std. Error z value Pr(>|z|)
ar1 -0.31441   0.17082 -1.8407 0.06567 .
ma1 -0.93734   0.11148 -8.4083 < 2e-16 ***
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Tahap selanjutnya dalam penyusunan model fungsi transfer adalah *prewhitening* dan korelasi silang. Korelasi silang menggambarkan struktur hubungan antara  $X_t$  dengan  $Y_t$ . Untuk mengidentifikasi pengaruh  $X_t$  terhadap  $Y_t$  maka deret  $X_t$  harus stasioner atau sudah distasionerkan. Dalam konteks pemodelan  $X_t$  terhadap  $Y_t$ , untuk membuat  $X_t$  stasioner tidak dengan pembedaan (*differencing*) namun dengan mengambil komponen *white noise* dari  $X_t$  (*prewhitening*). *Prewhitening* dilakukan terhadap deret input  $X_t$  yang didefinisikan sebagai alfa serta deret input  $Y_t$  yang didefinisikan sebagai beta. Hasil output untuk *prewhitening* dan korelasi silang berupa grafik ACF untuk beta dan alfa.



Gambar 9. Plot korelasi silang Produksi Teh dengan Luas Tanaman Menghasilkan Teh

Hasil plot korelasi silang digunakan untuk mengidentifikasi ordo  $r, s$ , dan  $b$ . Ordo  $r$  adalah panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$ , ordo  $s$  adalah panjang lag  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$ , dan ordo  $b$  adalah panjang jeda pengaruh  $X_t$  terhadap  $Y_t$ . Identifikasi ordo  $r, s$  dan  $b$  hanya dilihat pada lag yang positif.

Plot korelasi silang diatas menunjukkan bahwa hanya lag 0 yang keluar dari garis signifikansi, maka nilai  $b=0$ . Kemudian, tidak ada tambahan lagi lag yang signifikan maka nilai  $s=0$ . Mengingat data produksi teh dan luas tanaman menghasilkan merupakan data tahunan yang tidak mengandung musiman maka diasumsikan nilai  $r=0$ . Nilai  $b=0$  menunjukkan tidak ada jeda pengaruh antara luas tanaman menghasilkan pada waktu  $t$  terhadap produksi teh pada waktu  $t$ . Nilai  $s=0$  berarti ada korelasi antara produksi dan luas tanaman menghasilkan pada tahun yang sama. Dengan kata lain, dampak dari luas tanaman menghasilkan terhadap produksi dirasakan pada waktu yang sama ( $t$ ).

Tahap selanjutnya dilakukan pengepasan model, untuk nilai  $r, s$  dan  $b$ . Hasil pengujian fungsi transfer dengan nilai  $r=0, s=0$ , dan  $b=0$  menghasilkan nilai MAPE sebesar 7,05%

Tabel 12. Pengujian Orde  $r, s$ , dan  $b$  ARIMA (0,0,0) Produksi Teh

Series: train.h[, "Produksi_TOTAL"]						
Regression with ARIMA(0,0,0) errors						
Coefficients:						
intercept	xreg					
-5813.373	1.4736					
s.e.	19593.053	0.1891				
sigma^2 estimated as 149720164: log likelihood=-367.23						
AIC=740.45 AICc=741.25 BIC=745.03						
Training set error measures:						
	ME	RMSE	MAE	MPE	MAPE	MASE ACF1
Training set	3.381152e-11	11870.68	9612.217	-0.834711	7.050934	1.14012 0.2729892

Untuk menghasilkan orde yang paling tepat untuk menentukan orde Arima fungsi transfer dengan melakukan identifikasi model noise. Untuk menghasilkan model terbaik dengan menggunakan *auto-arima* pada R Studio, model maka noise yang disarankan adalah Arima (1,2,1). Model ini ternyata masih kurang tepat, karena menghasilkan MAPE yang cukup besar yaitu 100,00%

Tabel 13. Model Residual Tentatif Produksi Teh

```

Series: res
ARIMA(0,0,0) with zero mean

sigma^2 estimated as 140913096: log likelihood=-367.23
AIC=736.45 AICc=736.58 BIC=737.98

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 3.381152e-11 11870.68 9612.217 100    100    0.8909865 0.2729892

```

Tahap berikutnya adalah dilakukan pengepasan model dengan noise. Pengepasan model dengan noise dilakukan dengan menggabungkan nilai r,s,b dan hasil identifikasi model noise yaitu ARIMA(1,2,1). Selanjutnya diperoleh output model fungsi transfer ARIMA (1,2,1) (Tabel 14)

Tabel 14. Model Residual Tentatif ARIMA (1,2,1) Untuk Produksi Teh

```

Series: res
ARIMA(1,2,1)

Coefficients:
      ar1      ma1
      -0.3411 -1.0000
s.e.  0.1627  0.0844

sigma^2 estimated as 1.93e+08: log likelihood=-351.72
AIC=709.43 AICc=710.29 BIC=713.83
> coeftest(modelres)

z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ar1 -0.341129  0.162745 -2.0961 0.03607 *
ma1 -0.999997  0.084381 -11.8510 < 2e-16 ***
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

```

### Peramalan Berbasis Fungsi Transfer

Model fungsi transfer dengan model noise ARIMA (1, 2, 1), dilakukan peramalan berbasis nilai aktual dimana produksi teh diestimasi menggunakan data aktual luas tanaman menghasilkan periode 2015-2020. Meskipun data aktual produksi teh periode 2015-2020 telah ada, dilakukan peramalan produksi teh untuk mengecek *performance* model fungsi transfer.

Tabel 15. Pengujian Model Fungsi Transfer ARIMA (1,2,1)

```

Series: test.h[, "Produksi_TOTAL"]
Regression with ARIMA(1,2,1) errors

Coefficients:
    ar1    ma1  xreg
-0.3233 -0.9293 0.109
s.e.  0.0000 0.0000 0.000

sigma^2 estimated as 121818371: log likelihood=-44.02
AIC=90.04  AICc=92.04  BIC=89.43

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 4923.713 9376.369 5881.689 3.44786 4.140412 0.6708437 0.167878
    
```

Tujuan melakukan pemodelan fungsi transfer adalah untuk mendapatkan nilai ramalan periode ke depan, yakni produksi teh tahun 2021-2025.

Hasil estimasi produksi teh berbasis fungsi transfer dengan model noise ARIMA (1,2,1) menghasilkan MAPE sebesar 4,14%.

Tabel 16. Model Fungsi Transfer ARIMA (1,2,1)

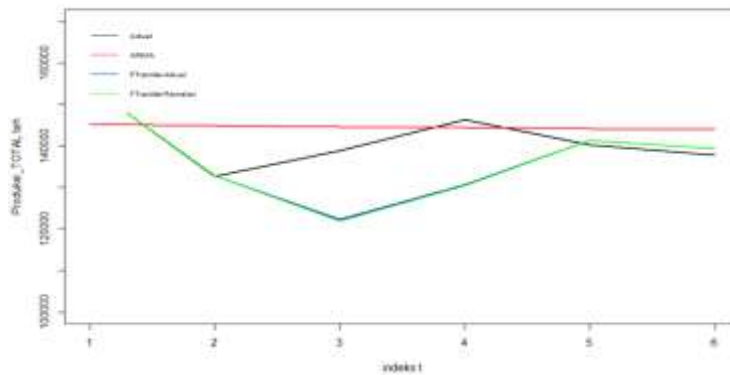
```

Series: train.h[, "Produksi_TOTAL"]
ARIMA(1,2,1)
Coefficients:
    ar1    ma1
-0.3240 -0.9205
s.e.  0.1685 0.0898

sigma^2 estimated as 118702187: log likelihood=-343.1
AIC=692.2  AICc=693.06  BIC=696.6

Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set -1262.16 10234.12 7478.167 -0.942272 5.324779 0.8869973 -0.1434023
    
```

Dilakukan perbandingan hasil estimasi terhadap data aktual produksi teh maupun hasil ramalan. Hasil ramalan tersebut ditampilkan dalam bentuk grafik pada Gambar 10.



Gambar 10. Perbandingan Hasil Ramalan Produksi Teh

Dari grafik di atas terlihat jika dibandingkan dengan data aktual produksi teh 2015-2020 (warna hitam), maka hasil peramalan dengan model ARIMA tanpa fungsi transfer (warna merah) tidak mengikuti pola data aktualnya. Di sisi lain, peramalan dengan fungsi transfer khususnya jika input luas tanaman menghasilkan yang digunakan adalah data aktual maka hasil ramalan produksinya (warna biru) sangat menyerupai pola data produksi aktual selama 6 tahun terakhir. Jika input luas tanaman menghasilkan teh yang digunakan adalah hasil ramalan, maka estimasi produksinya (warna hijau) hampir menyerupai pola data asli, meskipun tidak terlalu berimpit seperti estimasi menggunakan data actual.

Berdasarkan nilai MAPE yang dihasilkan maka model fungsi transfer ARIMA dengan input data aktual lebih bagus jika dibandingkan dengan model fungsi transfer ARIMA dengan input data ramalan maupun model ARIMA tanpa fungsi transfer. Namun demikian, karena tujuan penyusunan model adalah untuk mendapatkan angka estimasi tahun 2021 -2025, maka digunakan model Fungsi Transfer berbasis input data ramalan, yakni dengan terlebih dahulu dilakukan estimasi luas tanaman menghasilkan lima tahun ke depan menggunakan model ARIMA (1,2,1). Hasil estimasi luas tanaman menghasilkan tersebut kemudian menjadi input untuk meramalkan produksi teh lima tahun ke depan menggunakan model fungsi transfer.

Tabel 17. Hasil Peramalan Model Fungsi Transfer Arima

Time Series: Start = 41 End = 45 Frequency = 1 [1] 138357.9 137886.1 137737.6 137477.6 137256.2
---

Rata-rata pertumbuhan untuk 5 tahun kedepan yaitu periode 2021 hingga 2025 akan meningkat sebesar 0,58%. Tahun 2021 produksi teh diestimasi sebesar 138.696 ton. Produksi teh tahun 2022 sebesar 139.508 ton. Produksi teh diramalkan terus meningkat hingga tahun 2025. Tahun 2023 hasil estimasi produksi the sebesar 140.322 ton. Kemudian produksi teh mengalami peningkatan di tahun 2024 dan tahun 2025 masing-masing sebesar 141.136 ton dan 141.950 ton.

Tabel 18. Hasil Estimasi Produksi Teh Model Fungsi Trasfer Tahun 2021-2025

Tahun	Produksi Teh (Ton)	Pertumbuhan (%)
2021	138.696	
2022	139.508	0,59
2023	140.322	0,58
2024	141.136	0,58
2025	141.950	0,58
<b>Rata-rata pertumbuhan (%)</b>		<b>0,58</b>

- **Model Vector Auto Regressive (VAR)**

Beberapa tahapan yang dilakukan dalam penyusunan model VAR antara lain persiapan data, pembagian data *training* dan *testing*, pemilihan lag dan type, pengajuan asumsi, ramalan data *training*, *testing*, penghitungan mape, dan plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Dibahas juga interpretasi *Impulse Response Function* (IRF) dan *Variance Decomposition*.

Variabel yang digunakan untuk estimasi model VAR antara lain produksi teh dalam satuan ton, luas areal dalam satuan hektare, harga konsumen dalam satuan Rp/Kg, volume ekspor teh dalam satuan ton dan volume impor teh dalam satuan ton. Semua data ini bersumber dari Ditjen Perkebunan kecuali data harga konsumen bersumber dari Badan Pusat Statistik, selain itu series data yang digunakan dimulai dari tahun 1980 hingga 2020.

Komposisi variabel tersebut, komponen konstanta dan trend juga perlu dilakukan pengujian untuk mengetahui signifikan atau tidak keberadaannya. Jika keduanya signifikan, maka komponen



tersebut harus dimasukkan ke dalam model VAR dengan type “both”. Jika hanya konstanta yang signifikan, maka trend perlu dikeluarkan dari model VAR dengan model VAR type “const”. Jika hanya trend yang signifikan maka konstanta dikeluarkan dari model menggunakan model VAR type “trend”, dan jika keduanya tidak signifikan, maka type yang digunakan model VAR adalah “none”.

Keberadaan konstanta dan trend dapat dideteksi dari plot data awal, namun terkadang hal tersebut sulit dilakukan. Untuk mengantisipasi hal tersebut, sebaiknya dilakukan uji coba/*trial and error* menggunakan model VAR dari lag  $p=1$  s.d 6 dengan type “both” untuk mengetahui signifikan atau tidak keberadaannya. Untuk panjang lag maksimum bisa dilakukan trial error sampai tidak memungkinkan untuk dilakukan permodelan.

Pada data produksi teh, setelah dilakukan running model VAR dengan lag  $p=1$  type “both”, diperoleh informasi komponen konstanta dan trend signifikan, sehingga pada pengujian berikutnya dilakukan running model VAR dengan menggunakan lag  $(p) = 1$  s.d lag  $(p)=4$  dengan menyertakan komponen konstanta dan trend. Pemilihan lag  $p$  ditentukan dengan melihat banyaknya variabel yang signifikan dalam lag tersebut sekaligus memastikan harus ada variabel/peubah yang signifikan pada lag terpilih dimaksud. Setelah dilakukan run model ternyata ada 2 kandidat model VAR terbaik yaitu Var (1) type “both” dan Var (2) type “both”. Type “both” artinya persamaan Var itu mengandung constanta dan faktor trend.

Untuk model VAR yang pertama adalah Model VAR ( $p=1$ ) dengan konstanta dan trend (type=both). Hasil model Var(1) type both, untuk mengestimasi Produksi (t) teh menunjukkan bahwa koefisien constanta dan trend menunjukkan hasil yang signifikan pada tingkat kepercayaan 95%. Untuk variabel lain yang signifikan untuk mengestimasi produksi (t) antara lain Produksi, luas areal, dan harga konsumen pada lag-1. Ketiga variabel tersebut signifikan dengan tingkat kepercayaan 95%.

Model Var (1) type both ini menghasilkan nilai Adjusted R Square = 82,62%, artinya keragaman produksi dipengaruhi oleh variabel penjelasnya sebesar 82,62%. Nilai p-value untuk model produksi ini sangat kecil atau jauh lebih kecil dari 0,05, sehingga model layak untuk digunakan. Jumlah variabel yang signifikan ada 3 variabel dari total 6 variabel. banyak sehingga model ini layak digunakan dan menjadi kandidat terbaik untuk model VAR.

Tabel 19. Output Model VAR(1) type=both

```

VAR Estimation Results:
=====
Endogenous variables: Produksi_TOTAL, Lareal_TOTAL, harga_kons, vol_ekspor, vol_impor
Deterministic variables: both
Sample size: 33
Log Likelihood: -1700.827
Roots of the characteristic polynomial:
0.8956 0.8956 0.6397 0.1547 0.1177
Call:
VAR(y = teh[1:34, c(2, 3, 5, 7, 8)], p = 1, type = "both")

Estimation results for equation Produksi_TOTAL:
=====
Produksi_TOTAL = Produksi_TOTAL.l1 + Lareal_TOTAL.l1 + harga_kons.l1 + vol_ekspor.l1 +
vol_impor.l1 + const + trend

              Estimate Std. Error t value Pr(>|t|)
Produksi_TOTAL.l1  9.482e-02  2.402e-01  0.395  0.69631
Lareal_TOTAL.l1   7.069e-01  2.433e-01  2.906  0.00739 **
harga_kons.l1    -2.472e-02  3.315e-02  -0.746  0.46251
vol_ekspor.l1     8.254e-02  1.343e-01  0.615  0.54414
vol_impor.l1     -4.180e-01  5.656e-01  -0.739  0.46654
const             1.681e+04  1.988e+04  0.846  0.40549
trend             9.905e+02  4.169e+02  2.376  0.02517 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 8847 on 26 degrees of freedom
Multiple R-Squared: 0.8262, Adjusted R-squared: 0.7861
F-statistic: 20.6 on 6 and 26 DF, p-value: 9.744e-09

```

Tabel 20. Output Model VAR(2) type=both

```

t Model VAR(2) type=both
VAR Estimation Results:
=====
Endogenous variables: Produksi_TOTAL, Lareal_TOTAL, harga_kons, vol_ekspor, vol_impor
Deterministic variables: both
Sample size: 32
Log Likelihood: -1617.311
Roots of the characteristic polynomial:
0.8751 0.7855 0.7855 0.7645 0.7645 0.7197 0.6452 0.6452 0.5003 0.5003
Call:
VAR(y = teh[1:34, c(2, 3, 5, 7, 8)], p = 2, type = "both")

Estimation results for equation Produksi_TOTAL:
=====
Produksi_TOTAL = Produksi_TOTAL.l1 + Lareal_TOTAL.l1 + harga_kons.l1 + vol_ekspor.l1
+ vol_impor.l1 + Produksi_TOTAL.l2 + Lareal_TOTAL.l2 + harga_kons.l2 + vol_ekspor.l2 +
vol_impor.l2 + const + trend

      Estimate Std. Error t value Pr(>|t|)
Produksi_TOTAL.l1 1.706e-01 2.993e-01 0.570 0.5749
Lareal_TOTAL.l1   9.943e-01 4.119e-01 2.414 0.0255 *
harga_kons.l1    -1.582e-02 6.116e-02 -0.259 0.7985
vol_ekspor.l1     1.934e-02 1.907e-01 0.101 0.9202
vol_impor.l1     -1.065e+00 1.552e+00 -0.686 0.5006
Produksi_TOTAL.l2 -1.912e-01 2.866e-01 -0.667 0.5123
Lareal_TOTAL.l2  -3.026e-01 4.424e-01 -0.684 0.5018
harga_kons.l2    -1.596e-02 7.657e-02 -0.209 0.8369
vol_ekspor.l2     1.302e-01 1.870e-01 0.697 0.4941
vol_impor.l2      6.622e-01 2.000e+00 0.331 0.7440
const             2.477e+04 2.447e+04 1.012 0.3235
trend             1.291e+03 6.675e+02 1.934 0.0674 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 9657 on 20 degrees of freedom
Multiple R-Squared: 0.8174, Adjusted R-squared: 0.7169
F-statistic: 8.137 on 11 and 20 DF, p-value: 3.225e-05

```

Hasil model Var(2) type both, untuk mengestimasi Produksi (t) teh menunjukkan bahwa koefisien constanta dan trend menunjukkan hasil yang signifikan pada tingkat kepercayaan 99%. Untuk variabel lain yang signifikan untuk mengestimasi produksi (t) antara lain Produksi (t-1), luas areal

(t-1), harga konsumen (t-1), dan volume ekspor (t-1), serta volume impor teh (t-1) dengan tingkat kepercayaan 90% - 95%. Jumlah variabel sebanyak 5 variabel. Model VAR(2) type=both ini juga layak perlu diuji MAPE untuk melihat apakah lebih baik dari model VAR(1) type=both.

Selanjutnya dilakukan pengujian model VAR, dengan nilai p=3, p=4 type both, dan model VAR p=3, p=4 type konstan. Hasil pengujian untuk Var(3) type 'both' menghasilkan koefisien konstanta dan trend signifikan. Jika dibandingkan dengan p=2, untuk p=3 penambahan variabel yang signifikan relatif kecil. Selanjutnya dilakukan pengujian untuk nilai p=4, ternyata konstanta dan trend sudah tidak signifikan, sehingga berhenti pada p=4. Selanjutnya dilakukan pengujian untuk nilai p=3 dan p=4 tetapi type 'constant', menunjukkan hasil koefisien konstanta tidak signifikan, sehingga berhenti untuk dilakukan pengujian lebih lanjut. Berdasarkan pengujian ini yang terbaik adalah untuk nilai p=1 atau p=2 dengan type 'both'.

Pemeriksaan autokorelasi residual model menggunakan fungsi "serial.test" yang di dalamnya dilakukan pengujian *Portmanteau-and Breusch-Godfrey test*. Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi. Pengujian Jarque-Bera tests untuk menguji kenormalan, hasil pengujian menunjukkan Nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal. Pemeriksaan heteroskedastisitas model menggunakan fungsi "arch.test" yang di dalamnya dilakukan pengujian *ARCH-LM tests*. Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.

Tabel 21, Ouput Pengujian Asumsi VAR(2) type=both

```

Portmanteau Test (asymptotic)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 285.91, df = 350, p-value = 0.9948

$JB
  JB-Test (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 31.417, df = 10, p-value = 0.0005006
$Skewness
  Skewness only (multivariate)
data: Residuals of VAR object varhsheet.b2
Chi-squared = 18.299, df = 5, p-value = 0.002594

$Kurtosis
  Kurtosis only (multivariate)
data: Residuals of VAR object varhsheet.b2

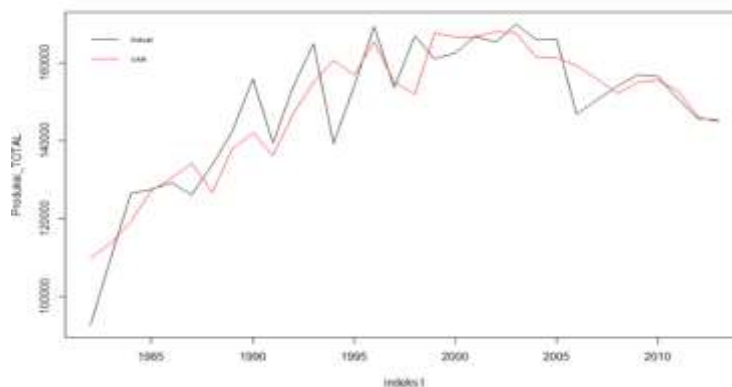
```

Chi-squared = 13.118, df = 5, p-value = 0.0223  
 ARCH (multivariate)  
 data: Residuals of VAR object varhsheet.b2  
 Chi-squared = 405, df = 1125, p-value = 1

Tabel 22. Nilai MAPE untuk Model VAR (2) type=both

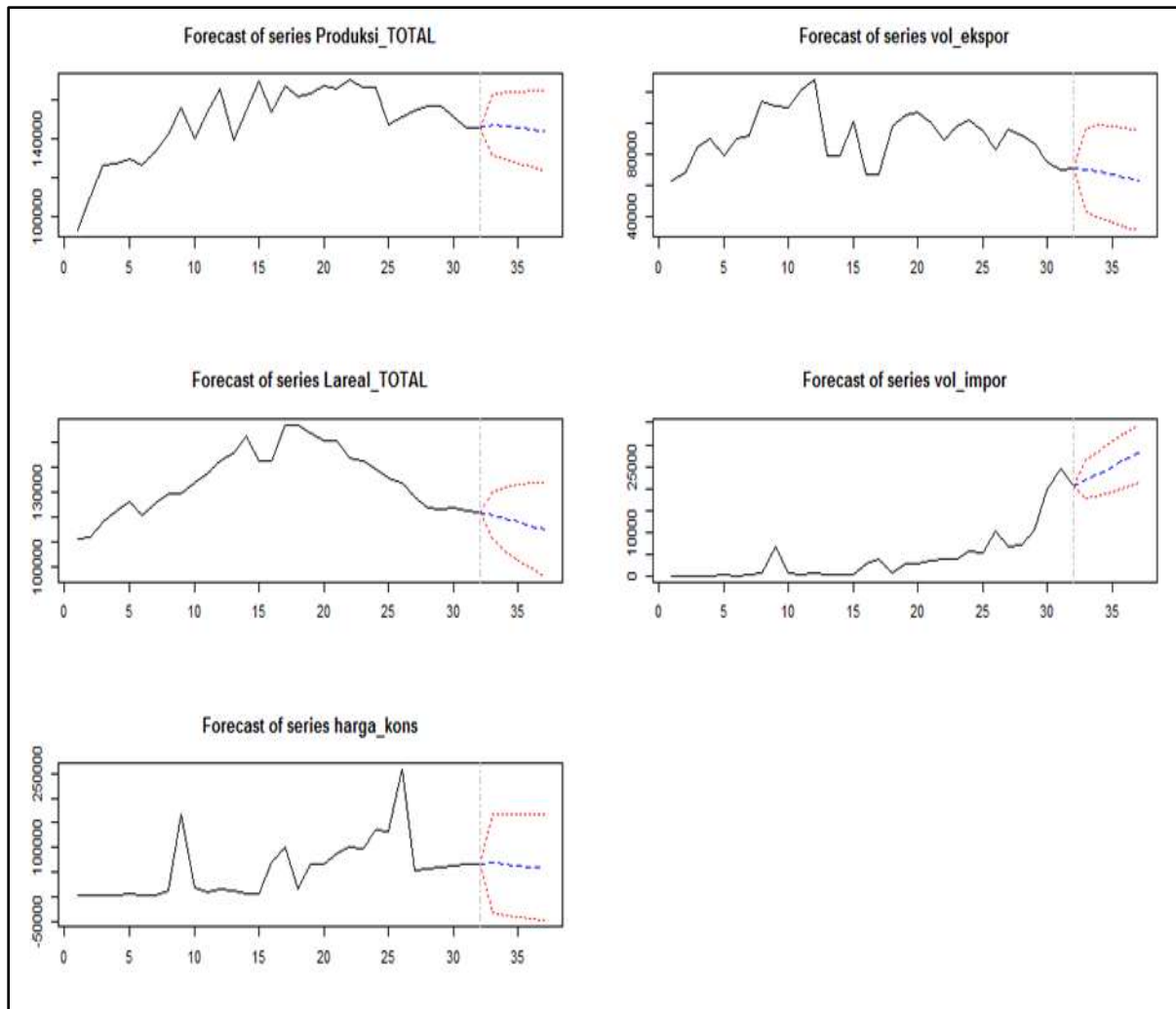
Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.1742	1.2219	2.6070	3.9705	5.3474	18.4368

Dari grafik di bawah dapat dilihat bahwa pergerakan ramalan pada data testing tidak mengikuti pergerakan data asli/aktual. Sehingga mungkin model VAR (1) type “both” bukan model yang terbaik. Hasil peramalan dengan model VAR (2) ini cenderung lebih tinggi dari data aktual, kecuali di tahun 2017 data aktual lebih tinggi dari data estimasi dengan VAR (2). Untuk data testing ini rata-rata penyimpangannya adalah sebesar 3,97%.



Gambar 11. Plot Ramalan dan Aktual VAR (2) Type “Both”

Keunggulan model VAR ini pada saat dilakukan permalan maka semua variabel dalam system akan ikut diramalkan. Hasil peramalan bukan hanya target variabel, tetapi semua variabel yang masuk dalam sistem. Hasil permalan 5 tahun ke depan untuk variabel produksi menunjukkan produksi akan menurun, begitu juga luas areal teh juga cenderung turun. Sebaliknya harga konsumen cenderung meningkat. Volume impor teh juga diperkirakan akan meningkat seiring dengan peningkatan produksi, sementara itu untuk volume impor teh diperkirakan akan tetap stagnan.



Gambar 12. Hasil Estimasi Model VAR (2) Type “Both”

Hasil estimasi produksi teh pada model VAR (2) Type “Both” berfluktuasi cenderung menurun. Rata-rata pertumbuhan untuk 5 tahun kedepan yaitu periode 2021 hingga 2025 akan turun sebesar 0,51%. Tahun 2021 produksi teh diestimasi sebesar 146.948 ton. Produksi teh tahun 2022 sebesar 146.327 ton. Produksi teh diramalkan terus meningkat hingga tahun 2025. Tahun 2023 hasil estimasi produksi the sebesar 145.583 ton. Kemudian produksi teh mengalami peningkatan di tahun 2024 dan tahun 2025 masing-masing sebesar 144.776 ton dan 143.948 ton.

Tabel 23. Hasil Estimasi Produksi Teh VAR (2), Type “Both”

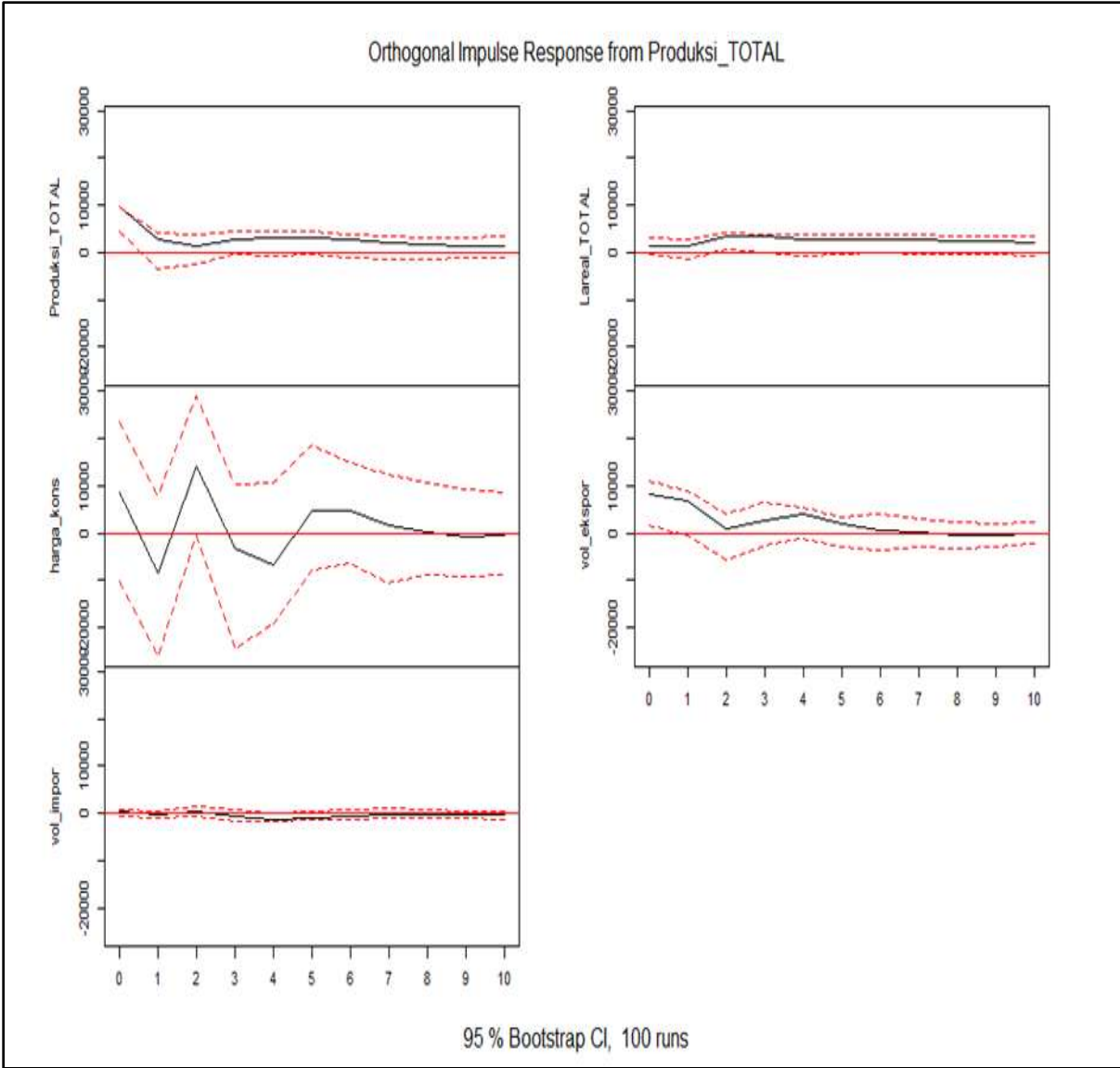
<b>Tahun</b>	<b>Produksi (Ton)</b>	<b>Pertumbuhan (%)</b>
2021	146.948	
2022	146.327	-0,42
2023	145.583	-0,51
2024	144.776	-0,55
2025	143.948	-0,57
<b>Rata-rata Pertumbuhan (%)</b>		<b>-0,51</b>

### **Interpretasi Impulse Response Function (IRF) dan Variance Decomposition**

Setelah diperoleh model terbaik, untuk permodelan VAR dapat dilakukan interpretasi lebih mendalam terkait Impulse Response Function (IRF) dan Variance Decomposition. Berikut interpretasi ketiga hal di atas untuk model terbaik VAR (2) type “both”.

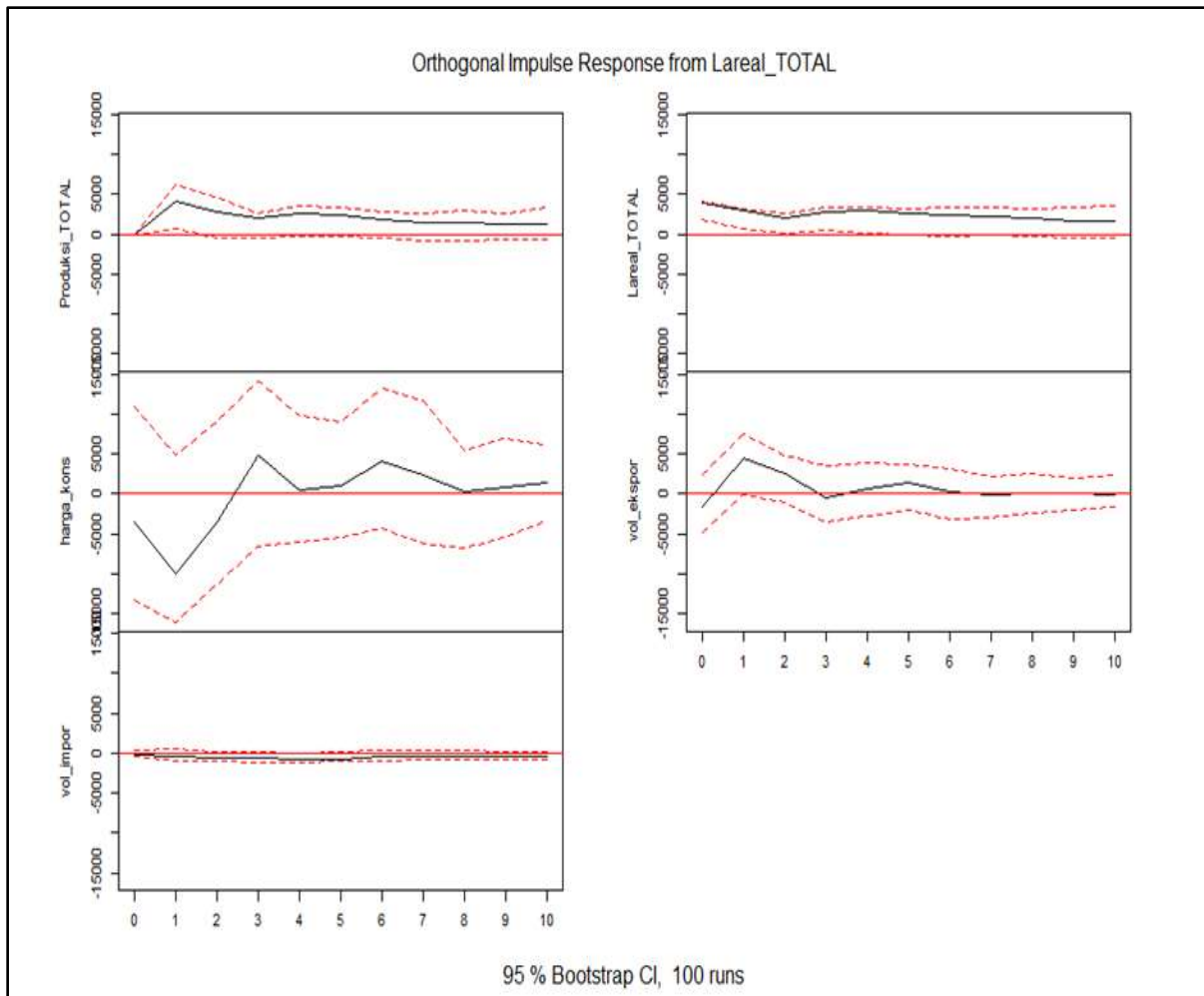
**Impulse Response Function** akan menjelaskan bagaimana perubahan atau gejolak yang terjadi pada suatu variabel/peubah di tahun tertentu akan berdampak pada variabel/peubah lain di tahun tertentu dan tahun-tahun setelahnya.

Dari grafik Impulse Response Function produksi model terbaik VAR (2) type “both”, dapat dilihat bahwa jika terjadi perubahan pada produksi di tahun tertentu maka akan berdampak pada produksi itu sendiri sampai 3 tahun ke depan. Dampak tersebut akan hilang setelah 3 tahun. Sedangkan dampak perubahan produksi tidak berdampak pada harga konsumen teh, luas areal, produksi teh, dan volume impor teh, namun masih berdampak pada volume ekspor kedepan. Perubahan produksi berdampak pada volume ekspor sampai dengan dua tahun dan akan hilang dampaknya setelah melewati tahun kedua tersebut.



Gambar 13. Impulse Respon Produksi Teh Model VAR (2) Type “Both”





Gambar 14. Impulse Respon Luas Areal Teh Model VAR (2) Type “Both”

- **Pemilihan Model Terbaik**

Salah satu dasar penentuan model terbaik dapat dilakukan dengan membandingkan nilai MAPE untuk data testing dan training yaitu dengan memilih nilai MAPE yang paling kecil, terutama untuk data testing. Selain MAPE yang terkecil, pola pergerakan ramalan juga harus diperhatikan. Pilihlah plot yang paling berhimpit/bersesuaian dengan data asli/aktual atau dengan kata lain performa hasil ramalan seiring dengan data historisnya.

Penyusunan angka estimasi produksi teh dilakukan dengan cara uji coba model statistik antara lain model ARIMA, Fungsi Transfer dan VAR.

Model pertama adalah ARIMA (1,2,2) dan ARIMA (3,2,5), dimana pada ARIMA (1,2,2) menghasilkan MAPE training sebesar 5,30% dan MAPE testing sebesar 7,83%

Hal ini menunjukkan bahwa kemampuan untuk melakukan estimasi dengan model ARIMA ini rata-rata akan mengalami kesalahan sekitar 7,83% lebih tinggi atau 7,83% lebih rendah. Untuk ARIMA (3,2,5) menghasilkan MAPE training sebesar 3,86% dan MAPE testing sebesar 6,23%

Hal ini menunjukkan bahwa kemampuan untuk melakukan estimasi dengan model ARIMA ini rata-rata akan mengalami kesalahan sekitar 6,23% lebih tinggi atau 6,23% lebih rendah.

Metode estimasi yang kedua adalah dengan model fungsi transfer, untuk melakukan estimasi produksi teh dengan variabel bebas adalah luas tanaman menghasilkan. Untuk model fungsi transfer ini menghasilkan MAPE data training 5,56%, sementara untuk MAPE data testing sebesar 4,27%. Model fungsi transfer ini menghasilkan MAPE yang lebih kecil dibandingkan model ARIMA, sehingga model fungsi transfer lebih akurat dalam melakukan estimasi. Hasil estimasi juga menunjukkan angka yang lebih realistis, dengan angka estimasi tahun 2021 sebesar 138,70 ribu ton. Namun untuk estimasi 5 tahun kedepan angka pertumbuhan sebesar 0,58%/tahun, sementara angka pertumbuhan 5 tahun sebelumnya sebesar 0,99%/tahun, jadi jauh lebih kecil dibandingkan data historisnya.

Untuk model estimasi yang terakhir adalah dengan model VAR (Vector Auto Regressive). Untuk model VAR ini menggunakan 5 variabel yaitu produksi, luas areal, harga konsumen, volume ekspor dan volume impor teh. Model yang terbaik untuk Model VAR adalah nilai  $p=2$  dan  $type="both"$ ,  $p=2$  artinya menggunakan variabel bebas sampai lag-2 dan ada konstanta dan factor trend. Estimasi produksi teh dengan menggunakan model VAR ini menghasilkan ketelitian yang cukup tinggi yaitu MAPE untuk data training 5,79% dan MAPE untuk data testing 6,74%. MAPE untuk data testing ini paling kecil dibandingkan dengan model-model lainnya, sehingga model

VAR ini disimpulkan paling kecil kesalahan estimasinya atau model VAR model terbaik untuk estimasi produksi teh nasional. Jika dibandingkan angka pertumbuhan produksi teh antara hasil estimasi 5 tahun kedepan dengan rerata pertumbuhan turun 0,51%, sementara angka pertumbuhan 5 tahun terakhir sebelumnya sebesar 0,99%. Angka hasil estimasi untuk produksi teh nasional tahun 2021 sebesar 146,95 ribu ton, sementara untuk angka sementara tahun 2020 sebesar 144,06 juta ton atau meningkat sebesar 2,00%.

Tabel 24. Perbandingan MAPE Produksi Teh Indonesia Model Arima, Fungsi Transfer dan VAR

	Pengujian MAPE	Model ARIMA				Fungsi Transfer		Model VAR		Hasil Pleno	
		ARIMA (1,2,2)	(%)	ARIMA (3,2,5)	(%)	Arima (0,2,1) Xreg=LTM	(%)	VAR (2) p=2 type=both	(%)		
		MAPE Training	5,30		3,86		5,56		5,79		
		MAPE Testing	7,83		6,23		4,27		6,74		
ATAP	2016	138.935		138.935		138.935		138.935			
	2017	146.251	5,27	146.251	5,27	146.251	5,27	146.251	5,27		
	2018	140.236	-4,11	140.236	-4,11	140.236	-4,11	140.236	-4,11		
	2019	137.803	-1,73	137.803	-1,73	137.803	-1,73	137.803	-1,73		
	2020	144.064	4,54	144.064	4,54	144.064	4,54	144.064	4,54		
Angka Estimasi (AESTI)	2021	148.050	2,77	136.931	-4,95	138.696	-3,73	146.948	2,00		
	2022	146.512	-1,04	134.032	-2,12	139.508	0,59	146.327	-0,42		
	2023	149.605	2,11	143.219	6,85	140.322	0,58	145.583	-0,51		
	2024	148.815	-0,53	134.503	-6,09	141.136	0,58	144.776	-0,55		
	2025	151.281	1,66	136.339	1,37	141.950	0,58	143.948	-0,57		
Rata-rata Pertumbuhan	ATAP 2016 - 2020		0,99		0,99		0,99		0,99		
	AESTI 2021 - 2025		0,55		0,00		0,58		-0,51		

Catatan Pengolah: Berdasarkan Perbandingan MAPE Data Training dan Testing maka Model Terbaik adalah : Fungsi Transfer Arima (0,2,1) Xreg= LTM

## KESIMPULAN

Model yang digunakan untuk penyusunan angka estimasi produksi teh antara lain model *Autoregresif Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Autoregressive* (VAR). Sementara data yang terdapat pada statistik perkebunan merupakan hasil sinkronisasi dan validasi yang dilaksanakan secara berjenjang dari tingkat kabupaten/kota sampai tingkat nasional dengan mengacu pada Pedoman Pengelolaan Data Komoditas Perkebunan (PDKP) yang diterbitkan oleh Direktorat Jenderal Perkebunan. Secara periodik, data perkebunan disajikan dalam 3 status angka yaitu Angka Tetap (ATAP), Angka Sementara (ASEM) dan Angka Estimasi (AESTI). Angka Tetap merupakan angka hasil rekapitulasi dari pelaporan yang sudah tetap, sehingga tidak dilakukan estimasi, sedangkan untuk penentuan Angka Sementara dan Estimasi perlu dilakukan estimasi dengan metode estimasi yang paling relevan dan tepat (PDKP, 2013). Data Angka Tetap (ATAP) merupakan data 2 tahun yang lalu (n-2), Angka Sementara (ASEM)

merupakan data tahun lalu ( $n-1$ ), dan Angka Estimasi (AESTI) merupakan data tahun yang berjalan ( $n$ ).

Untuk analisis ini data dibagi menjadi 2 kelompok, yaitu data training tahun 1980 – 2014, dan data testing tahun 2015 – 2020. Data training untuk penyusunan model, sedangkan data testing untuk uji coba model dalam melakukan estimasi 6 tahun kedepan. Untuk estimasi produksi teh alternatif model pertama adalah Model ARIMA. Model ARIMA terbaik adalah ARIMA (1,2,2), menghasilkan MAPE untuk data training 5,30%, dan MAPE data testing 7,83%. Model yang kedua dengan menggunakan Fungsi Transfer dengan variabel input luas tanaman menghasilkan, dengan MAPE untuk data training sebesar 5,56% dan MAPE data testing 4,27%. Untuk model yang ketiga model VAR(2) type 'both' ada pengaruh trend dan konstanta, menghasilkan MAPE data training 5,79% dan data MAPE data testing 6,74%.

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen. Model VAR berlaku pada saat nilai setiap variabel dalam sebuah system tidak hanya bergantung pada lag-nya sendiri, namun juga pada nilai lag variabel lain.

Tahapan yang dilakukan dalam penyusunan model VAR adalah sebagai berikut: persiapan data, pembagian data training dan testing, pemilihan lag dan type, pengajuan asumsi, ramalan data training, testing, penghitungan mape, dan plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Disamping itu dibahas Interpretasi Impulse Response Function (IRF) dan Variance Decomposition.

Berdasarkan perbandingan besarnya MAPE baik data testing maupun data training dan hasil estimasi produksi 5 tahun kedepan, maka model terbaik yang terpilih adalah model VAR(2) tipe 'both' karena menghasilkan akurasi yang cukup tinggi sehingga MAPE rata-rata sebesar 4,92%. Hasil estimasi produksi teh nasional dalam bentuk daun kering untuk model VAR (2) type 'both' untuk tahun 2021 sebesar 146,95 ribu ton, tahun 2022 sebesar 146,33 ribu ton, tahun 2023 sebesar 145,58 ribu ton, tahun 2024 sebesar 144,78 ribu ton, dan tahun 2025 sebesar 143,95 ribu ton. Laju pertumbuhan estimasi produksi teh nasional selama 5 tahun kedepan rata-rata 0,99%.

## DAFTAR PUSTAKA

- Budiar, Aindra. 2013. Peramalan Produksi TBS Kelapa Sawit Dengan Model Arima dan Model Fungsi Transfer Input Ganda. Institut Pertanian Bogor, Bogor
- Gujarati, D.N. and D.C. Porter, 2010. Dasar-dasar Ekonometrika. Jakarta: Penerbit Salemba Empat, Buku 2, Edisi 5.
- Heizer, J., Render, B. & Munson, C., 2011. Operations Management Sustainability and Supply Chain Management. Boston: Pearson.
- Kementerian Pertanian. 2013. Pedoman Teknis Pengembangan Tanaman Teh Tahun 2014. Jakarta: Direktorat Jenderal Perkebunan Kementerian Pertanian.
- Kementerian Pertanian. 2019. Statistik Perkebunan Indonesia 2018-2020 (Teh). Jakarta: Direktorat Jenderal Perkebunan.
- Mardiyanto, Atqo. 2000. Kajian Peramalan Dengan Model Struktural dan Non Struktural (VAR dan ARIMA). Institut Pertanian Bogor, Bogor
- Makridakis S, Wheelwright SC, McGee VE. 1999. Metode dan Aplikasi Peramalan. Suminto H, penerjemah. Binarupa Aksara, Jakarta.
- M. Firdaus 2020. Aplikasi Ekonometrika dengan E-Views, Stata dan R. Institut Pertanian Bogor (IPB) Press.
- Tsay, R.S., 2014. Multivariate Time Series Analysis With R and Financial Applications. New Jersey: John Wiley & Sons.



# **KAJIAN METODOLOGI ESTIMASI DATA PRODUKSI CENGKEH**

Vera Junita Siagian, Fungsional Statistisi, Pusdatin Kementan

## **ABSTRAK**

Pada tahun 2021, Pusdatin bekerjasama dengan Ditjen Perkebunan dan BPS mengkaji 3 metode estimasi yang nantinya didapatkan metode yang lebih akurat, lebih objektif dan lebih baik secara statistik dari metode sebelumnya dan diharapkan hasil estimasinya cenderung mendekati angka aktualnya. Melalui kegiatan pengembangan metode estimasi data perkebunan tahun 2021, diharapkan akan dihasilkan metode yang paling sesuai untuk mengestimasi produksi cengkeh nasional dengan membandingkan besaran MAPE (Mean Absolut Percentage Error) dan kerealistisan hasil permalan dengan data series sebelumnya

Tujuan dari makalah ini adalah untuk membandingkan efektivitas metode tersebut dalam hasil estimasi produksi cengkeh dan menentukan metode terbaik dalam estimasi produksi cengkeh nasional. Metode yang diterapkan adalah ARIMA (*Autoregressive Integrated Moving Average*), Fungsi Transfer dan VAR (*Vector Auto Reggression*) dengan menggunakan software R Studio

## **ABSTRACT**

*In 2021, Center for Agriculture Data and Informasi System (CADIS) together with the Directorate General of Plantation and the BPS-Statistic Indonesia examined 3 methods of estimation that will be obtained more accurately, more objectively and statistically better than the previous method and it is expected that the estimated results tend to be close to the actual number. Through the development of plantation data estimation method in 2021, it is expected that the most suitable method will be produced to estimate national clove production by comparing the amount of MAPE (Mean Absolut Percentage Error) and the reality of the results of the permalan with the previous series data*

*The purpose of this paper is to compare the effectiveness of such methods in the estimated production of cloves and determine the best method in estimating national clove production. The methods applied are ARIMA (Autoregressive Integrated Moving Average), Transfer Function and VAR (Vector Auto Reggression) using R Studio software*

## PENDAHULUAN

Cengkeh atau *Syzygium aromaticum* merupakan jenis tumbuhan perdu yang memiliki batang pohon besar dan berkayu keras dari keluarga pohon *Myrtaceae*. Cengkeh adalah tanaman asli Indonesia, banyak digunakan sebagai bumbu masakan pedas di negara-negara Eropa, dan sebagai bahan utama rokok kretek khas Indonesia. Cengkeh merupakan salah satu komoditas sub sektor perkebunan yang sebagian besar (97,43%) diusahakan oleh perkebunan rakyat. Hasil utama tanaman cengkeh adalah bunganya yang dipanen pada saat kelopak bunga belum mekar. Bunga cengkeh kering merupakan salah satu bahan baku utama untuk rokok kretek yang merupakan rokok khas Indonesia.

Indonesia memiliki potensi besar sebagai salah satu negara penghasil cengkeh di dunia. Dalam lima tahun terakhir, produksi cengkeh nasional menunjukkan tren yang meningkat. Produksi cengkeh Indonesia pada tahun 2021 diperkirakan mencapai 140.997 ton. Dibandingkan tahun 2020 (Angka Sementara), produksi naik sebesar 0,13% yaitu sebesar 140.812 ton. Dilihat dari produktivitas, tahun 2021 produktivitas cengkeh nasional diperkirakan sebesar 416 kg/ha. Dibandingkan tahun lalu yaitu tahun 2020 sedikit naik yaitu sebesar 416 kg/ha.

Untuk menyusun angka estimasi (AESTI) produksi cengkeh nasional menurut Buku Pedoman Pelaksanaan Pengeolaan Data Komoditas Perkebunan (PDKP), metode yang dilakukan selama ini adalah metode pemulusan eksponensial tunggal atau Single Exponential Smoothing (SES) dan pemulusan eksponensial ganda atau DSS (Double Exponential Smoothing). Untuk pemilihan model terbaik yaitu dengan membandingkan besaran MAPE (Mean Absolut Percentage Error) dan kerealistisan hasil permalan dibandingkan dengan data series sebelumnya. Menurut T. Hani Handoko (2011), *Exponential Smoothing* adalah suatu tipe teknik peramalan rata-rata bergerak yang melakukan penimbangan terhadap data masa lalu dengan cara eksponensial sehingga data paling akhir mempunyai bobot atau timbangan lebih besar dalam rata-rata bergerak.

Pada tahun 2021 ini, Pusdatin berkerja sama dengan Ditjen perkebunan dan BPS mengkaji 3 (tiga) metode estimasi guna menghasilkan angka yang nilainya semakin mendekati nilai sebenarnya. Ke tiga metode yang di kaji yaitu ARIMA (*Autoregressive Integrated Moving Average*), Fungsi Transfer dan VAR (*Vector Auto Reggression*) dengan menggunakan software R Studio



## MATERI DAN METODE

Data yang digunakan dalam kajian ini adalah data sekunder yang bersumber dari Direktorat Jenderal Perkebunan dan Badan Pusat Statistik (BPS). Untuk data produksi cengkeh bersumber dari Direktorat Jenderal Perkebunan sementara untuk data harga produsen dan volume ekspor bersumber dari BPS dengan series data tahunan yaitu dari tahun 1983 sampai tahun 2020 (Angka Sementara)

### Model Arima

ARIMA sering juga disebut metode runtun waktu Box-Jenkins. ARIMA sangat baik ketepatannya untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang ketepatan peramalannya kurang baik. Biasanya akan cenderung flat (mendatar/konstan) untuk periode yang cukup panjang. Model Autoregresif Integrated Moving Average (ARIMA) adalah model yang secara penuh mengabaikan independen variabel dalam membuat peramalan. ARIMA menggunakan nilai masa lalu dan sekarang dari variabel dependen untuk menghasilkan peramalan jangka pendek yang akurat. ARIMA cocok jika observasi dari deret waktu (time series) secara statistik berhubungan satu sama lain (dependent) [https://daps.bps.go.id/ile\\_artikel/77/arima.pdf](https://daps.bps.go.id/ile_artikel/77/arima.pdf)

Metode ARIMA dibagi kedalam tiga kelompok model, yaitu *autoregressive model* (AR), *moving average model* (MA) dan model campuran yang memiliki karakteristik kedua model di atas yaitu *autoregressive integrated moving average* (ARIMA).

#### *Autoregressive Model* (AR)

AR adalah suatu model yang menjelaskan pergerakan suatu peubah melalui peubah itu sendiri di masa lalu. Bentuk umum model autoregressive dengan ordo p (AR(p)) atau model ARIMA (p,0,0) dinyatakan sebagai berikut:

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t

$Y_{t-p}$  = data *time series* pada kurun waktu ke (t-P)

$\mu$  = suatu konstanta

$\theta_1 \dots \theta_p$  = parameter autoregressive ke-p

$\varepsilon_t$  = nilai kesalahan pada waktu ke t

#### *Moving Average* (MA)

MA adalah suatu model yang melihat pergerakan variabelnya melalui sisaannya di masa lalu. Bentuk model MA dengan ordo  $q$  atau MA ( $q$ ) atau model ARIMA ( $0, d, q$ ) dapat ditulis sebagai berikut:

$$Y_t = \mu - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke- $t$

$\phi_1 \dots \phi_q$  = parameter-parameter moving average

$\varepsilon_{t-q}$  = nilai kesalahan pada waktu ke ( $t-q$ )

*Autoregressive Intergrated Moving Average (ARIMA)*

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk modelnya adalah :

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke- $t$

$Y_{t-p}$  = data *time series* pada kurun waktu ke ( $t-p$ )

$\mu$  = suatu konstanta

$\theta_1 \theta_2 \dots \theta_p \phi_1 \phi_2 \dots \phi_q$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke- $(t-q)$

### Model Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai dari prediksi masa depan dari suatu deret berkala (disebut deret output atau  $Y_t$ ) didasarkan pada nilai masa lalu dari deret itu sendiri ( $Y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $X_t$ ) dengan deret output tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke- $t$ , tetapi juga pada waktu  $t+1, t+2, \dots, t+k$ . Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara variabel *input* dan variabel *output*.

Model umum Fungsi Transfer:

$$y_t = D(B)x_t + N_t \quad \longrightarrow \quad y_t = \frac{\theta_s(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t$$

Dimana:

$b \rightarrow$  panjang jeda pengaruh  $X_t$  terhadap  $Y_t$

- r → panjang lag Y periode sebelumnya yang masih mempengaruhi  $Y_t$
- s → panjang jeda X periode sebelumnya yang masih mempengaruhi  $Y_t$
- p → ordo AR bagi noise  $N_t$
- q → ordo MA bagi noise  $N_t$

#### Model Var

VAR (*Vector Auto Regression*) digunakan untuk memproyeksikan sebuah sistem dengan variabel runtut waktu dan untuk menganalisis dampak dinamis dari faktor gangguan yang terdapat dalam sistem variabel tersebut. Pada dasarnya Analisis VAR sama dengan suatu model persamaan simultan, karena dalam Analisis VAR kita mempertimbangkan beberapa variabel endogen secara bersama-sama dalam suatu model. Sebenarnya analisis ini mirip dengan model persamaan simultan biasa. Hanya saja di dalam Analisis VAR masing-masing variabel selain diterangkan oleh nilainya di masa lampau, juga dipengaruhi oleh nilai masa lalu dari semua variabel endogen lainnya dalam model yang diamati. Selain itu, dalam analisis VAR biasanya tidak ada variabel eksogen di dalam model.

Pemeriksaan lag digunakan untuk menentukan panjang lag optimal yang akan digunakan selanjutnya menemukan estimasi parameter untuk model Vector Autoregressive (VAR). Dalam model VAR, panjang lag menunjukkan derajat bebas. Model terbaik adalah model yang memiliki nilai Akaike Information Criterion (AIC) terkecil.

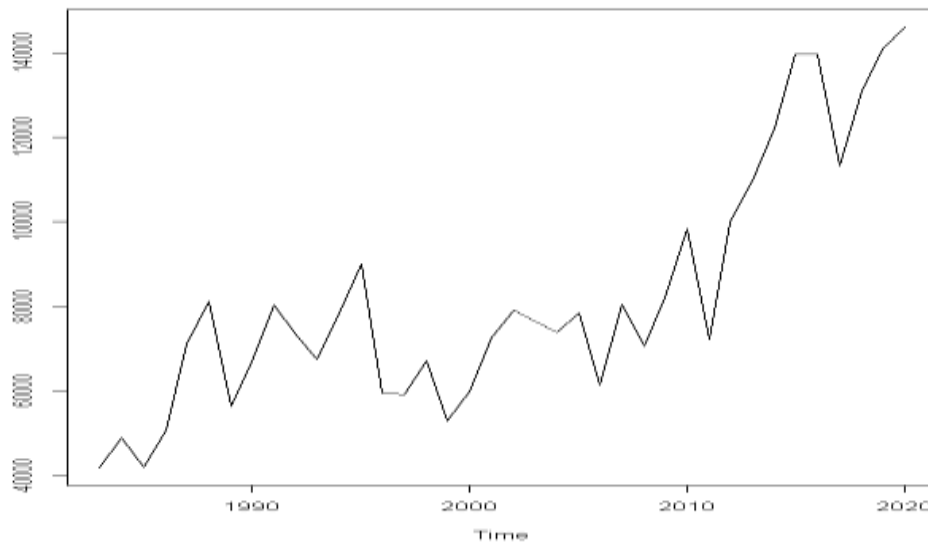
### HASIL DAN PEMBAHASAN

R Studio adalah bahasa pemrograman dan sistem perangkat lunak yang dirancang khusus untuk mengerjakan segala hal terkait komputasi statistik. Bahasa pemrograman ini pertama kali dikembangkan pada tahun 1993 oleh dua orang pakar statistik yaitu Ross Ihaka dan Robert Gentleman di Auckland University, New Zealand. Sampai saat ini, bahasa pemrograman R terus berkembang secara pesat seiring dengan semakin populernya terminologi “*Big Data*” dan meningkatnya kebutuhan perusahaan akan data *scientist* untuk mengolah dan menganalisis data di perusahaan tersebut sebagai dasar pengambilan kebijakan dan mengotomatisasi proses bisnis menjadi data *driven*. Bahasa pemrograman seperti Python dan R telah menjadi pilihan utama bagi para peneliti maupun praktisi di bidang data *science* untuk mengolah dan menganalisis data baik itu untuk kepentingan penelitian maupun bisnis. Oleh karena itu, bagi

seorang pemula di bidang data *science*, R merupakan bahasa pemrograman yang sangat *recommended* untuk dikuasai.

### Model Arima

Syarat utama dalam melakukan pemodelan ARIMA adalah kestasioneran data. Kestasioneran data dapat diketahui secara visual (plot datanya) dan juga dengan uji statistik. Dari plot data gambar 1 terlihat produksi cengkeh tahun 1983-2019 memiliki trend sehingga terindikasi tidak stasioner.



**Gambar 1.** Plot Produksi Cengkeh Tahun 1983-2020

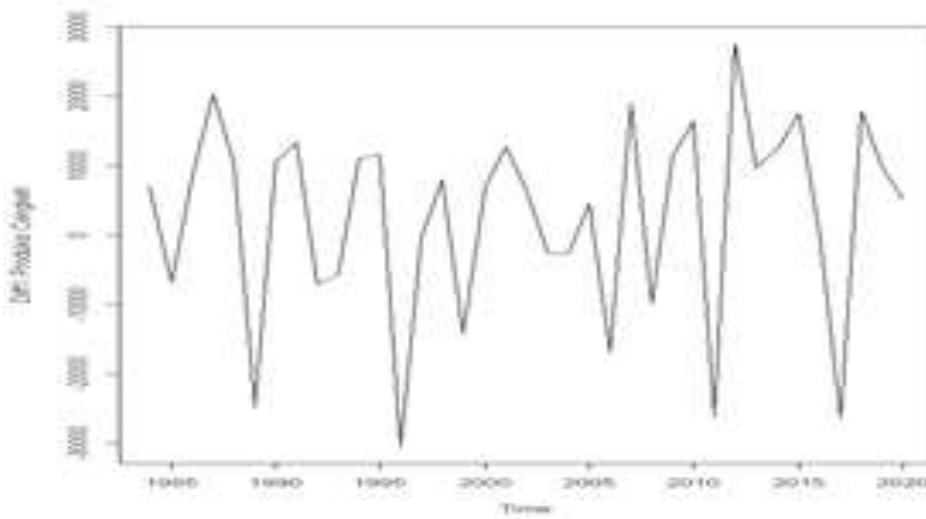
Berdasarkan gambar 1 terlihat bahwa data produksi cengkeh belum stasioner. Untuk lebih memastikan apakah data stasioner digunakan uji ADF (*Augmented Dickey-Fuller*). Jika test statistics lebih besar dari critical value maka data tersebut tidak stasioner. Berdasarkan hasil uji *Augmented Dickey-Fuller* dapat dilihat seperti hasil dibawah ini.

**Tabel 1.** Hasil Uji Augmented Dickey-Fuller

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
Test regression trend  
Test regression trend  
Call:  
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)  
Residuals:  
  Min   1Q Median   3Q   Max  
-28529 -6686  3442  9289 20567  
Coefficients:  
      Estimate Std. Error t value Pr(>|t|)  
(Intercept) 13093.4490  8412.2658  1.556  0.1294  
z.lag.1      -0.2999   0.1611 -1.861  0.0719 .  
tt           743.0080  383.2038  1.939  0.0614 .  
z.diff.lag   -0.1245   0.1826 -0.682  0.5004  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
Residual standard error: 13770 on 32 degrees of freedom  
Multiple R-squared:  0.1801,    Adjusted R-squared:  0.1032  
F-statistic: 2.343 on 3 and 32 DF,  p-value: 0.09161  
Value of test-statistic is: -1.8612 2.0631 1.9964  
Critical values for test statistics:  
  1pct  5pct 10pct  
tau3 -4.15 -3.50 -3.18  
phi2  7.02  5.13  4.31  
phi3  9.31  6.73  5.61
```

Dari hasil uji Augmented Dickey-Fuller diatas terlihat bahwa nilai tes statistic adalah -1,8612 dan nilai critical tau 3 untuk 1% sebesar -4,15, taraf uji 5% sebesar -3,50 dan taraf uji 10% sebesar -3,18. Dari hasil uji Augmented Dickey-Fuller ini dapat disimpulkan data tidak stasioner hingga pada taraf 1%, 5% dan 10% yang berarti tolak  $H_0$ . Suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melakukan differencing, maka akan dilakukan differencing 1 kali. Setelah differencing satu kali, secara visual sudah terlihat data stasioner (Gambar 2) dan dari hasil uji Augmented Dickey-Fuller terlihat bahwa nilai tes statistic adalah -5.9826 dan nilai critical tau 1 untuk 1% sebesar -2,62, taraf uji 5% sebesar -

1,95 dan taraf uji 10% sebesar -1,61% sehingga dapat disimpulkan data sudah stasioner pada taraf 1%, 5% dan 10%.



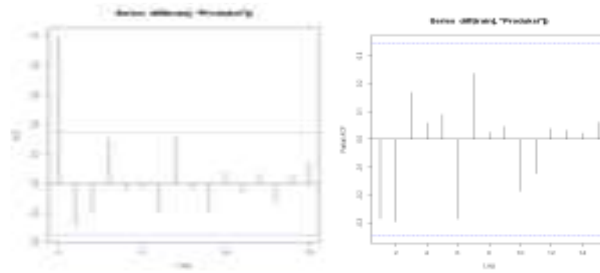
**Gambar 2.** Plot Produksi Cengkeh Tahun 1983-2020 Diff1

**Tabel 2.** Hasil Uji Augmented Dickey-Fuller

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
Call:  
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
Residuals:  
  Min   1Q Median   3Q   Max  
-23937 -2846  7598 13237 24534  
Coefficients:  
      Estimate Std. Error t value Pr(>|t|)  
z.lag.1  -1.5818   0.2642  -5.986  1e-06 ***  
z.diff.lag  0.2838   0.1678   1.691   0.1  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
Residual standard error: 14190 on 33 degrees of freedom  
Multiple R-squared:  0.646,    Adjusted R-squared:  0.6245  
F-statistic: 30.11 on 2 and 33 DF, p-value: 3.623e-08  
Value of test-statistic is: -5.9862  
Critical values for test statistics:  
  1pct  5pct 10pct  
tau1 -2.62 -1.95 -1.61
```

### Identifikasi Model

Pada tahap ini akan ditentukan nilai p,d dan q dimana penentuan p dan q dapat dilakukan dengan beberapa cara yaitu dengan melihat korelogram autokorelasi (ACF) dan korelogram autokorelasi parsial (PACF), autoarima serta ar maselect. Sebelum menentukan nilai p,d,dan q terlebih dahulu data produksi cengkeh dibagi menjadi data training yaitu data dari tahun 1983-2014 dan data testing mulai tahun 2015-2020. Data training digunakan untuk membangun model, sedangkan data testing digunakan untuk validasi dan pengepasan model. Dari gambar plot ACF pada Gambar 1 menunjukkan nyata pada lag=0 , sementara untuk lag=1 dan seterusnya berada dibawah garis selang kepercayaan dan tidak menunjukkan pola *cut off* maupun *tail off* sehingga agak kesulitan untuk mengidentifikasi model ARIMA nya.



**Gambar 3.** Plot ACF dan PACF Produksi Cengkeh

Apabila terjadi kesulitan mengidentifikasi model tentatif ARIMA menggunakan plot ACF dan PACFnya maka dapat digunakan *autoarima*. Hasil *autoarima* adalah ARIMA (0,1,1) dengan Standard error model sebesar 0,1693 dan nilai MAPE data training sebesar 14,58653% seperti terlihat pada tabel dibawah ini.

**Tabel 3.** Model Arima Berdasarkan Automodel

```
Series: train[, "Produksi"]
ARIMA(0,1,1) with drift
Coefficients:
      ma1  drift
-0.3970 2769.961
s.e. 0.1693 1436.407
sigma^2 estimated as 182791568: log likelihood=-348.84
AIC=703.68 AICc=704.54 BIC=708.08
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Training set 34.21727 12890.86 10548.54 -2.698447 14.58653 0.8577462 0.0005415266
```

Selain cara diatas dapat juga dijalankan dengan perintah `armaselect`, maka R-Studio akan menunjukkan sepuluh (10) model tentative yang diurutkan berdasarkan nilai SBC terkecil. Sepuluh model tentatif yang ditampilkan berdasarkan `armaselect` yaitu:



**Tabel 4.** Sepuluh Model Tentatif

	p	q	sbc
[1,]	5	5	501.9701
[2,]	4	5	541.0991
[3,]	2	2	582.5727
[4,]	3	1	583.3990
[5,]	1	2	584.6655
[6,]	4	1	585.0510
[7,]	4	3	585.3204
[8,]	1	1	585.6550
[9,]	1	3	586.0668
[10,]	4	2	586.3882

Dari hasil model autoarima dan 10 model tentative hasil *armaselect* dicobakan satu per satu dengan melihat mape terkecil dan kelayakan modelnya serta kelogisan hasil ramalan datanya. Setelah dicobakan satu persatu maka mape yang terkecil adalah ARIMA (2,1,2)

**Tabel 5.** Model Arima Berdasarkan Automodel

Call:

```
arima(x = train[, "Produksi"], order = c(2, 1, 2))
```

Coefficients:

```
ar1          ar2      ma1      ma2
```

```
-0.2338 -0.8947 -0.0557 0.9993
```

```
s.e. 0.1082 0.0912 0.1175 0.2651
```

```
sigma^2 estimated as 1.37e+08: log likelihood = -347.31, aic = 704.62
```

```
library(lmtest)
```

```
coefstest(model1)
```

z test of coefficients:

```
Estimate Std. Error z value Pr(>|z|)
```

```
ar1 -0.233805 0.108197 -2.1609 0.0307017 *
```

```
ar2 -0.894683 0.091228 -9.8071 < 2.2e-16 ***
```

```
ma1 -0.055684 0.117513 -0.4739 0.6356031
```

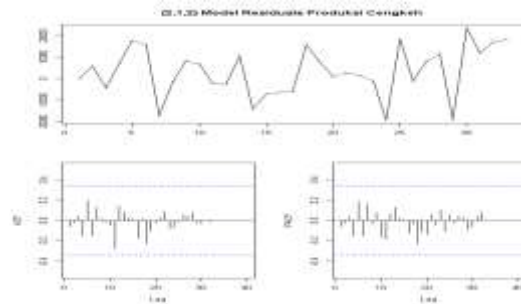
```
ma2 0.999275 0.265078 3.7697 0.0001634 ***
```

```
---
```

```
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Dari Hasil penelusuran model ARIMA (2,1,2) diatas menunjukkan parameter ar 2 dan ma 2 adalah sangat nyata. Kemudian dilihat sebaran sisaan LJung-Box dari lag 5 sampai lag 30. Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan LJung-Box. Hasil pengujian nilai p-value pada lag 5 sampai dengan lag 30 tidak ada yang signifikan, sehingga dapat

disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (2,1,2) layak digunakan. Hasil sisaan dengan Uji Ljung-Box dan hasil plot sisaan disajikan pada gambar dan table dibawah ini



**Gambar 4.** Plot Sisaan Model Arima (2,1,2)

Hasil plot sisaan model ARIMA (2,1,2) menunjukkan bahwa plot ACF dan PACF sudah tidak menunjukkan pola *cut off* atau *tail off* yang merujuk ke model ARIMA tertentu sehingga model ARIMA (2,1,2) sudah cukup layak. Berdasarkan hasil *Uji Ljung-Box* mengindikasikan autokorelasi sisaan tidak signifikan hingga 30 lag. Oleh karenanya model ARIMA(2,1,2) sudah cukup baik mengepas data produksi cengkeh Indonesia.

Tabel 6. Hasil *Uji Ljung-Box*

lags	statistic	df	p-value
5	2.722463	5	0.7426807
10	4.502952	10	0.9218197
15	10.039997	15	0.8172165
20	17.120482	20	0.6451369
25	19.539339	25	0.7705532
30	22.426322	30	0.8380904

Kemudian dilakukan uji coba peramalan dengan menggunakan data testing. Hasil peramalan data testing selama lima tahun kedepan yaitu tahun 2021 – 2025 tersaji di bawah ini

**Tabel 6.** Hasil Peramalan Data Testing

Time Series:

Start = 34

End = 38

Frequency = 1

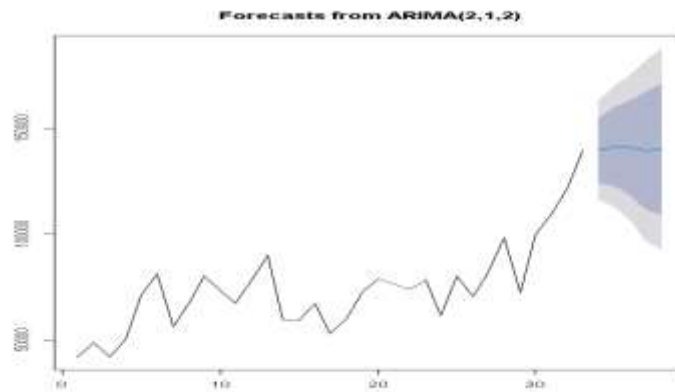
[1] 139877.3 141676.4 141044.4 139582.5 140489.8

accuracy(ramalan\_produksi,test[, "Produksi"])

Kemudian dilakukan peramalan untuk tahun 2021-2025

	ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
Training set	260.524	11526.91	9393.978	2.574737	<b>12.618782</b>	0.7638642	-0.06152918
Test set	-6410.471	13746.83	9107.567	-5.675658	<b>7.535093</b>	0.7405749	NA

Plot data dari tahun 1983-2014 (set data training) dan garis berwarna biru merupakan hasil peramalan data set testing (2015-2020), yang dilengkapi dengan arsiran selang kepercayaannya (daerah biru muda).



**Gambar 5.** Plot Ramalan Produksi Cengkeh Data Testing

Dengan model ARIMA (2,1,2) dilakukan pengepasan seluruh data dengan hasil sebagai berikut

**Tabel 7.** Hasil Pengepasan Seluruh Data

Call:

```
arima(x = data_produksi[, "Produksi"], order = c(2, 1, 2))
```

Coefficients:

```
      ar1    ar2    ma1    ma2
-0.1490 -0.8764 -0.1070 1.0000
s.e. 0.1149 0.0989 0.0851 0.1417
sigma^2 estimated as 154530267: log likelihood = -403.31, aic = 816.62
```

Training set error measures:

```
      ME      RMSE    MAE    MPE    MAPE    MASE    ACF1
Training set 3084.857 12266.37 9952.46 2.202629 12.45561 0.8131849 -0.06651474
```

Dari hasil pengepasan untuk seluruh data yaitu tahun 1983-2020 maka dihasilkan MAPE sebesar 12,45% Dengan menggunakan model ARIMA (2,1,2) kemudian dilakukan peramalan untuk tahun 2021-2025. Hasil peramalan dan plot dengan ARIMA (2,1,2) adalah sebagai berikut:

**Tabel 8..** Hasil Peramalan Model ARIMA (2,1,2)

Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95	
39	150296.0	133997.0	166595.0	125368.9	175223.2
40	156972.6	136632.9	177312.3	125865.7	188079.5
41	152228.	127511.1	176945.8	114426.5	190030.4
42	147084.1	116857.0	177311.2	100855.7	193312.5
43	152008.5	118155.2	185861.7	100234.4	203782.6

Dari hasil model autoarima dan 10 model tentative hasil *armaselect* dicobakan satu per satu dengan melihat mape terkecil dan kelayakan modelnya serta kelogisan hasil ramalan datanya. Setelah dicobakan satu persatu maka mape yang terkecil adalah ARIMA (2,1,2)

**Tabel 9.** Model Arima Berdasarkan Automodel

Call:

```
arima(x = train[, "Produksi"], order = c(2, 1, 2))
```

Coefficients:

```
ar1          ar2      ma1          ma2
-0.2338 -0.8947 -0.0557 0.9993
s.e. 0.1082 0.0912 0.1175 0.2651
```

sigma^2 estimated as 1.37e+08: log likelihood = -347.31, aic = 704.62

```
library(lmtest)
```

```
coefest(model1)
```

z test of coefficients:

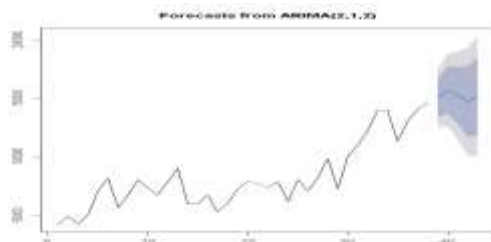
```
Estimate Std. Error z value Pr(>|z|)
ar1 -0.233805 0.108197 -2.1609 0.0307017 *
ar2 -0.894683 0.091228 -9.8071 < 2.2e-16 ***
ma1 -0.055684 0.117513 -0.4739 0.6356031
ma2 0.999275 0.265078 3.7697 0.0001634 ***
```

---

Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dari Hasil penelusuran model ARIMA (2,1,2) diatas menunjukkan parameter ar 2 dan ma 2 adalah sangat nyata. Kemudian dilihat sebaran sisaan LJung-Box dari lag 5 sampai lag 30. Salah satu syarat kebaikan model ARIMA adalah sebaran sisaan LJung-Box. Hasil pengujian nilai p-value pada lag 5 sampai dengan lag 30 tidak ada yang signifikan, sehingga dapat disimpulkan bahwa sisaan bersifat random dan tidak ada autorelasi. Hal ini menunjukkan bahwa ARIMA (2,1,2) layak digunakan. Hasil sisaan dengan Uji LJung-Box dan hasil plot sisaan disajikan pada gambar dan table dibawah inii

(2,1,2) adalah, pada tahun 2021 sebesar 150.296 ton, tahun 2022 sebesar 156.972 ton, tahun 2023 sebesar 152.228 ton, tahun 2024 sebesar 147.084 ton dan tahun 2025 sebesar 152.008 ton. Untuk plot estimasi produksi cengkeh tahun 2021-2025 seperti gambar 6.



**Gambar 6.** Plot Hasil Estimasi Tahun 2020-2025

Setelah dilakukan beberapa tahapan penelusuran model maka model ARIMA yang dipilih adalah ARIMA (1.1.2) dengan nilai AIC 787,77 dengan mape 12,85. Hasil estimasi produksi cengkeh tahun 2020-2024 adalah sebagai berikut:

**Tabel 9.** Hasil Estimasi Produksi Cengkeh 2021-2025 dengan ARIMA(2,1,2)

Tahun	Produksi (Ton)	Pertumbuhan (%)
2021	150,296	
2022	156,972	4.44
2023	152,228	-3.02
2024	147,084	-3.38
2025	152,008	3.35
Rata-rata pertumbuhan (%)		0.35
Mape Traini	12,62%	
Mape Testir	7,54%	

### Model Fungsi Transfer

Model fungsi transfer merupakan gabungan dari karakteristik analisis regresi berganda dengan karakteristik deret berkala ARIMA (Autoregressive Integrated Moving Average). Dalam model ini, selain mengandung keterkaitan dengan kejadian pada waktu-waktu sebelumnya, juga keterkaitan yang sifatnya kausal dengan variabel lain sebagaimana pada model regresi. Model fungsi transfer merupakan suatu model yang menggambarkan nilai prediksi dari suatu time series (deret output atau  $Y_t$ ) berdasarkan nilai-nilai dari deret itu sendiri ( $Y_t$ ) dan berdasarkan pula pada data time series yang mempunyai hubungan (deret input atau  $X_t$ ) dengan deret output. Model fungsi transfer yang bersifat dinamis berpengaruh tidak hanya pada hubungan linier antara waktu ke- $t$  input  $X_t$  dan waktu ke- $t$  output  $Y_t$ , tetapi juga saat input  $X_t$  dengan saat  $t, t+1, \dots, t+k$  pada output  $Y_t$  (Fitriani et al., 2013). Dalam melakukan estimasi dengan fungsi transfer diperlukan dua variabel, yaitu output  $Y_t$  dan input  $X_t$ . Input  $X_t$  merupakan variabel yang diduga paling berpengaruh terhadap output  $Y_t$ . Bentuk umum fungsi transfer single input ( $X_t$ ) dengan single output ( $Y_t$ ).  $Y_t = v(B) X_t + N_t$

Pada analisis ini estimasi untuk produksi cengkeh (output  $Y_t$ ) yang diduga dipengaruhi oleh harga cengkeh di tingkat produsen.

Tahap pembentukan model fungsi transfer untuk produksi cengkeh adalah sebagai berikut:

- Mempersiapkan deret input ( $X_t$ ) dan output ( $Y_t$ ) yang stasioner.
- Pembagian series data awal menjadi series data training dan testing
- Penentuan Model ARMA untuk deret input.
- Prewhitening deret input dan output

- Mendeteksi dan mengukur hubungan kekuatan antara  $\alpha_t$  dan  $\beta_t$  dengan menggunakan CCF (Cross Correlation Function).

- Penetapan  $(b, r, s)$  yang menghubungkan deret *input* dan deret *output* dimana  $b$  adalah panjang jeda

pengaruh  $X_t$  terhadap  $Y_t$ ,  $r$  adalah panjang lag  $Y$  periode sebelumnya yang masih mempengaruhi  $Y_t$ ,  $s$  adalah panjang lag  $X$  periode sebelumnya yang masih mempengaruhi  $Y_t$

- Penaksiran awal deret noise ( $\pi_t$ )

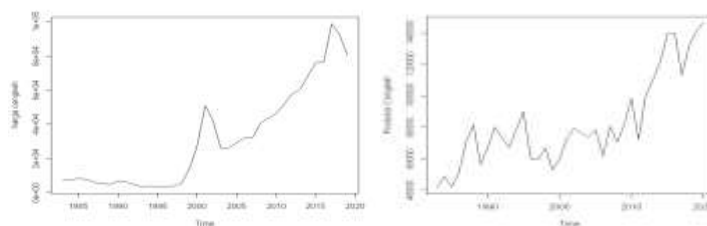
- Penentuan model ARIMA deret noise

- Estimasi parameter dengan fungsi transfer

- Uji kesesuaian model fungsi transfer

- Menggunakan model fungsi transfer untuk peramalan.

Eksplorasi data dilakukan dengan menampilkan plot data produksi dan harga cengkeh. Data produksi dan harga cengkeh dimulai tahun 1983-2020 yaitu sebanyak 38 series akan dibagi menjadi series data training untuk periode tahun 1983-2014 dan series data testing untuk periode tahun 2015-2020. Plot data produksi dan harga cengkeh dari tahun 1983-2020.



**Gambar 7.** Plot data harga Cengkeh tahun 1983-2020

Dari plot data harga cengkeh terlihat bahwa adanya peningkatan dari tahun ke tahun sehingga terindikasi data tidak stasioner. Untuk pengujiannya dapat dilakukan dengan uji Augment Dickey-Fuller. Hasil Augmen Dickey Fuller adalah sebagai berikut:

**Tabel 10. Uji Stasioner Diff 1**

```
#####
# Augmented Dickey-Fuller Test Unit Root Test #
#####
Test regression trend
Call:
lm(formula = z.diff ~ z.lag.1 + 1 + tt + z.diff.lag)
Residuals:
    Min     1Q   Median     3Q    Max
-12902.8 -4184.7 -653.6  2357.0 23577.0

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -4222.3391  3059.1298  -1.380  0.17708
z.lag.1      -0.3063    0.1034  -2.963  0.00571 **
tt           768.6126  279.5474   2.749  0.00973 **
z.diff.lag   0.3810    0.1657   2.300  0.02814 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 7135 on 32 degrees of freedom
Multiple R-squared:  0.2652,    Adjusted R-squared:  0.1963
F-statistic: 3.849 on 3 and 32 DF,  p-value: 0.0185
Value of test-statistic is: -2.9626 3.2291 4.397
Critical values for test statistics:
    1pct 5pct 10pct
tau3 -4.15 -3.50 -3.18
phi2  7.02  5.13  4.31
phi3  9.31  6.73  5.61
```

Dari hasil uji Augmented Dickey-Fuller terlihat bahwa data tidak stasioner karena nilai *tes statistic* lebih besar dari nilai *critical* dimana nilai *tes statistic nya* adalah **-2,9626** dan nilai *critical tau 3* untuk 1% sebesar **-4,15**, taraf uji 5% sebesar **-3,50** dan taraf uji 10% sebesar **-3,18**. Dari hasil uji Augmented Dickey-Fuller ini dapat disimpulkan data tidak stasioner hingga pada taraf 1%, 5% dan 10% yang berarti tolak  $H_0$ . Untuk menstasionerkan data sehingga dilakukan differencing 1. Hasil uji Augmented Dickey-Fuller Diff 1 adalah sebagai berikut:

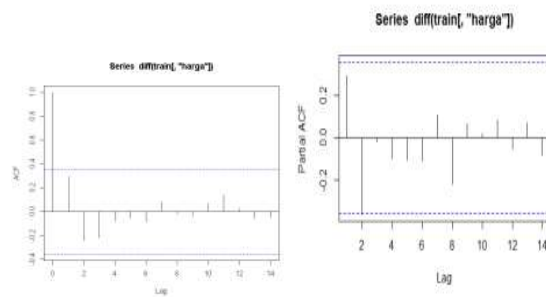


**Tabel 11.** Uji Stasioner Diff 2

```
#####  
# Augmented Dickey-Fuller Test Unit Root Test #  
#####  
Test regression none  
Call:  
lm(formula = z.diff ~ z.lag.1 - 1 + z.diff.lag)  
Residuals:  
  Min   1Q Median   3Q   Max  
-28072 -1789  422  3249 19163  
Coefficients:  
      Estimate Std. Error t value Pr(>|t|)  
z.lag.1  -1.6524   0.2662  -6.208 5.95e-07 ***  
z.diff.lag  0.3197   0.1686   1.896  0.0671 .  
---  
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1  
Residual standard error: 9148 on 32 degrees of freedom  
Multiple R-squared:  0.664,  
Adjusted R-squared:  0.643  
F-statistic:31.62 on 2 and32 DF, p-value: 2.638e-08  
Value of test-statistic is: -6.2079  
Critical values for test statistics:  
  1pct  5pct 10pct  
tau1 -2.62 -1.95 -1.61
```

Dari hasil uji Augmented Dickey-Fuller setelah di differencing 1 kali terlihat bahwa nilai *tes statistic* sudah lebih kecil dibandingkan nilai *critical*ya dimana nilai *tes statistic* adalah -6.2079 sementara nilai critical tau 1 untuk 1% sebesar -2,62, taraf uji 5% sebesar -1,95 dan taraf uji 10% sebesar -1,61%. Hal ini dapat disimpulkan data sudah stasioner hingga pada taraf uji 1%, 5% dan 10%.

Stasioneritas data dapat juga dilihat dari plot ACF dan PACF. Jika lag pada ACF maupun PACF menunjukkan *tail off* maka data tidak stasioner. Berikut adalah plot ACF dan PACF harga cengkeh setelah differencing 1.



**Gambar 8.** Plot ACF dan PACF Harga Cengkeh

Berdasarkan plot di atas, terlihat bahwa plot ACF cenderung cut off dimana hanya lag 0 yang keluar dari *confidence interval*. Plot PACF menunjukkan cut off pada lag 2, ditunjukkan berada di luar garis *confidence interval*. Hal tersebut memperkuat hasil uji ADF yaitu data harga cengkeh telah stasioner.

#### Pencarian Model Tentatif Variabel Input

Untuk mencari model tentatif variable input (harga) dapat dilakukan dengan penelusuran menggunakan model ARIMA. Model terbaik dapat dipilih melalui *autoarima* yang tersedia pada R Studio maupun *arma selec*. Data yang digunakan untuk memilih model terbaik adalah series data trainingnya. Hasil dari *autoarima* adalah sebagai berikut:

**Tabel 12.** Hasil Model *Autoarima*

```
Series: train[, "harga"]
ARIMA(0,1,1)
Coefficients:
      ma1
      0.5051
s.e. 0.1682
sigma^2 estimated as 38826306: log likelihood=-314.48
AIC=632.97 AICc=633.39 BIC=635.83
Training set error measures:
      ME      RMSE      MAE      MPE      MAPE      MASE      ACF1
Trainingset1340.543 6033.213 3532.181 3.339486 15.07852 0.8722615 -0.05599509
```

Hasil *autoarima* untuk harga cengkeh adalah ARIMA (0,1,1) dengan Standard error sebesar 0,1682, nilai MAPE data training sebesar 15.07% dan nilai AIC 632,97 seperti terlihat pada tabel diatas.

Selanjutnya dilakukan pengujian *coefst* pada model ARIMA(0,1,1) untuk mengetahui signifikansi dari koefisien MA pada model tersebut, output *coefst*nya adalah :

**Tabel 13.** Hasil *Coefst* untuk Arima ( 0,1,1)

```
Call:
arima(x = train[, "harga"], order = c(0, 1, 1))
Coefficients:
      ma1
      0.5051
s.e. 0.1682
sigma^2 estimated as 37573843: log likelihood = -314.48, aic = 632.97
z test of coefficients:
      Estimate Std. Error z value Pr(>|z|)
ma1  0.50513   0.16825  3.0023 0.002679 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Berdasarkan output di atas terlihat bahwa komponen *ma1* signifikan, pada taraf 1%, 5% dan 10% dengan *mape* 15,18%. Selain menggunakan *script auto arima* model tentatif dapat juga dijalankan dengan perintah *armaselect*. R-Studio akan menunjukkan sepuluh (10) model tentatif yang diurutkan berdasarkan nilai AIC terkecil.dari ordo  $p=1$  s/d 5 dan  $q=1$  s/d 5. Sepuluh model tentatif yang ditampilkan pada lag 1 untuk *arma select* adalah sebagai berikut:

**Tabel 14.** Hasil Model *Armaselect Diff1*

	p	q	sbc
[1,]	5	5	453.8245
[2,]	1	5	513.4337
[3,]	3	5	514.1288
[4,]	2	5	515.2249
[5,]	4	5	515.7160
[6,]	0	5	517.6856
[7,]	3	4	526.7095
[8,]	4	4	527.3080
[9,]	5	4	530.7385
[10,]	2	4	534.4641

Model terbaik adalah model yang memiliki nilai SBC terkecil. Model pertama yang disarankan adalah adalah nilai  $p=5$  dan  $q=5$ . Karena data harga cengkeh di differencing satu kali berarti  $d=1$ , artinya model yang direkomendasikan adalah ARIMA (5,1,5). Model terbaik

kedua adalah ARIMA (1,1,5) sampai model terbaik ke sepuluh. Untuk mengetahui model terbaik dapat dilihat dari nilai AIC, mape dan koefisien ar dan ma nya nyata. Meskipun ARIMA (5,1,5) diurutan pertama dengan sbc yang lebih kecil, belum tentu merupakan terbaik sehingga perlu di ujicobakan semua model yang disajikan oleh *armaselect*.

Setelah di uji cobakan dari 10 model tentatif diatas, mape data testing yang didapatkan diatas 10% sehingga dicobakan untuk lag 2. Model tentatif dari hasil *armaselect* untuk lag 2 adalah:

**Tabel 15.** Hasil Model Armaselect Diff2

	p	q	sbc
[1,]	5	5	346.0552
[2,]	3	5	448.7169
[3,]	4	5	452.0989
[4,]	2	5	463.9719
[5,]	5	4	479.0701
[6,]	1	5	490.6346
[7,]	3	4	494.7115
[8,]	4	4	497.6300
[9,]	2	4	502.0853
[10,]	1	3	506.5748

Dari 10 model tentatif *armaselect* lag 2 yang telah diuji cobakan maka ARIMA yang dipilih adalah ARIMA (1,2,3)

**Tabel 16.** Hasil Coeffttest Arima (1,2,3)

Call:

```
arima(x = train[, "harga"], order = c(1, 2, 3))
```

Coefficients:

```

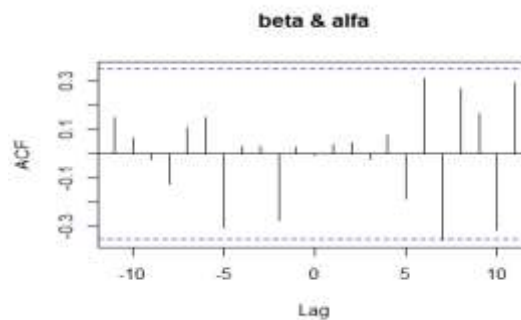
ar1    ma1    ma2    ma3
0.7048 -1.3891 -0.1349 0.5771
s.e. 0.1827 0.2493 0.3494 0.2019
sigma^2 estimated as 31896033: log likelihood = -304.12, aic = 618.23
z test of coefficients:
```

```

Estimate Std. Error z value Pr(>|z|)
ar1 0.70480 0.18270 3.8577 0.0001145 ***
ma1 -1.38905 0.24926 -5.5726 2.509e-08 ***
ma2 -0.13486 0.34937 -0.3860 0.6994874
ma3 0.57707 0.20191 2.8581 0.0042616 **
```

Dari hasil z test arima (1,2,3) terlihat ordo ar 1 dan ma 1 signifikan pada taraf 1%, 5% dan 10%, sementara untuk ordo ma 3 signifikan pada taraf 1% dan 5%. Oleh karena ARIMA (1,2,2) memiliki AIC yang lebih kecil maka selanjutnya dapat digunakan untuk peramalan ptduksi cengkeh.

Tahapan selanjutnya untuk penyusunan model fungsi transfer ini adalah dengan prewhitening dan korelasi silang. Korelasi silang menggambarkan struktur hubungan antara  $X_t$  dengan  $Y_t$ . Untuk mengidentifikasi pengaruh  $X_t$  terhadap  $Y_t$  maka deret  $X_t$  harus stasioner atau sudah distasionerkan. Dalam konteks pemodelan  $X_t$  terhadap  $Y_t$ , untuk membuat  $X_t$  stasioner tidak dengan pembedaan (*differencing*) namun dengan mengambil komponen *white noise* dari  $X_t$  (*prewhitening*). *Prewhitening* dilakukan terhadap deret input  $X_t$  yang didefinisikan sebagai alfa serta deret input  $Y_t$  yang didefinisikan sebagai beta. Hasil grafik ACF untuk beta dan alfa.



**Gambar 9.** Plot Korelasi Silang

Berdasarkan grafik ACF di atas, terlihat bahwa tidak ada yang keluar dari garis signifikansi,. Karena lag signifikan pertama kali pada lag 0 maka nilai  $b=0$ . Selain itu, tidak ada tambahan lagi lag yang signifikan maka nilai  $s=0$ . Mengingat data ptduksi Cengkeh dan harga Cengkeh merupakan data tahunan yang tidak mengandung musiman maka diasumsikan nilai  $r=0$ . Nilai  $b=0$  menunjukkan tidak ada jeda pengaruh antara harga pada waktu  $t$  terhadap ptduksi pada waktu  $t$ . Nilai  $s=0$  berarti ada korelasi antara ptduksi dan harga cengkeh pada tahun yang sama. Dengan kata lain, dampak dari harga cengkeh terhadap produksi dirasakan pada waktu yang sama ( $t$ ).

Tahap selanjutnya dilakukan pengepasan model, untuk nilai  $r,s$  dan  $b$ . Hasil pengujian fungsi transfer dengan nilai  $r=0, s=0$ , dan  $b=0$  menghasilkan nilai MAPE yaitu 14,36%

**Tabel 17.** Pengujian Ordo r,s dan b ARIMA (0,0,0)

Series: train[, "Produksi"]

Regression with ARIMA(0,0,0) errors

Coefficients:

intercept xreg

59245.692 0.5781

s.e. 3631.412 0.1181

sigma<sup>2</sup> estimated as 195607937: log likelihood=-349.84

AIC=705.68 AICc=706.54 BIC=710.08

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set -1.477929e-11 13541.88 11312.4 -3.814881 **16.84568** 0.9326015 0.3713585

Untuk menghasilkan orde yang paling tepat dan menentukan orde arima fungsi transfer dengan melakukan identifikasi model noise. Untuk menghasilkan model terbaik dengan menggunakan auto-arima pada R Studio, model noise yang disarankan adalah Arima (1,0,0). Model ini ternyata masih kurang tepat, karena menghasilkan MAPE yang cukup besar yaitu 132,58%.

**Tabel 18.** Model Residual Tetntatif Autoarima

ARIMA(1,0,0) with zero mean

Coefficients:

ar1

0.4329

s.e. 0.1727

sigma<sup>2</sup> estimated as 157728286: log likelihood=-347.02

AIC=698.05 AICc=698.46 BIC=700.98

Training set error measures:

ME RMSE MAE MPE MAPE MASE ACF1

Training set 382.4049 12361.2 10619.11 122.4435 **132.5826** 0.8957333 -0.04023018

Dengan autoarima disarankan ARIMA (1,0,0) dengan mape 132,58%. Karena mape nya sangat tinggi sehingga di cari model alternated dengan *arma selec*. Hasil armaselect residual adalah:

**Tabel 19.** Model Residual Tetntatif

	p	q	sbc
[1,]	5	5	524.2167
[2,]	1	4	583.9228
[3,]	2	4	586.8061
[4,]	1	5	589.6450
[5,]	3	4	590.1632
[6,]	3	5	590.4835
[7,]	1	1	590.9206
[8,]	1	2	591.2835
[9,]	2	1	591.2929
[10,]	0	4	591.8627

Setelah diujicobakan model tentatife yang disarankan oleh armaselect, maka model arima uang dipilih adalah ARIMA (1,1,2) dengan hasil sebagai berikut:

**Tabel 20.** Model Residual ARIMA(1,1,2)

```
Series: res
ARIMA(1,1,2)
Coefficients:
      ar1      ma1      ma2
    -0.6424  0.4484 -0.5510
s.e.  0.1709  0.1897  0.1756
sigma^2 estimated as 155850093: log likelihood=-335.88
AIC=679.76  AICc=681.3  BIC=685.5
z test of coefficients:
Estimate Std. Error z value Pr(>|z|)
ar1 -0.64238  0.17092 -3.7584 0.000171 ***
ma1  0.44843  0.18973  2.3636 0.018100 *
ma2 -0.55099  0.17563 -3.1373 0.001705 **
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Dari hasil z test arima (1,1,2) terlihat ordo ar 1 signifikan pada taraf 1%, 5% dan 10%, ordo ma 1 signifikan pada taraf 5% dan ordo ma 2 signifikan pada taraf 10% dan 5%. Oleh karena ARIMA (1,1,2) memiliki AIC yang lebih kecil maka selanjutnya digunakan untuk deret *noise*.

Tahap selanjutnya dilakukan pengepasan model dengan noise. Pengepasan model dengan menggabungkan nilai r,s dan b dan hasil identifikasi model noise yaitu ARIMA (1,1,2) dengan hasil sebagai berikut:

**Tabel 21.** Pengepasan Model ASRIMA (1,12)

Series: train[, "Produksi"]

Regression with ARIMA(1,1,2) errors

Coefficients:

ar1	ma1	ma2	xreg	
-0.6497	0.4605	-0.5394	0.4198	
s.e.	0.1701	0.1909	0.1767	0.2842

sigma^2 estimated as 160101200: log likelihood=-335.72  
AIC=681.45 AICc=683.85 BIC=688.62

Training set error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
----	------	-----	-----	------	------	------

Training set 2361.085 10995.93 8796.638 1.763051 **12.2849** 0.7252007 -0.2004686z test of coefficients:

Kemudian dilakukan pengujian coeftest untuk mengetahui tingkat signifikansi dari masing-masing koefisien, seperti terlihat pada dibawah ini

**Tabel 22.** Pengujian Koefisien Model Fungsi Transfer ARIMA (1,1,2)

	Estimate	Std. Error	z value	Pr(> z )
ar1	-0.64968	0.17011	-3.8191	0.0001339 ***
ma1	0.46053	0.19095	2.4119	0.0158713 *
ma2	-0.53936	0.17674	-3.0517	0.0022753 **
xreg	0.41983	0.28420	1.4772	0.1396137

---  
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Dari hasil pengujian di atas, ditemukan bahwa komponen AR1 dan ma 2 signifikan akan tetapi komponen fungsi transfer xreg tidak signifikan. Dari hasil pengepasan model noise dengan ARIMA (1,1,2) diperoleh mape 12,282% dan nilai AIC =681,45. Kemudian ar 1, ma 1 dan ma 2 signifikan pada selang kepercayaan 95%,

**Peramalan berbasis fungsi transfer, data Input nilai Aktual**

Berdasarkan model fungsi transfer dengan model noise ARIMA (1, 1, 2), dilakukan peramalan berbasis nilai aktual dimana produksi cengkeh diestimasi menggunakan data aktual harga cengkeh periode 2015-2020. Meskipun data aktual produksi cengkeh periode 2015-2020 telah ada, dilakukan peramalan produksi cengkeh untuk mengecek *performance/kelayakan* model fungsi transfer

Pengujian MAPE model fungsi transfer ARIMA (1,1,2) dengan factor input harga cengkeh dan Data Aktual tahun 2015-2020



### Tabel 23. Fungsi Transfer Nilai Aktual

Series: test[, "Produksi"]

Regression with ARIMA(2,1,2) errors

Coefficients:

	ar1	ar2	ma1	ma2	xreg
	-0.2733	-0.9405	0.0179	0.9998	0.1711
s.e.	0.0000	0.0000	0.0000	0.0000	0.0000

sigma^2 estimated as 148812944: log likelihood=-55.9  
AIC=113.81 AICc=115.14 BIC=113.42

Training set error measures:

	ME	RMSE	MAE	MPE	<b>MAPE</b>	MASE	ACF1
Training set	1574.098	13191.34	9671.702	0.4170526	<b>7.565361</b>	0.815448	-0.01654499

Series: test[, "Produksi"]

Regression with ARIMA(2,1,2) errors

Coefficients:

	ar1	ar2	ma1	ma2	xreg
	-0.2733	-0.9405	0.0179	0.9998	0.1711
s.e.	0.0000	0.0000	0.0000	0.0000	0.0000

sigma^2 estimated as 148812944: log likelihood=-55.9  
AIC=113.81 AICc=115.14 BIC=113.42

Training set error measures:

	ME	RMSE	MAE	MPE	<b>MAPE</b>	MASE	ACF1
Training set	1574.098	13191.34	9671.702	0.4170526	<b>7.565361</b>	0.815448	-0.01654499

Regression with ARIMA(2,1,2) errors

Coefficients:

	ar1	ar2	ma1	ma2	xreg
	-0.2733	-0.9405	0.0179	0.9998	0.1711
s.e.	0.0000	0.0000	0.0000	0.0000	0.0000

sigma^2 estimated as 148812944: log likelihood=-55.9  
AIC=113.81 AICc=115.14 BIC=113.42

Training set error measures:

	ME	RMSE	MAE	MPE	<b>MAPE</b>	MASE	ACF1
Training set	1574.098	13191.34	9671.702	0.4170526	<b>7.565361</b>	0.815448	-0.01654499

Peramalan produksi cengkeh periode 2015-2020 menggunakan fungsi transfer ARIMA (1,1,2) dengan input harga cengkeh **data aktual** menghasilkan MAPE 7,56%.

## Tabel 24. Fungsi Transfer Nilai Ramalan

Series: test[, "Produksi"]

Regression with ARIMA(2,1,2) errors

Coefficients:

ar1	ar2	ma1	ma2	xreg
-0.2733	-0.9405	0.0179	0.9998	0.1711
s.e.	0.0000	0.0000	0.0000	0.0000

sigma<sup>2</sup> estimated as 148812944: log likelihood=-55.18  
AIC=112.36 AICc=113.69 BIC=111.97

Training set error measures:

ME	RMSE	MAE	MPE	MAPE	MASE	ACF1
1442.507	11408.22	8449.071	0.4304519	6.621199	0.7123645	-0.06405457

Tujuan melakukan pemodelan fungsi transfer adalah untuk mendapatkan nilai ramalan periode ke depan, yakni produksi cengkeh tahun 2021-2025. Karena data series input harga cengkeh tersedia hingga tahun 2020, maka perlu dilakukan peramalan harga cengkeh terlebih dahulu atau dengan kata lain peramalan produksi dilakukan berbasis nilai ramalan harga cengkeh.

Oleh karenanya, untuk pengujian data testing terlebih dahulu dilakukan estimasi harga cengkeh periode 2015-2020 menggunakan model ARIMA (1,2,3) sebagaimana yang telah diperoleh dari tahap pencarian model tentatif untuk variabel input, sebagai variabel input harga cengkeh. Pemilihan variabel input harga cengkeh karena secara tidak langsung berpengaruh juga pada produksi cengkeh nasional. Selanjutnya dilakukan peramalan produksi cengkeh dengan fungsi transfer ARIMA (1,1,2) sebagai model terbaik berdasarkan tahapan pengepasan model dengan noise. Peramalan produksi dengan fungsi transfer ARIMA (1,1,2) menggunakan nilai ramalan harga cengkeh yang telah diestimasi dengan ARIMA (1,1,2).

### *Pengepasan Model Arima Output*

Setelah dilakukan peramalan produksi cengkeh baik menggunakan input yaitu harga produsen data aktual maupun ramalan, tahapan berikutnya adalah pengepasan model arima output. Pengepasan model ARIMA output dimaksudkan untuk membandingkan hasil ramalan produksi baik berdasarkan data training (1983-2014) maupun data testing (2015-2020). Hasil Pengepasan model

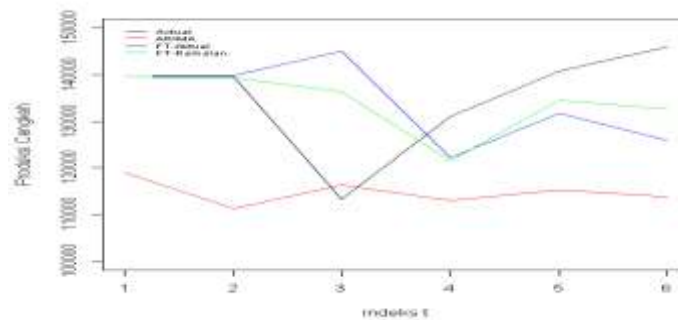
## Tabel 25. Pengepasan Model

```

Series: train[, "Produksi"]
ARIMA(1,1,2)
Coefficients:
    ar1    ma1    ma2
-0.6601  0.5310 -0.4690
s.e.  0.1705  0.1905  0.1746
sigma^2 estimated as 1.65e+08: log likelihood=-336.71
AIC=681.42  AICc=682.96  BIC=687.15
Training set error measures:
            ME  RMSE  MAE  MPE  MAPE  MASE  ACF1
Training set 3513.664 12016.52 10179.59 2.917689 14.03883 0.8392123 -0.1449769
Test set    20341.840 23365.08 21383.23 14.429246 15.34938 1.7628475    NA

```

Untuk membandingkan ketepatan model estimasi, dilakukan perbandingan hasil estimasi terhadap data aktual pada tahun 2015-2020 (data testing). Hasil ramalan dibandingkan dalam sebuah grafik seperti Gambar dibawah



**Gambar 10.** Perbandingan Hasil Ramalan

Setelah dicari model terbaik untuk meramalkan produksi cengkeh makan akan diestimasi produksi cengkeh lima tahun ke depan (2021-2025). Hasil ramalan lima tahun ke depan:

Tabel 26. Hasil Estimasi Tahun 2021-2025

Tahun	Produksi Pertumbuhan	
	(Ton)	(%)
2021	157,425	
2022	153,954	-2.20
2023	163,804	6.40
2024	164,589	0.48
2025	171,545	4.23
Rata-rata pertumbuhan (%)		2.22
Mape Training	14.04%	
Mape Testing	15,35%	

## Model VAR

Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini tidak perlu menentukan mana variabel endogen dan variabel eksogen karena semua variabel merupakan variabel endogen. Dalam pemodelan produksi cengkeh menggunakan tiga variabel yaitu produksi cengkeh, harga cengkeh dan volume ekspor cengkeh dengan series tahun 1983-2020 dengan format data CSV

Tahapan-tahapan penyusunan VAR adalah : persiapan data, pembagian data training dan testing, pemilihan lag dan type, pengajuan asumsi, ramalan data training, testing, penghitungan MAPE, dan plot, pemilihan model terbaik, dan pengepasan model untuk seluruh data dan peramalannya. Terakhir menggunakan *Impulse Response Function* (IRF) untuk menelusuri guncangan atau shock suatu variabel terhadap variabel lainnya.

Series data yang digunakan mulai tahun 1983-2000. Data ini akan dibagi menjadi 2 set yaitu set data training mulai tahun 1983 - 2014 dan set data testing mulai tahun 2015 -2020. Untuk pemilihan lag ( $p$ ) dan bentuk type yang digunakan perlu dilakukan pengujian. Jika keduanya signifikan, maka komponen tersebut termasuk ke type "both". Jika konstantanya yang signifikan maka termasuk model VAR type "const". Jika keduanya tidak signifikan maka type yang digunakan adalah type "non". Untuk pemilihan lag  $p$  nya dilakukan uji coba, trial and error dengan menggunakan model VAR dari lag  $p=1$  s.d 5 dengan type "both" untuk mengetahui signifikan atau tidak keberadaannya. Setelah di uji cobakan dari lag  $p=1$  s-5 baik untuk type cons, both dan non maka ada 2 type yang dipilih dengan memperhatikan signifikannya yaitu  $p=1$  type "both"

Tabel 27. Output VAR p=1 type “both” :

VAR Estimation Results:

```

=====
Endogenous variables: Produksi, harga, volek
Deterministic variables: both
Sample size: 32
Log Likelihood: -975.533
Roots of the characteristic polynomial:
0.7232 0.7232 0.002026
Call:
VAR(y = cengkeh[1:33, c(2, 3, 4)], p = 1, type = "both")
Estimation results for equation Produksi:
=====
Produksi = Produksi.l1 + harga.l1 + volek.l1 + const + trend
      Estimate Std. Error t value Pr(>|t|)
Produksi.l1  0.4906  0.1687  2.909 0.00717 **
harga.l1     0.4727  0.2179  2.169 0.03904 *
volek.l1     -0.9431  0.5234 -1.802 0.08273 .
const       29408.0389 9991.1534  2.943 0.00660 **
trend        247.6390  542.9977  0.456 0.65199
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 12130 on 27 degrees of freedom
Multiple R-Squared:  0.7062,    Adjusted R-squared: 0.6626
F-statistic: 16.22 on 4 and 27 DF,  p-value: 6.954e-07
Estimation results for equation harga:
=====
harga = Produksi.l1 + harga.l1 + volek.l1 + const + trend
      Estimate Std. Error t value Pr(>|t|)
Produksi.l1 -0.07974  0.08865 -0.900 0.3763
harga.l1     0.83875  0.11455  7.322 7.08e-08 ***
volek.l1     -0.09951  0.27510 -0.362 0.7204
const        741.74381 5251.47728  0.141 0.8887
trend        651.06672 285.40648  2.281 0.0306 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
Residual standard error: 6373 on 27 degrees of freedom
Multiple R-Squared:  0.9295,    Adjusted R-squared: 0.9191
F-statistic:  89 on 4 and 27 DF,  p-value: 3.819e-15
Estimation results for equation volek:
=====

```

### Pengujian Asumsi

Selanjutnya dilakukan pengujian asumsi p=1 type both”.untuk non autokorelasi, normalitas, dan homoskedastisitas pada sisaan model VAR p=1 type “both”

- Pemeriksaan autokorelasi residual model menggunakan fungsi “serial.test” yang di dalamnya dilakukan pengujian Portmanteau-and Breusch-Godfrey test. Hasil pemeriksaan autokorelasi adalah

Portmanteau Test (asymptotic)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 103.67, df = 135, p-value = 0.9792

Nilai p-value lebih besar dari 0,05 maka sisaan saling bebas atau asumsi non autokorelasi terpenuhi.

Pengujian Jarque-Bera tests untuk menguji kenormalan.

JB-Test (multivariate)

-

data: Residuals of VAR object varhsheet.b1

Chi-squared = 109.76, df = 6, p-value < 2.2e-16

\$Skewness

Skewness only (multivariate)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 27.351, df = 3, p-value = 4.97e-06

\$Kurtosis

Kurtosis only (multivariate)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 82.414, df = 3, p-value < 2.2e-16

hasil pengujian menunjukkan Nilai p-value lebih kecil dari 0,05, namun karena jumlah data yang digunakan cukup banyak, maka series tersebut dapat dianggap normal.

- Pemeriksaan heteroskedastisitas model menggunakan fungsi “arch.test” yang di dalamnya dilakukan pengujian ARCH- LM tests .

ARCH (multivariate)

data: Residuals of VAR object varhsheet.b1

Chi-squared = 162, df = 180, p-value = 0.8281

Nilai p-value lebih besar dari 0,05 maka ragam sisaan homogen atau asumsi homoskedastisitas terpenuhi.

Selanjutnya dilakukan peramalan data, baik untuk data training maupun untuk data testing sekaligus dilakukan penghitungan MAPE. MAPE (Mean Absolute Percentage Error) merupakan rata-rata dari keseluruhan persentase kesalahan (selisih) antara data aktual dengan

data hasil peramalan. Untuk menguji suatu model lebih baik dengan model yang lain, maka dilakukan pengujian model dengan membandingkan Nilai MAPE baik untuk data training maupun data testing. Untuk MAPE VAR p=1 type both

Tabel 28

Mape data testing

Min. 1st Qu. Median Mean 3rd Qu. Max.

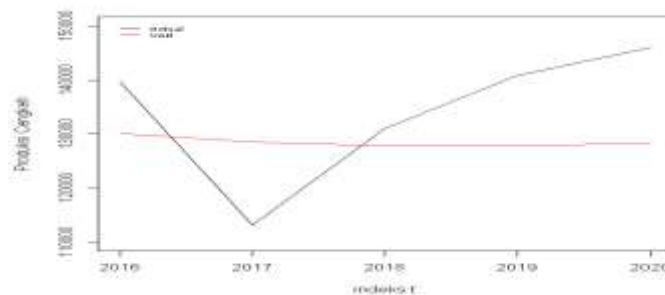
2.441 6.838 9.246 8.845 12.111 13.586

Min. 1st Qu. Median Mean 3rd Qu. Max.

0.2567 5.2827 9.6832 12.5374 17.0834 40.0469

Nilai ini dapat diartikan bahwa rata-rata seluruh persentase kesalahan antara data actual dengan data hasil ramalan pada data testing adalah 12,70%

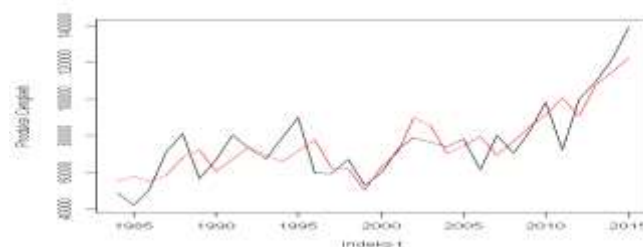
Kemudian dilakukan plot data ramalan dan data actual tahun 2016-2020 atau plot data testing



Gambar 11. Plot Ramalan dan Aktual Data Testing

Dari grafik plot di atas dapat dilihat bahwa pergerakan ramalan pada data testing mengikuti pergerakan data aktual. Sehingga model VAR p=1 type “both” merupakan model terbaik.

Plot data ramalan dan data actual tahun 1983-2015 atau plot data training



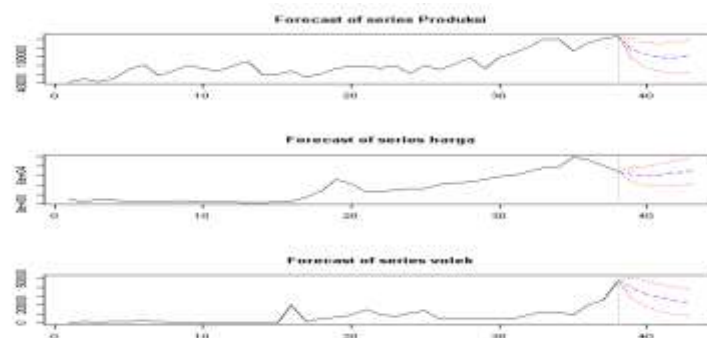
Gambar 12. Plot Ramalan dan Aktual Seluruh Data

Dari grafik plot di atas untuk plot ramalan dan aktual data training dapat dilihat bahwa pergerakan ramalan pada data testing mengikuti pergerakan data aktual.

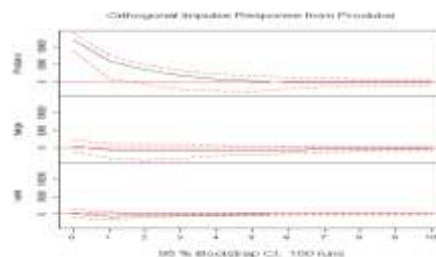
Kemudian dilakukan ramalan untuk tahun 2021-2025.

Tabel 29. Hasil ramalan tahun 2021-2025

Produksi cengkeh		
Tahun	Produksi (Ton)	Pertumbuhan (%)
2021	147,469	
2022	142,010	-3.16
2023	137,501	-3.72
2024	133,034	-3.25
2025	127,815	-3.92
Rata-rata pertumbuhan (%)		-3.51
Mape Training	12,54%	
Mape Testing	8,85%	



Gambar 13. Plot Ramalan Model VAR p=1 both



Gambar 14. Plot Orthogonal Impulse Response from Produksi

Impulse Response Function akan menjelaskan bagaimana perubahan atau gejolak yang terjadi pada suatu variabel/peubah di tahun tertentu akan berdampak pada variabel/peubah lain di tahun tertentu dan tahun-tahun setelahnya.

Dari grafik Impulse Response Function produksi model terbaik VAR (1) type “both”, dapat dilihat bahwa jika terjadi apabila ada perubahan produksi cengkeh Indonesia pada tahun tertentu, maka produksi cengkeh akan merespon hingga 3 tahun ke depan.

#### Pemilihan Model Terbaik Estimasi Produksi Cengkeh



Salah satu dasar penentuan model terbaik dapat dilakukan dengan membandingkan nilai MAPE untuk data testing dan training yaitu dengan memilih nilai MAPE yang paling kecil, terutama untuk data testing. Selain MAPE yang terkecil, pola pergerakan ramalan juga harus diperhatikan. Dipilih plot yang paling berhimpit/bersesuaian dengan data asli/aktual atau dengan kata lain performa hasil ramalan seiring dengan data historisnya.

Tabel 30. Perbandingan MAPE Model Arima, F Transfer dan VAR

No	Uraian Model	Mape		Hasil Estimasi Luas Areal Cengkeh (Ha)					Pertumb. %
		Testing	Training	2021	2022	2023	2024	2025	
1	ARIMA (2,1,2)	7,54	12,62	150296	156972	152228	147084	152006	0,35
2	Fungsi Transfer	12,49	15,74	157425	153954	163804	164589	171545	2,22
3	VAR	8,85	12,54	147469	142810	137501	133034	127815	-3,51

### KESIMPULAN

Jika dilihat dari mape terkecil maka ada 2 model terbaik yaitu ARIMA (2,1,2) dan VAR (1) type both. Jika dilihat dari histori datanya selama 5 tahun terakhir, produksi cengkeh mengalami peningkatan 1,08% per tahun. Sementara produksi cengkeh dari hasil model Var lima tahunkedepan yaitu tahun 2021-2025 turun 3,51 per tahun sehingga dari 4 model yang diuji cobakan model yang dipilih adalah ARIMA (2,1,2). Model Arima ini ini menghasilkan MAPE untuk data training 12,62% dan MAPE untuk data testing 7,54%. Jika dilihat dari histori datanya 5 tahun kedepan yaitu tahun 2021-2025 mengalami peningkatan 0,35% dan MAPE untuk data testing paling kecil dibandingkan dengan model-model lainnya, sehingga model **ARIMA (2,1,2)** disimpulkan paling kecil kesalahan estimasinya atau model terbaik dibandingkan model fungsi transfer dan model VAR.

### DAFTAR PUSTAKA

- Kementerian Pertanian. 2013. Pedoman Pelaksanaan Pengelolaan Data Komoditas Perkebunan (PDKP). Dirjen Perkebunan – Kementerian Pertanian
- Kementerian Pertanian. 2021. Statistik Perkebunan Indonesia 2019-2021 (Cengkeh). Jakarta: Direktorat Jenderal Perkebunan
- Makridakis. (1999). Metode dan Aplikasi Peramalan. Jakarta: Binarupa Aksara.



# KAJIAN MODEL ESTIMASI PRODUKSI LADA DI INDONESIA

## *Model Study for Estimating Pepper Production in Indonesia*

Roydatul Zikria<sup>1\*</sup>

<sup>1</sup>*Pusat Data dan Sistem Informasi Pertanian, Kementerian Pertanian*

*Jl. Harsono RM No. 3 Gedung D Lantai 4, Jakarta Selatan, Indonesia*

\*Korespondensi penulis. E-mail: roydatul\_z@pertanian.go.id

### ABSTRACT

Indonesian pepper production has fluctuated for the last five years. Estimation of pepper production in the next few years uses ARIMA, transfer function and VAR model. Regarding estimation of those models, VAR is selected as the best model to estimate pepper production. In detail, the model is VAR(1) with none type using area, export quantity, and import quantity as input variables. VAR model produces MAPE for training data and testing data which are respectively 9.09 and 4.81. Estimation of pepper production for the next five years (2021-2025) increases by 1.24% each year. The estimation of pepper production in 2021 is 88,103 ton while in 2025 is estimated 92,563 ton.

**Keywords:** *estimation, ARIMA, transfer function, VAR*

### ABSTRAK

Produksi lada di Indonesia cenderung berfluktuasi selama lima tahun terakhir. Untuk mengestimasi produksi lada beberapa tahun ke depan, penelitian ini menggunakan model ARIMA, fungsi transfer dan VAR. Hasil estimasi dari ketiga model tersebut dibandingkan dan terpilih VAR sebagai model estimasi terbaik. Model VAR yang terpilih adalah VAR(1) *type none* dengan input variabel yaitu luas areal, volume ekspor, dan volume impor. Hasil estimasi dengan model VAR menghasilkan MAPE data training sebesar 9,09 dan MAPE data testing 4,81. Produksi lada lima tahun ke depan (2021-2025) diramalkan meningkat 1,24% per tahun. Tahun 2021 produksi lada diramalkan sebesar 88.103 ton kemudian meningkat di tahun 2025 menjadi 92.563 ton.

**Kata kunci:** *estimasi, ARIMA, fungsi transfer, VAR*

## PENDAHULUAN

Salah satu komoditas perkebunan yang berperan penting dalam perekonomian Indonesia adalah lada. Usahatani lada mampu menyumbang negara, menyediakan lapangan pekerjaan serta menjadi bahan baku industri maupun konsumsi langsung (Kemala (2006). Pengembangan usahatani lada di Indonesia sebagian besar berada di Luar Jawa seperti Kepulauan Bangka Belitung, Lampung, Sumatera Selatan, Sulawesi Selatan dan Kalimantan Timur. Berdasarkan data Kementerian Pertanian (2020) kelima provinsi tersebut berkontribusi sebesar 70,42% terhadap produksi lada di Indonesia. Jenis lada yang paling banyak diproduksi di Indonesia adalah lada hitam yang berasal dari Lampung atau dikenal dengan nama *Lampung black pepper* dan lada putih atau *Muntok white pepper* yang dihasilkan di Kepulauan Bangka Belitung (Nurdjannah, 2006).

Pengusahaan lada di Indonesia didominasi oleh Perkebunan Rakyat (PR) dan sebagian kecil diusahakan oleh Perkebunan Besar Swasta (PBS). Menurut Zikria (2020) sebesar 97,97% produksi lada di Indonesia selama periode 2011-2020 didominasi oleh perkebunan rakyat, sedangkan kontribusi lada yang berasal dari PBS hanya 2,04%. Pengusahaan lada oleh perkebunan rakyat umumnya masih dilakukan secara tradisional (Damanik, 2001). Menurut Kardinan et al. (2018) budidaya lada secara tradisional dilakukan petani dimana penggunaan bahan kimia sebagai asupan tambahan menjadi andalan dan merupakan garansi keberhasilan petani. Oleh karena itu, ketika pupuk dan pestisida sulit diperoleh maka asupan tanaman lada berkurang yang mengakibatkan produksinya turun.

Selama sepuluh tahun terakhir (2011-2020) produksi lada di Indonesia cenderung berfluktuasi. Tahun 2011 produksi lada sebesar 87.089 ton dan turun menjadi 86.083 ton pada tahun 2020 atau terjadi penurunan 0,06% per tahun (Kementerian Pertanian, 2020). Meskipun demikian, tidak menutup kemungkinan bahwa produksi lada beberapa tahun ke depan diproyeksikan naik selama didukung dengan kebijakan pengembangan baik ekstensifikasi maupun intensifikasi. Estimasi produksi lada beberapa tahun ke depan sangat penting utamanya sebagai bahan penentu kebijakan maupun *Early Warning System (EWS)* mengingat masih terdapat lag antara data Angka Tetap (ATAP) Perkebunan dengan data di tahun berjalan (*real time*).

Estimasi produksi pada tahun berjalan dan beberapa periode ke depan sangat penting untuk dilakukan. Hal ini karena informasi tersebut menjadi bahan dalam penentuan kebijakan di subsektor perkebunan. Direktorat Jenderal Perkebunan telah mengupayakan penyediaan data

yang *near real time* dengan melakukan penyusunan Angka Estimasi (AESTI) pada tahun berjalan. Namun estimasi produksi komoditas perkebunan selama lima tahun ke depan masih belum tersedia. Estimasi Ditjen Perkebunan hanya dilakukan untuk satu tahun ke depan menggunakan model *univariate* seperti *Double Exponential Smoothing (DES)*. Salah satu kelemahan dari model *univariate* yaitu variabel yang digunakan hanya satu misalnya produksi. Akibatnya hasil analisis hanya mampu memberikan gambaran terhadap satu variabel saja tanpa adanya intervensi dari variabel lain. Padahal produksi komoditas perkebunan tidak terlepas dari pengaruh variabel-variabel lain seperti luas areal, ekspor-impor serta variabel lainnya. Selain itu, untuk menentukan kebijakan subsektor perkebunan seperti peningkatan produksi, diperlukan informasi variabel input lain yang diduga turut berpengaruh terhadap produksi komoditas perkebunan sehingga intervensi kebijakan dapat dilakukan pada variabel input tersebut. Oleh karena itu diperlukan model yang mampu menyajikan analisis mendalam dalam mengestimasi produksi dengan melibatkan variabel input lain, misalnya model *multivariate*.

Penelitian ini menyajikan hasil estimasi produksi lada di Indonesia dengan model *univariate* maupun *multivariate*. Terdapat tiga model yang digunakan dalam mengestimasi produksi lada antara lain *Autoregressive Integrated Moving Average (ARIMA)*, fungsi transfer dan *Vector Autoregression (VAR)*. Model ARIMA menghasilkan estimasi produksi lada tanpa ada pengaruh dari variabel lain. Model fungsi transfer menghasilkan angka estimasi produksi dengan memasukkan intervensi dari satu variabel pendukung yang dianggap paling berpengaruh terhadap produksi. Model VAR mengestimasi produksi dengan mempertimbangkan pengaruh dari beberapa variabel lain atau terdapat lebih dari satu variabel pendukung yang diduga berpengaruh terhadap produksinya. Hasil estimasi dari ketiga model tersebut akan dibandingkan untuk selanjutnya ditentukan model terbaik untuk meramalkan produksi lada di Indonesia beberapa tahun ke depan. Pemilihan model terbaik dilakukan dengan membandingkan tingkat kesalahan (*error*) terkecil yang dihasilkan oleh masing-masing model. Selain itu, model terbaik yang dipilih juga mempertimbangkan kelogisan hasil estimasi dibandingkan perkembangan produksi pada periode sebelumnya. Program kebijakan yang dilakukan oleh Ditjen Perkebunan pada tahun berjalan juga menjadi pertimbangan dalam menentukan model beserta hasil estimasinya. Hal ini karena intervensi di tahun berjalan tersebut diduga akan berdampak positif terhadap produksi lada beberapa tahun ke depan.

## METODE PENELITIAN

### Kerangka Pemikiran

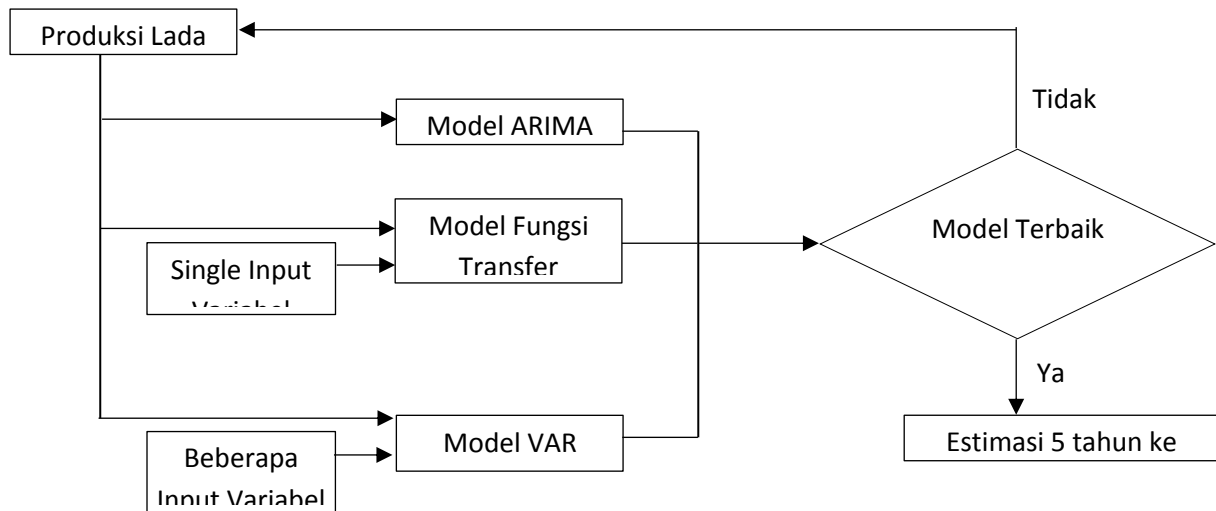
Kerangka pemikiran dalam penelitian ini menggambarkan hasil estimasi produksi lada dengan tiga model yaitu ARIMA, fungsi transfer dan VAR. Model ARIMA umumnya digunakan untuk peramalan jangka pendek, sedangkan untuk peramalan jangka panjang maka model ini kurang baik ketepatan hasil estimasinya. Estimasi dengan model ARIMA hanya menggunakan variabel dependen dan mengabaikan variabel independennya. Dengan kata lain, untuk mengestimasi produksi lada beberapa tahun ke depan maka variabel yang digunakan hanya produksi itu sendiri.

Model fungsi transfer menggambarkan nilai ramalan masa depan dari suatu deret berkala (deret output) yang didasarkan pada nilai-nilai masa lalu dari deret itu sendiri serta didasarkan pula pada suatu deret berkala yang berhubungan (deret input). Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linear antara waktu ke- $t$  dengan deret/variabel input, tetapi juga terdapat hubungan antara variabel input dengan variabel output pada waktu ke- $t$ ,  $t+1$ , ...,  $t+k$ . Pada fungsi transfer terdapat rangkaian output yang mungkin dipengaruhi oleh rangkaian *multiple* input. Untuk kasus *single input* variabel pada fungsi transfer, dapat menggunakan metode korelasi silang. Penelitian ini menggunakan *single input* variabel yaitu volume ekspor untuk meramalkan produksi lada.

Model VAR menggunakan pendekatan non-struktural atau tidak mendasarkan pada teori ekonomi tertentu dalam melakukan peramalan. Model ini memperlakukan seluruh peubah secara simetris tanpa mempermasalahkan variabel dependen dan independen atau dengan kata lain model VAR memperlakukan seluruh variabel sebagai variabel dependen/endogen, karena pada kenyataannya suatu variabel dapat bertindak sebagai endogen atau dependen dari suatu peubah lain (Gujarati & Porter, 2010). Untuk meramalkan produksi lada beberapa tahun ke depan, penelitian ini menggunakan beberapa variabel antara lain luas areal, volume ekspor, dan volume impor.

Pembentukan model estimasi produksi lada dilakukan dengan membagi series data aktual menjadi data training dan data testing. Data training digunakan untuk menentukan model estimasi dan meramalkan data testing yang sebenarnya sudah tersedia data aktualnya. Hasil ramalan data testing tersebut kemudian dibandingkan dengan data aktual untuk dihitung tingkat kesalahan (*error*) hasil ramalan. Model terbaik untuk estimasi adalah model dengan tingkat *error* yang paling kecil, dalam hal ini ditunjukkan oleh nilai *Mean Absolute Percentage Error*

(MAPE) terkecil. Selain MAPE, pemilihan model terbaik juga mempertimbangkan kelogisan hasil ramalan dengan historis data sebelumnya. Berdasarkan hasil identifikasi model ARIMA, fungsi transfer dan VAR, dipilih model terbaik untuk meramalkan produksi lada di Indonesia selama lima tahun ke depan. Secara umum tahapan penelitian ini disajikan melalui kerangka konseptual penelitian pada Gambar 1.



Gambar 1. Kerangka Konseptual Penelitian

### Pengumpulan Data

Data yang digunakan pada analisis ini bersumber dari Direktorat Jenderal Perkebunan, Kementerian Pertanian. Variabel yang digunakan antara lain produksi lada, luas areal lada, volume ekspor lada, dan volume impor lada. Produksi lada yang digunakan merupakan total produksi lada baik yang diusahakan oleh Perkebunan Rakyat (PR), Perkebunan Besar Negara (PBN) dan Perkebunan Besar Swasta (PBS). Luas areal lada yang digunakan merupakan penjumlahan dari luas Tanaman Belum Menghasilkan (TBM), Tanaman Menghasilkan (TM) dan Tanaman Rusak/Tanaman Tidak Menghasilkan (TR/TTM). Volume ekspor maupun volume impor dihitung berdasarkan enam kode HS yaitu 09041110, 09041120, 09041190, 09041210, 09041220, 09041290. Series data yang digunakan pada penelitian ini yaitu data tahun 1976-2020, dimana data tersebut seluruhnya merupakan Angka Tetap (ATAP). Berdasarkan series data tersebut, selanjutnya dilakukan pengelompokan data training untuk periode 1976-2014 dan data testing untuk periode 2015-2020, sehingga diperoleh total observasi sebanyak 45. Berdasarkan model terbaik yang terpilih, dilakukan estimasi produksi lada di Indonesia selama lima tahun ke depan yaitu 2021-2025.

## Analisis Data

Secara empiris, penelitian ini membandingkan hasil estimasi produksi lada dengan tiga model estimasi yaitu ARIMA, fungsi transfer dan VAR. Tahapan penelitian dimulai dengan mencari model estimasi berdasarkan historis data training untuk meramalkan data testing. Selanjutnya hasil estimasi data testing dibandingkan dengan nilai aktual produksinya untuk mengetahui tingkat kesalahan berdasarkan nilai MAPE. Berdasarkan nilai MAPE yang dihasilkan oleh ketiga model estimasi tersebut dipilih model ramalan dengan MAPE terkecil. Model dengan MAPE terkecil dipilih sebagai model terbaik untuk mengestimasi produksi lada selama lima tahun ke depan. Model terbaik yang terpilih juga harus memenuhi asumsi statistik yang ditetapkan di masing-masing model. Pengolahan data untuk estimasi produksi lada baik dengan model ARIMA, fungsi transfer maupun VAR dilakukan dengan program RStudio.

## Estimasi dengan Model ARIMA

Model ARIMA dibagi ke dalam tiga kelompok model yaitu *Autoregressive Model (AR)*, *Moving Average Model (MA)* dan *Autoregressive Integrated Moving Average Model (ARIMA)*. Model AR menjelaskan pergerakan suatu peubah itu sendiri di masa lalu. Model AR ordo ke- $p$  untuk mengestimasi produksi lada atau dapat ditulis ARIMA ( $p, 0, 0$ ) sebagai berikut:

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t \dots \dots \dots (1)$$

dimana:

- $Y_t$  = produksi lada pada waktu ke- $t$
- $Y_{t-p}$  = produksi lada pada kurun waktu ke ( $t-p$ )
- $\mu$  = suatu konstanta
- $\theta_1 \dots \theta_p$  = parameter autoregressive ke- $p$
- $\varepsilon_t$  = nilai kesalahan pada waktu ke- $t$

Model MA menjelaskan pergerakan peubahnya melalui sisaannya di masa lalu. Model MA dengan ordo  $q$  untuk mengestimasi produksi lada atau ARIMA ( $0,0,q$ ) ditulis sebagai berikut:

$$Y_t = \mu - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t \dots \dots \dots (2)$$

dimana:

- $Y_t$  = produksi lada pada waktu ke- $t$
- $\phi_1 \dots \phi_q$  = parameter-parameter moving average



$\varepsilon_{t-q}$  = nilai kesalahan pada waktu ke  $(t-q)$

Model ARIMA merupakan model dari fungsi linear nilai lampau beserta nilai sekarang dan sisaan lampainya. Bentuk model ARIMA  $(p,d,q)$  untuk mengestimasi produksi lada ditulis sebagai berikut:

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t \dots \dots \dots (3)$$

dimana:

$Y_t$  = produksi lada pada waktu ke- $t$

$Y_{t-p}$  = produksi lada pada kurun waktu ke  $(t-p)$

$\mu$  = suatu konstanta

$\theta_1 \theta_2 \dots \theta_p \phi_1 \phi_2 \dots \phi_q$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke- $(t-q)$

Penggunaan model ARIMA mensyaratkan series data yang stasioner. Suatu deret waktu yang tidak stasioner harus diubah menjadi data stasioner dengan melakukan *differencing*. *Differencing* yaitu menghitung perubahan atau selisih nilai observasi. Data yang telah dilakukan *differencing* perlu dicek kembali apakah telah stasioner atau belum. Pengecekan stasioneritas data dapat dilihat dengan cara melihat sebaran data, menggunakan uji *Augmented Dickey-Fuller (ADF) Test Unit Root Test* dan melihat dari perilaku autokorelasi berdasarkan plot *Autocorrelation Function (ACF)* dan *Partial Autocorrelation Function (PACF)*.

Berdasarkan sebaran datanya, data yang telah stasioner menyebar secara acak dan tidak memiliki pola-pola tertentu baik pola musiman maupun *trend*. Pengecekan stasioneritas dengan uji ADF memiliki hipotesis sebagai berikut:

Hipotesis: .....(4)

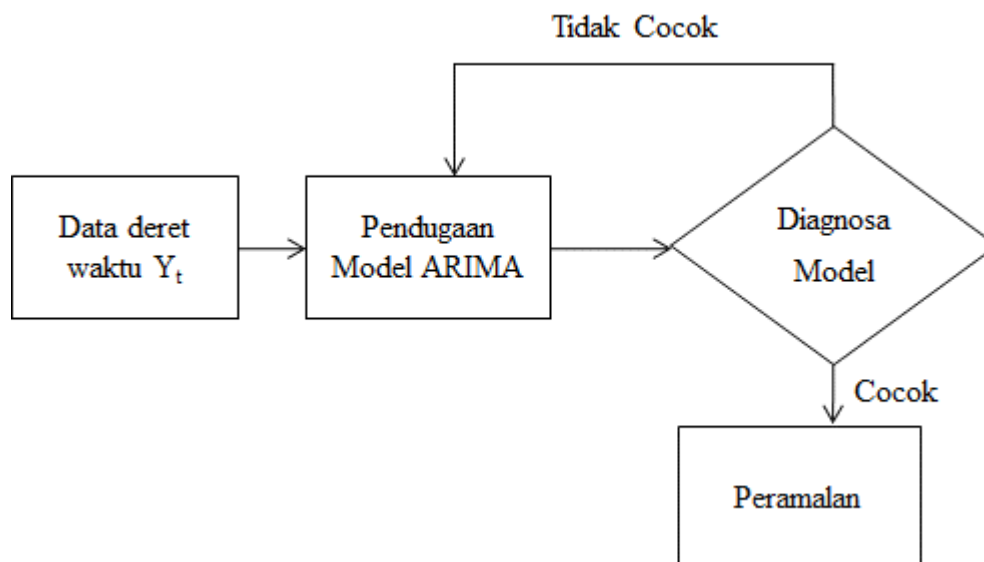
$H_0$  : Data tidak stasioner

$H_1$  : Data stasioner

Jika nilai *test-statistic* pada uji ADF lebih kecil dari *critical value for test-statistic* baik pada taraf  $(\alpha)$  1%, 5% atau 10% maka  $H_0$  ditolak yang berarti data telah stasioner. Pengecekan stasioneritas dari perilaku *autokorelasi* dilihat dari plot ACF dan PACF. Jika pada kedua plot tersebut tidak banyak lag yang keluar dari *confidence interval* maka data telah stasioner.

Pada data yang telah stasioner dilakukan tahapan pendugaan model ARIMA menggunakan fungsi *auto.arima* atau *armaselect* yang tersedia pada program RStudio. Program tersebut akan memberikan rekomendasi model terbaik untuk mengestimasi produksi lada. Berdasarkan model terbaik yang terpilih, kemudian dilakukan pemeriksaan sisaan menggunakan pengujian LJungBox. Jika autokorelasi sisaan tidak signifikan yang ditandai dengan nilai p-value yang lebih besar dari 5% atau 10%, maka model ARIMA tersebut sudah cukup baik untuk mengepas data produksi lada.

Model ARIMA yang terpilih digunakan untuk mengestimasi data testing. Hasil ramalan data testing selanjutnya dibandingkan dengan data aktualnya untuk mengecek akurasi hasil ramalan. Akurasi hasil ramalan model ARIMA ditunjukkan oleh MAPE data training dan data testing. Jika model terpilih dirasa telah menghasilkan MAPE yang kecil, maka model tersebut dapat digunakan untuk mengestimasi produksi lada untuk beberapa periode ke depan. Selain MAPE terkecil, estimasi ke depan juga perlu mempertimbangkan kelogisan antara historis data dengan hasil estimasinya. Pemodelan untuk estimasi produksi lada dengan ARIMA dilakukan melalui beberapa langkah sebagaimana disajikan pada Gambar 2 berikut:



Gambar 2. Langkah-langkah Estimasi Produksi Lada dengan Model ARIMA

### Estimasi dengan Model Fungsi Transfer

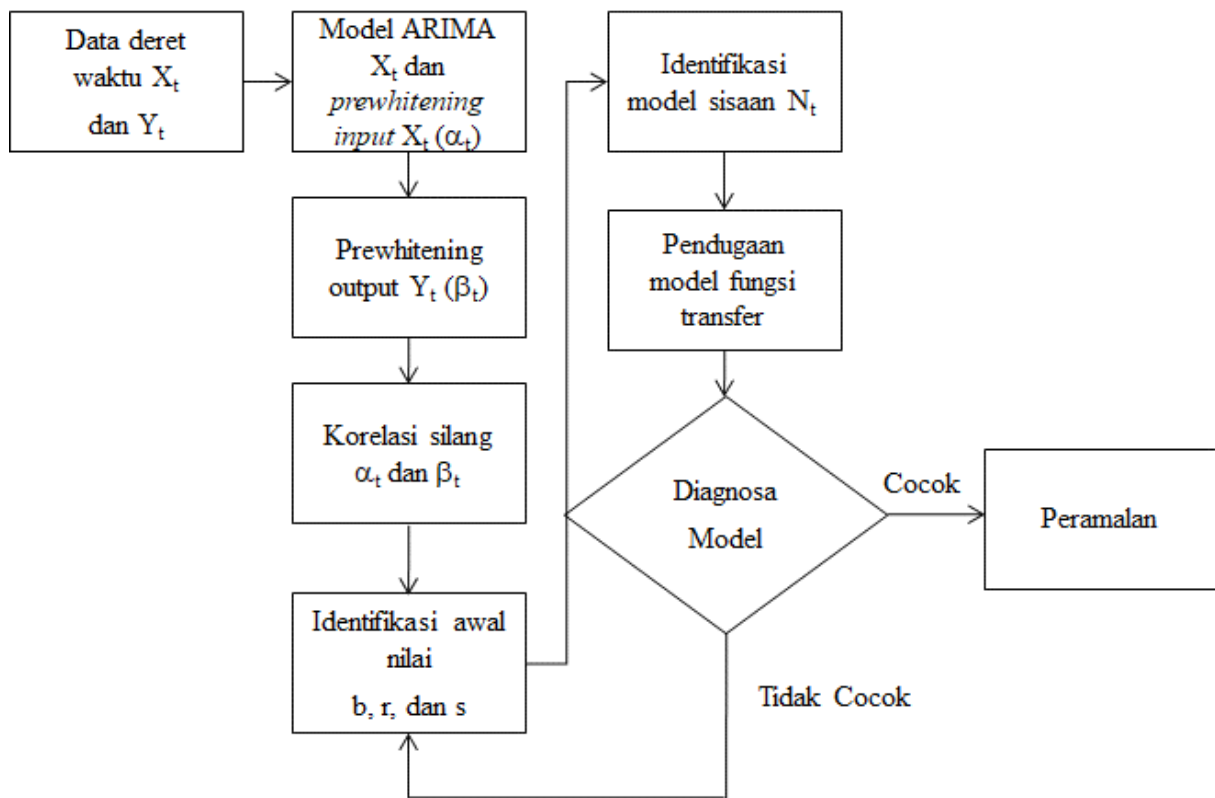
Produksi lada di Indonesia tidak terlepas dari pengaruh variabel lain. Menurut Hamdani et al. (2015) salah satu variabel yang berpengaruh terhadap produksi lada di Indonesia adalah volume ekspor. Analisis ini menggunakan volume ekspor sebagai variabel input dalam mengestimasi produksi (variabel output) menggunakan model fungsi transfer. Model fungsi transfer pada penelitian ini menggambarkan ramalan produksi lada yang didasarkan pada nilai-nilai masa lalu dari produksi itu sendiri, serta didasarkan pada volume ekspor (variabel input) dan gangguan/noise. Model fungsi transfer untuk mengestimasi produksi lada dituliskan sebagai berikut:

$$y_t = v(B)x_t + N_t \quad \Rightarrow \quad y_t = \frac{\omega_s(B)}{\delta_r(B)} x_{t-s} + \frac{\theta_p(B)}{\phi_q(B)} \varepsilon_t \quad \dots\dots\dots(5)$$

dimana:

- $y_t$  = produksi lada tahun ke-t
- $x_t$  = volume ekspor lada tahun ke-t
- $b$  = panjang jeda pengaruh volume ekspor terhadap produksi lada
- $r$  = panjang lag produksi lada periode sebelumnya yang masih mempengaruhi produksi lada tahun-t
- $s$  = panjang jeda volume ekspor lada periode sebelumnya yang masih mempengaruhi produksi lada tahun-t
- $p$  = ordo AR bagi noise  $N_t$
- $q$  = ordo MA bagi noise  $N_t$

Pemodelan untuk estimasi produksi lada dengan fungsi transfer dilakukan melalui beberapa langkah sebagaimana disajikan pada Gambar 3 berikut:



Gambar 3. Langkah-langkah Estimasi Produksi Lada dengan Model Fungsi Transfer

### Estimasi dengan Model VAR

Pemodelan dengan *Vector Autoregression (VAR)* memperlakukan seluruh peubah secara simetris tanpa mempermasalahkan apakah variabel tersebut merupakan variabel dependen maupun independen. Penelitian ini menggunakan produksi lada sebagai variabel dependen. Produksi lada tersebut diduga dipengaruhi oleh beberapa variabel, tidak seperti model fungsi transfer yang hanya menggunakan volume ekspor saja sebagai variabel inputnya. Selain volume ekspor (Hamdani et al., 2015), variabel input lain yang diduga juga berpengaruh terhadap produksi lada di Indonesia adalah luas areal (Fatma et al., 2020; Nursalam, 2020). Meskipun lada merupakan komoditas ekspor, namun tidak menutup kemungkinan Indonesia juga mengimpor lada dari negara lain. Selain itu, terdapat fenomena lada ekspor asal Indonesia yang ditolak (*reject*) oleh negara tujuan ekspor dikarenakan isu mutu dan keamanan pangan (Ditjenbun.pertanian.go.id, 2019). Oleh karena itu, penelitian ini menambahkan variabel volume impor untuk mengestimasi produksi lada di Indonesia.

Model VAR termasuk kategori model sistem, dimana ketika tidak ada kepastian untuk menentukan bahwa suatu peubah adalah eksogen (independen) maka suatu perluasan analisis fungsi perpindahan alami akan memperlakukan masing-masing peubah secara simetris (Enders,

2004). Sebagai contoh, pada kasus-kasus peubah yang membiarkan alur waktu atau *time path*  $\{s_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{y_t\}$  dan membiarkan *time path*  $\{y_t\}$  dipengaruhi oleh nilai saat ini dan waktu sebelumnya dari  $\{s_t\}$ . Di dalam sistem *bivariate*, hubungan tersebut dapat digambarkan seperti pada persamaan berikut:

$$\begin{aligned} s_t &= b_{10} - b_{12}y_t + \gamma_{11}s_{t-1} + \gamma_{12}y_{t-1} + \varepsilon_{s_t} \\ y_t &= b_{20} - b_{21}s_t + \gamma_{21}s_{t-1} + \gamma_{22}y_{t-1} + \varepsilon_{y_t} \end{aligned} \dots\dots\dots(6)$$

Dengan mengasumsikan bahwa kedua peubah  $s_t$  dan  $y_t$  adalah stasioner:  $\varepsilon_{s_t}$  dan  $\varepsilon_{y_t}$  adalah *disturbances* yang memiliki rata-rata nol dan matriks kovarians terbatas atau bersifat *white noise* dengan standar deviasi yang berurutan  $\sigma_s$  dan  $\sigma_y$ : serta  $\{\varepsilon_{s_t}\}$  dan  $\{\varepsilon_{y_t}\}$  adalah *disturbances* yang independen dengan rata-rata nol dan kovarian terbatas (*uncorrelated white-noise disturbances*). Kedua persamaan di atas merupakan orde pertama VAR, karena panjang *lag* nya hanya satu. Agar Persamaan (6) lebih mudah dipahami dan digunakan sebagai alat analisis maka ditransformasikan dengan menggunakan matriks aljabar, dan hasilnya dapat dituliskan secara bersama seperti pada persamaan di bawah ini:

$$\begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \begin{bmatrix} s_t \\ y_t \end{bmatrix} = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} + \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \begin{bmatrix} s_{t-1} \\ y_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix} \text{ atau dengan bentuk lain:}$$

$$Bx_t = \Gamma_0 + \Gamma_1 X_{t-1} + \varepsilon_t \dots\dots\dots(7)$$

dimana:

$$B = \begin{bmatrix} 1 & b_{12} \\ b_{21} & 1 \end{bmatrix} \quad x_t = \begin{bmatrix} s_t \\ y_t \end{bmatrix} \quad \Gamma_0 = \begin{bmatrix} b_{10} \\ b_{20} \end{bmatrix} \quad \Gamma_1 = \begin{bmatrix} \gamma_{11} & \gamma_{12} \\ \gamma_{21} & \gamma_{22} \end{bmatrix} \quad \varepsilon_t = \begin{bmatrix} \varepsilon_{s_t} \\ \varepsilon_{y_t} \end{bmatrix}$$

Dengan melakukan pengalihan antara persamaan (7) dengan  $B^{-1}$  atau invers matriks B, maka akan dapat ditentukan model VAR dalam bentuk standar, seperti dituliskan pada persamaan di bawah ini:

$$x_t = A_0 + A_1 x_{t-1} + \ell_t \dots\dots\dots(8)$$

dimana:

$$A_0 = B^{-1} \Gamma_0$$

$$A_1 = B^{-1} \Gamma_1$$

$$\ell_t = B^{-1} \varepsilon_t$$

Pada penelitian ini  $X_t$  merupakan matriks yang dibentuk dari produksi lada, luas areal, volume ekspor, dan volume impor.

Model VAR didasarkan pada beberapa asumsi antara lain:

**- Sisaan mengikuti fungsi distribusi normal**

Uji normalitas pada model VAR didasarkan pada nilai *Jarque-Bera (JB) test (multivariate)*, *Skewness only (multivariate)*, dan *Kurtosis only (multivariate)*. Hipotesis yang mendasari yaitu:

Hipotesis: .....(9)

$H_0$  : Sisaan terdistribusi normal

$H_1$  : *Otherwise*

Jika nilai p-value dari *JB test, Skewness only dan Kurtosis only* lebih besar dari nilai  $\alpha$  maka  $H_0$  diterima, dengan kata lain asumsi normalitas terpenuhi.

**- Varians sisaan konstan untuk setiap data pengamatan (homoskedastisitas)**

Asumsi homoskedastisitas pada model VAR didasarkan pada nilai *ARCH (multivariate)* dengan hipotesis sebagai berikut:

Hipotesis: .....(10)

$H_0$  : Homoskedastisitas

$H_1$  : Heterokedastisitas

Jika nilai p-value pada *ARCH (multivariate)* lebih besar dari nilai  $\alpha$  maka  $H_0$  diterima, dengan kata lain asumsi homoskedastisitas terpenuhi.

**- Tidak terdapat autokorelasi antar sisaan untuk setiap data pengamatan**

*Pormanteau Test (asymptotic)* digunakan untuk menguji ada tidaknya autokorelasi antar sisaan pada data amatan. Hipotesis yang digunakan sebagai berikut:

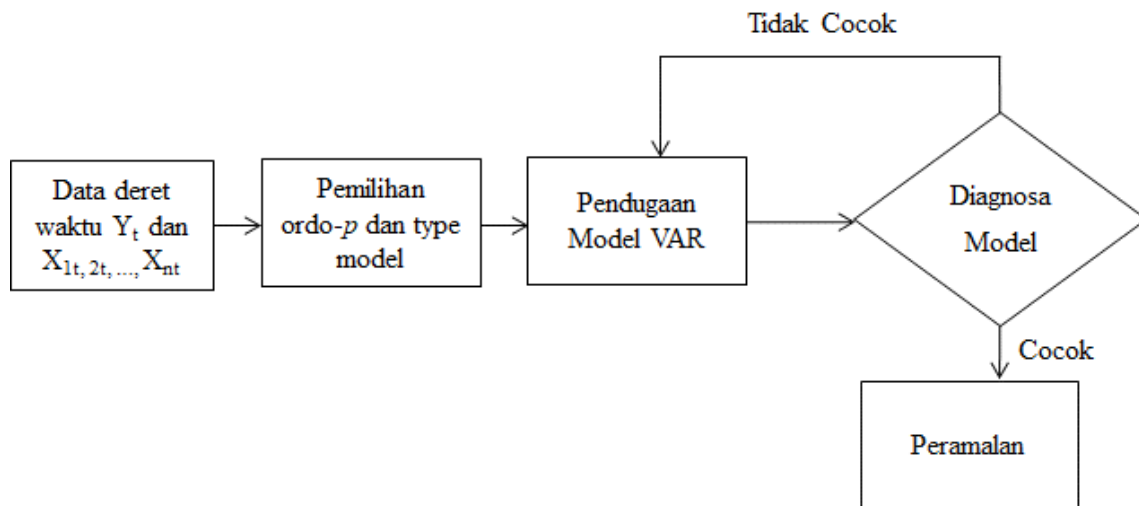
Hipotesis: .....(11)

$H_0$  : Tidak ada autokorelasi antar sisaan

$H_1$  : *Otherwise*

Jika nilai p-value hasil uji *Pormanteau Test* lebih besar dari nilai  $\alpha$  maka  $H_0$  diterima, dengan kata lain asumsi tidak adanya autokorelasi antar sisaan telah terpenuhi.

Pemodelan untuk estimasi produksi lada dengan model VAR dilakukan melalui beberapa langkah sebagaimana disajikan pada Gambar 4 berikut:



Gambar 4. Langkah-langkah Estimasi Produksi Lada dengan Model VAR

## HASIL DAN PEMBAHASAN

### Keragaan Lada di Indonesia

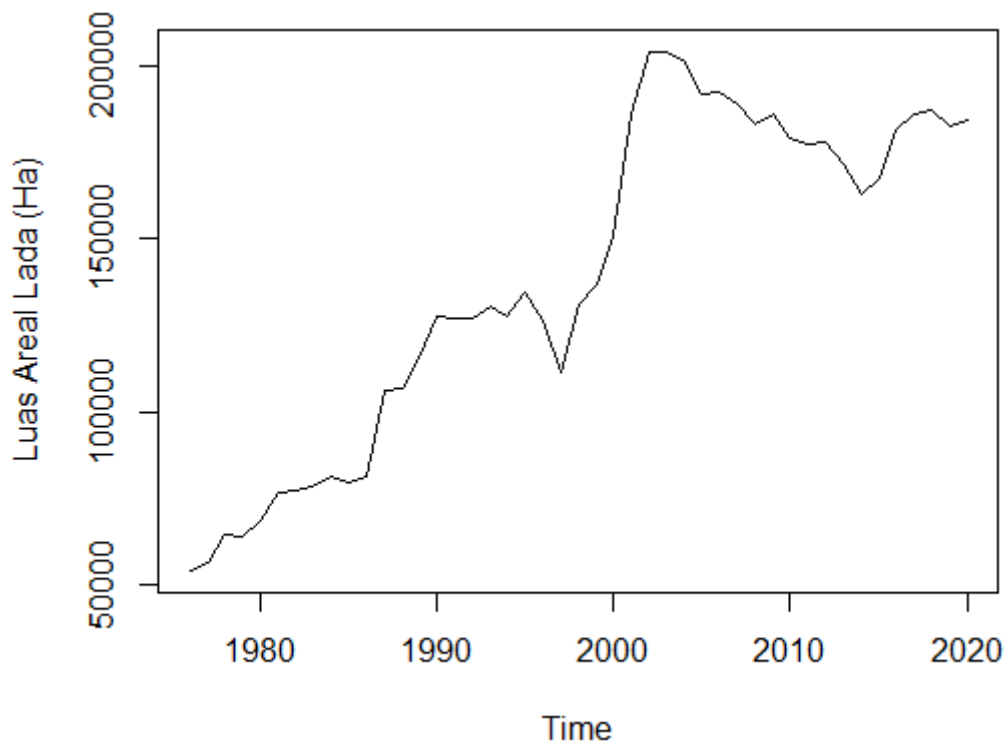
Lada merupakan salah satu komoditas unggulan sub sektor perkebunan dengan nilai ekonomis tinggi. Pengembangan usahatani lada di Indonesia sebagian besar berada di Luar Jawa seperti Kepulauan Bangka Belitung, Lampung, Sumatera Selatan, Sulawesi Selatan dan Kalimantan Timur. Berdasarkan data Kementerian Pertanian (2020) kelima provinsi tersebut berkontribusi sekitar 71% terhadap produksi lada di Indonesia. Jenis lada yang paling banyak diproduksi di Indonesia adalah lada hitam yang berasal dari Lampung atau dikenal dengan nama *Lampung black pepper* dan lada putih atau *Muntok white pepper* yang dihasilkan di Kepulauan Bangka Belitung (Nurdjannah, 2006). Perkembangan produksi lada dari tahun 1976 hingga tahun 2020 cenderung berfluktuasi (Gambar 5). Pada tahun 1976 produksi lada sebesar 36.854 ton kemudian meningkat menjadi 86.083 ton pada tahun 2020. Selama sepuluh tahun terakhir (2011-2020) produksi lada tercatat turun sebesar 0,06% per tahun, dengan produksi tertinggi dicapai pada tahun 2013 sebesar 91.039 ton.



Gambar 5. Perkembangan Produksi Lada di Indonesia Tahun 1976-2020



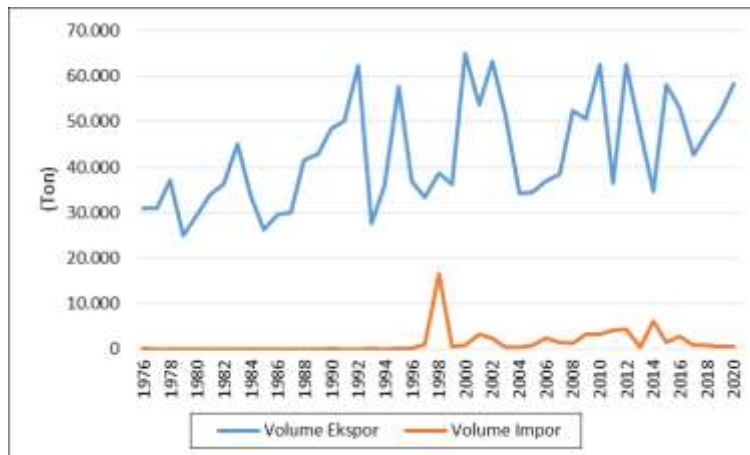
Fluktuasi produksi lada terjadi seiring dengan perkembangan luas arealnya yang juga berfluktuasi dari tahun ke tahun (Gambar 6). Pada tahun 1976 luas areal lada baik meliputi TBM, TM maupun TR/TTM sebesar 53.817 ha kemudian mengalami peningkatan menjadi 191.635 ha pada tahun 2020. Dengan rata-rata penambahan luas areal lada sebesar 0,92% per tahun selama sepuluh tahun terakhir, pengembangan areal lada di Indonesia melalui program ekstensifikasi masih memungkinkan untuk dilakukan. Komposisi luasan lada di Indonesia sendiri didominasi oleh TM diikuti dengan TBM dan TR/TTM.



Gambar 6. Perkembangan Luas Areal Lada di Indonesia Tahun 1976-2020

Sebagai salah satu produsen lada terbesar di dunia, Indonesia juga merupakan negara yang paling banyak mengekspor lada. Meskipun volume ekspor lada cenderung berfluktuatif dari tahun ke tahun, jika dibandingkan dengan volume impornya maka kuantitas lada yang diekspor masih jauh lebih banyak (Gambar 7). Ekspor impor lada yang digunakan pada analisis ini merujuk pada enam kode HS. Pada tahun 1976 volume ekspor lada sebesar 30.831 ton dengan volume impornya hanya 128 ton. Di tahun 2020 volume ekspor lada meningkat menjadi 58.378 ton dengan volume impor sebesar 603 ton. Jika dibandingkan antara produksi dengan

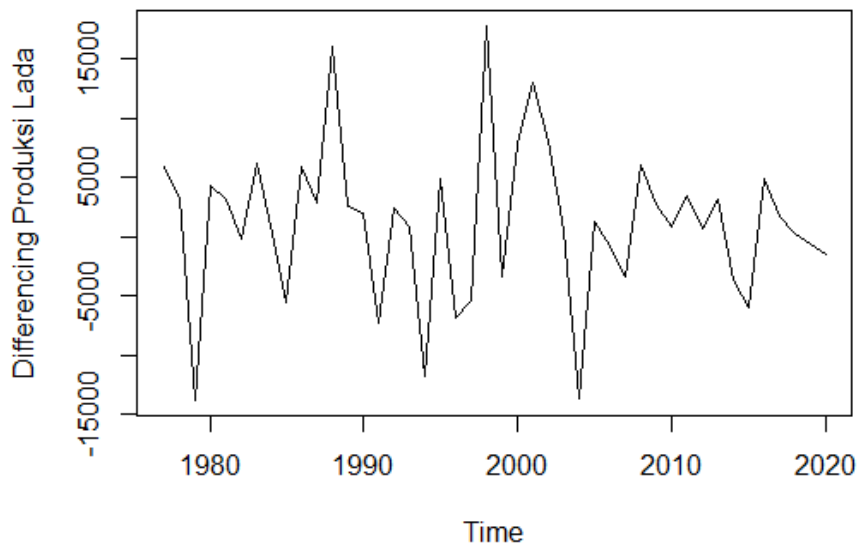
volume ekspor lada, terlihat bahwa lebih dari 50% produksi lada Indonesia diekspor untuk memenuhi permintaan lada dunia.



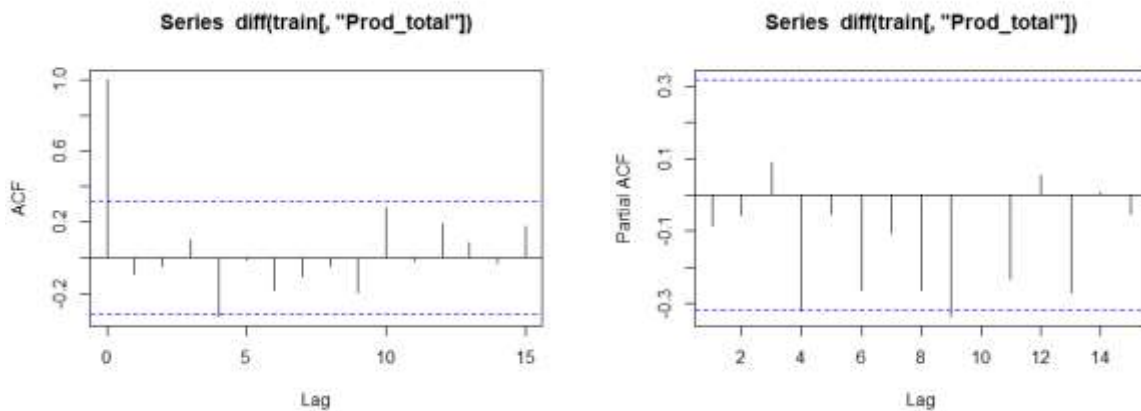
Gambar 7. Perkembangan Volume Ekspor dan Volume Impor Lada di Indonesia Tahun 1976-2020

### Hasil Estimasi Produksi Lada dengan Model ARIMA

Pemodelan ARIMA dilakukan untuk data yang telah stasioner. Berdasarkan plot data serta hasil uji ADF terindikasi bahwa data produksi lada belum stasioner, sehingga dilakukan *differencing* satu kali. *Differencing* adalah menghitung perubahan atau selisih nilai observasi. Setelah dilakukan *differencing*, data tersebut menjadi stasioner (Gambar 8) yang ditunjukkan dengan *value of test-statistic* sebesar -4,86, dimana nilai tersebut lebih kecil dari *critical values for test-statistics* baik untuk  $\alpha$  sebesar 1% (-2,62), 5% (-1,95) maupun 10% (-1,61). Stasioneritas data tersebut juga didukung dengan plot ACF dan PACF, dimana pada plot tersebut tidak banyak lag yang keluar dari *confidence interval* (Gambar 9).



Gambar 8. Plot Produksi Lada Setelah *Differencing*



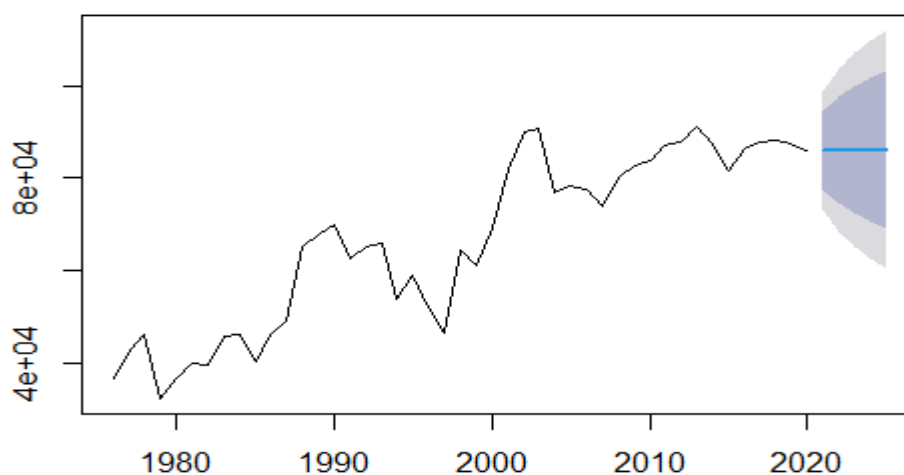
Gambar 9. Plot ACF dan PACF Setelah Dilakukan *Differencing* pada Data Produksi Lada

Dengan menggunakan program RStudio, diperoleh model terbaik untuk mengestimasi produksi lada yaitu ARIMA (1,1,1). Setelah diperoleh model ARIMA (1,1,1), dilakukan diagnosa model untuk mengecek kecocokan model dalam mengestimasi produksi lada. Diagnosa kecocokan model ARIMA didasarkan pada ada tidaknya autokorelasi antar sisaan. Autokorelasi antar sisaan tersebut diperiksa menggunakan uji LJungBox. Berdasarkan hasil pemeriksaan autokorelasi antar sisaan dengan uji LJungBox, ditemukan bahwa autokorelasi antar sisaan tidak signifikan pada seluruh lag yang ditandai dengan nilai p-value yang lebih besar dari 5%. Hal ini menunjukkan bahwa tidak ada autokorelasi antar sisaan pada model ARIMA (1,1,1). Dengan kata lain, model ARIMA (1,1,1) sudah cukup baik untuk mengepas data produksi lada. Selain itu, model ini sudah cocok digunakan untuk mengestimasi produksi lada.

Setelah dilakukan diagnosa model, maka perlu dicek akurasi model untuk peramalan. Pada model ARIMA (1,1,1) dilakukan estimasi produksi lada periode 2015-2020. Hasil estimasi tersebut selanjutnya dibandingkan dengan produksi aktualnya sehingga diperoleh persentase kesalahan rata-rata secara mutlak (MAPE). MAPE data training menggambarkan tingkat kesalahan model berdasarkan series data training yang digunakan. MAPE data testing menggambarkan tingkat kesalahan hasil ramalan dibandingkan data aktual. Model ARIMA (1,1,1) menghasilkan MAPE data training sebesar 8,98 dan MAPE data testing sebesar 1,98. Umumnya, model ARIMA yang baik memiliki nilai MAPE data testing yang kecil. Pada model ARIMA (1,1,1) dihasilkan MAPE data testing yang cukup kecil. MAPE tersebut menunjukkan bahwa tingkat kesalahan hasil ramalan produksi lada dengan model ARIMA (1,1,1) sebesar 1,98% terhadap produksi aktualnya.

Selama lima tahun terakhir (2016-2020) produksi lada di Indonesia turun 0,06% per tahun. Data Kementerian Pertanian (2020) menunjukkan bahwa pada tahun 2020 produksi lada sebesar 86.083 ton. Meskipun demikian, dengan model ARIMA (1,1,1) produksi lada di Indonesia selama lima tahun ke depan diperkirakan naik 0,01% per tahun. Pada tahun 2021 produksi lada diramalkan sebesar 86.104 ton kemudian meningkat menjadi 86.119 ton pada tahun 2022. Tahun 2023 produksi lada kembali diramalkan meningkat menjadi 86.130 ton. Peningkatan produksi tersebut terjadi hingga tahun 2024 dan 2025, dengan estimasi produksi masing-masing sebesar 86.137 ton dan 86.143 ton. Hasil estimasi produksi lada dengan model ARIMA (1,1,1) disajikan pada Gambar 10.

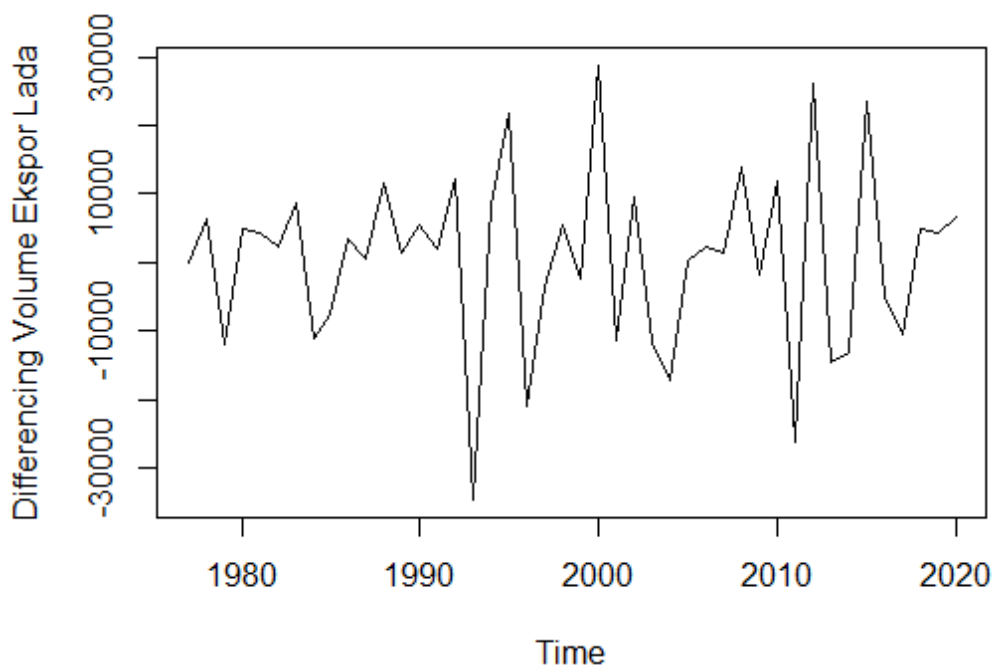
### Forecasts from ARIMA(1,1,1)



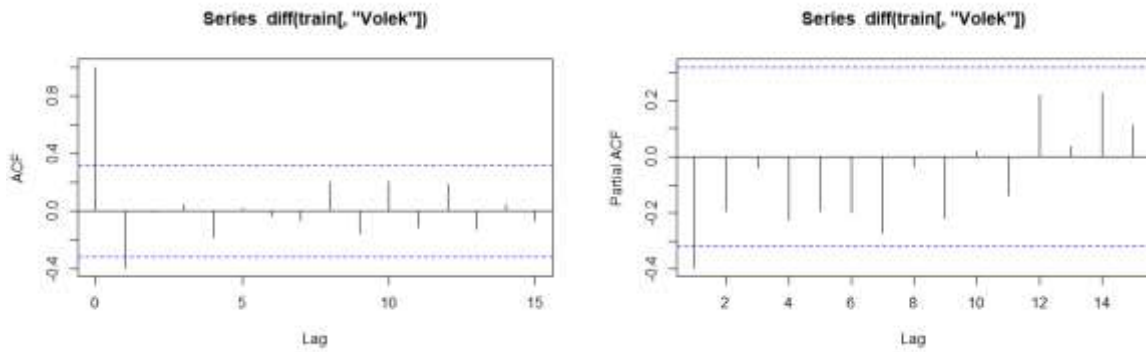
Gambar 10. Hasil Estimasi Produksi Lada dengan Model ARIMA (1,1,1)

## Hasil Estimasi Produksi Lada dengan Model Fungsi Transfer

Secara umum, estimasi produksi lada menggunakan model fungsi transfer terdiri dari beberapa tahapan antara lain peramalan variabel input dengan model ARIMA, *prewhitening* input dan output, korelasi silang, identifikasi nilai  $b$ ,  $r$ , dan  $s$ , identifikasi model sisaan/noise, pendugaan model fungsi transfer, diagnosa model, dan peramalan. Tahapan pertama estimasi produksi lada ( $Y_t$ ) dengan model fungsi transfer adalah peramalan nilai variabel input ( $X_t$ ) yaitu volume ekspor menggunakan model ARIMA. Estimasi dengan model ARIMA mensyaratkan stasioneritas pada data. Dikarenakan data volume ekspor belum stasioner, maka dilakukan *differencing* satu kali sehingga menghasilkan data yang telah stasioner (Gambar 11). Berdasarkan pengujian dengan *Augmented Dickey-Fuller Test Unit Root Test* diperoleh *value of test-statistic* -5,89 yang lebih kecil dari *critical value for test statistics* baik untuk  $\alpha$  sebesar 1% (-2,62), 5% (-1,95) maupun 10% (-1,61). Hal ini menunjukkan bahwa data volume ekspor lada telah stasioner. Stasioneritas volume ekspor lada juga ditunjukkan dengan plot ACF dan PACF, dimana pada plot tersebut tidak banyak lag yang keluar dari *confidence interval* (Gambar 12). Setelah data stasioner maka dilakukan pendugaan model. Model ARIMA terbaik untuk mengesimasi volume ekspor lada adalah ARIMA (3,1,5) dengan AIC model sebesar 824,9.



Gambar 11. Plot Volume Ekspor Lada Setelah *Differencing*

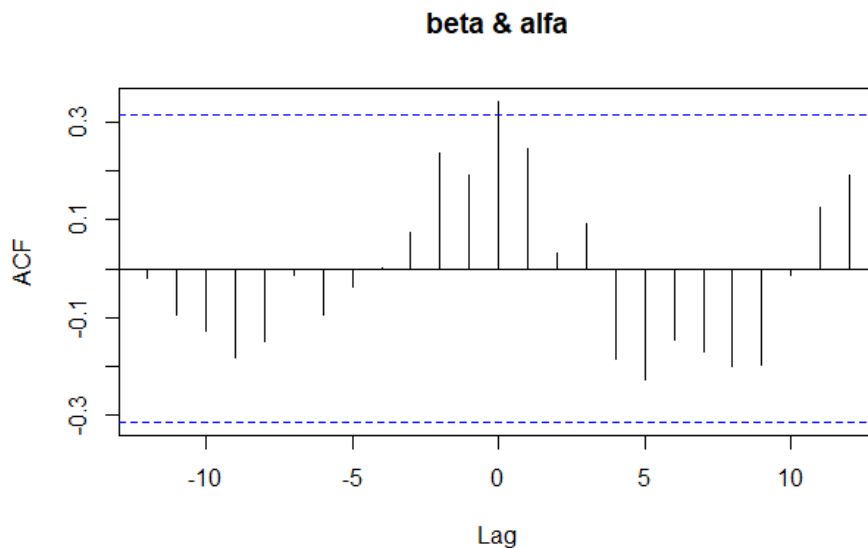


Gambar 12. Plot ACF dan PACF Setelah Dilakukan *Differencing* pada Volume Ekspor Lada

Tahapan kedua dalam pemodelan fungsi transfer adalah *prewhitening*. *Prewhitening* adalah pembentukan deret data yang tidak dipengaruhi oleh faktor luar melalui pemodelan time series ARIMA. *Prewhitening* dilakukan terhadap input dan output variabel yaitu volume ekspor lada dan produksi lada. *Prewhitening* tersebut dibentuk dari nilai residual input dan output variabel hasil pemodelan dengan ARIMA (3,1,5). *Prewhitening* input variabel (volume ekspor) selanjutnya disebut  $\alpha_t$  sedangkan *prewhitening* output variabel (produksi) disebut  $\beta_t$ .

Tahapan ketiga dalam estimasi produksi lada dengan fungsi transfer yaitu korelasi silang. Korelasi silang dilakukan antara  $\alpha_t$  dan  $\beta_t$ . Dari hasil korelasi silang tersebut diperoleh plot ACF sebagaimana disajikan pada Gambar 13. Plot ACF tersebut selanjutnya digunakan untuk mengidentifikasi nilai  $b$ ,  $r$ , dan  $s$  pada tahapan berikutnya.

Tahapan keempat pada model fungsi transfer adalah identifikasi nilai  $b$ ,  $r$ , dan  $s$ . Berdasarkan plot ACF pada Gambar 13 diperoleh nilai  $b=0$  yang ditunjukkan dengan lag pertama kali signifikan pada lag 0. Interpretasi dari nilai  $b=0$  yaitu tidak ada jeda pengaruh dampak volume ekspor lada terhadap produksi lada. Nilai  $r$  diasumsikan 0 karena data produksi lada maupun volume ekspor lada merupakan data tahunan yang tidak mengandung pola musiman. Selanjutnya dilakukan identifikasi nilai  $s$  dan diperoleh nilai  $s=0$ . Hal ini ditunjukkan dengan tidak ada tambahan lag yang signifikan setelah lag 0. Interpretasi dari nilai  $s=0$  yaitu korelasi antara volume ekspor lada dengan produksi lada terjadi di tahun yang sama.



Gambar 13. Plot ACF Hasil Korelasi Silang  $\alpha_t$  dan  $\beta_t$

Tahapan kelima yaitu identifikasi model sisaan/noise ( $N_t$ ). Identifikasi tersebut dilakukan dengan memanfaatkan fungsi *auto.arima* maupun *armaselect* pada RStudio. Model noise yang direkomendasikan berdasarkan fungsi *auto.arima* yaitu ARIMA (0,1,1) dengan AIC sebesar 814,26. Di sisi lain, salah satu model noise yang direkomendasikan oleh fungsi *armaselect* adalah ARIMA (1,1,0) dengan AIC sebesar 817,45. *Akaike Information Criteria* (AIC) merupakan kriteria yang digunakan untuk mengidentifikasi kecocokan model ARIMA. Suatu model dikatakan baik jika nilai AIC semakin kecil. Oleh karena itu, pada tahap identifikasi model noise dipilih ARIMA (0,1,1) untuk pemodelan fungsi transfer.

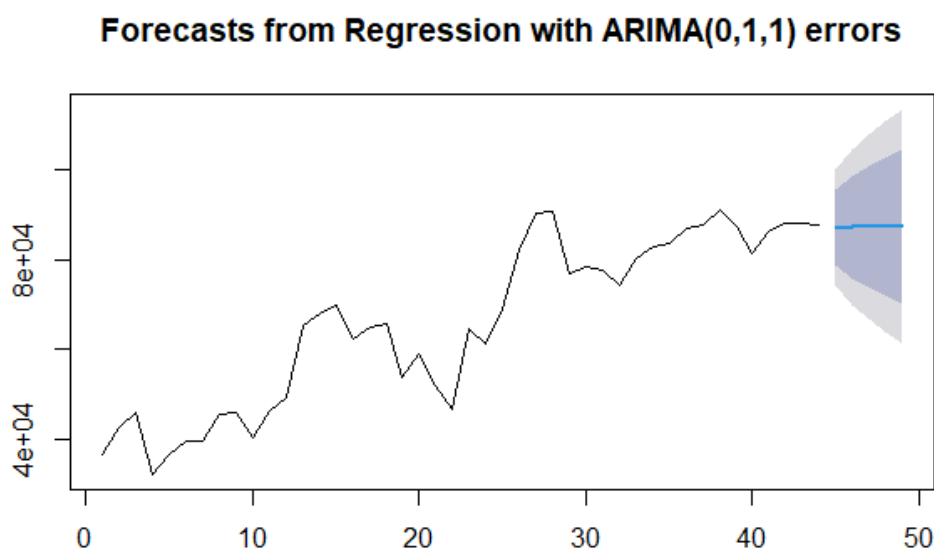
Tahapan keenam yaitu pendugaan model fungsi transfer. Model fungsi transfer yang diduga cocok untuk mengestimasi produksi lada adalah fungsi transfer ARIMA (0,1,1) dengan input variabelnya volume ekspor. Volume ekspor sendiri terlebih dahulu diestimasi dengan model ARIMA (3,1,5). Model fungsi transfer ARIMA (0,1,1) ini menghasilkan MAPE data training sebesar 8,06. MAPE data training diperoleh dari pemodelan dengan series data tahun 1976-2014. Meskipun MAPE yang dihasilkan cukup kecil, hal ini belum dapat dijadikan acuan bahwa model fungsi transfer tersebut sudah tetap. Oleh karena itu perlu dilakukan diagnosa model untuk mengetahui akurasi hasil ramalan dibandingkan dengan data aktual.

Tahapan ketujuh adalah diagnosa model fungsi transfer. Diagnosa model dilakukan dengan meramalkan produksi lada periode 2015-2020 (data testing). Hasil ramalan data testing tersebut selanjutnya dibandingkan dengan nilai aktual produksi lada. Estimasi data testing tersebut menghasilkan MAPE sebesar 1,88. MAPE tersebut menunjukkan bahwa tingkat

kesalahan hasil ramalan produksi lada periode 2015-2020 dengan model fungsi transfer ARIMA (0,1,1) sebesar 1,88% jika dibandingkan dengan produksi riilnya. Selain itu, pengujian menggunakan statistik *z test of coefficients* menunjukkan bahwa koefisien MA(1) dan input variabel (volume ekspor) berpengaruh signifikan pada model fungsi transfer ARIMA (0,1,1). Oleh karena itu, model fungsi transfer tersebut dianggap layak untuk mengestimasi produksi lada beberapa tahun ke depan.

Tahapan terakhir yaitu estimasi produksi lada selama lima tahun ke depan dengan fungsi transfer ARIMA (0,1,1). Produksi lada selama periode 2021-2025 diramalkan naik 0,06% per tahun, meskipun selama lima tahun terakhir (2016-2020) produksinya turun 0,06% per tahun. Hasil estimasi produksi lada pada tahun 2021 sebesar 87.086 ton kemudian meningkat menjadi 87.126 ton pada tahun 2022. Tahun 2023 produksi lada diramalkan kembali naik menjadi 87.386 ton. Namun sejak tahun 2024 hingga 2025 produksi lada Indonesia diprediksi turun masing-masing menjadi 87.360 ton dan 87.303 ton. (Gambar 15). Perkembangan produksi lada beserta hasil estimasinya selama lima tahun ke depan disajikan

pada Gambar 14



Gambar 14. Hasil Estimasi Produksi Lada dengan Model Fungsi Transfer

### Hasil Estimasi Produksi Lada dengan Model VAR

Estimasi dengan model VAR diawali dengan pemilihan ordo-*p* serta type model VAR yang paling baik. Type model VAR terbaik ditandai dengan banyaknya variabel yang signifikan pada ordo yang telah ditentukan. Beberapa type model VAR antara lain *both*, *const*, *trend*, dan *none*. Type *both* berarti terdapat konstanta dan trend pada model. Type *const* berarti terdapat



konstanta pada model. Type *trend* berarti terdapat trend pada model. Type *none* artinya tidak terdapat konstanta maupun trend pada model. Ordo- $p$  sendiri bernilai bulat positif seperti  $p=1$ ,  $p=2, \dots$  dan seterusnya.

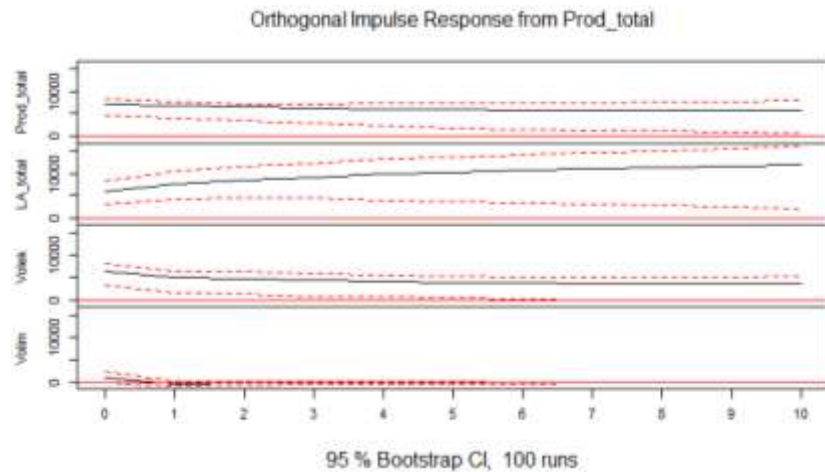
Model VAR yang diduga cocok untuk estimasi produksi lada yaitu VAR dengan ordo-1 atau VAR(1) dengan type *none* yaitu tanpa ada konstanta maupun trend. Pada model VAR(1) beberapa input variabel yang digunakan yaitu luas areal, volume ekspor, dan volume impor, sedangkan produksi menjadi output variabel. Model VAR(1) type *none* tersebut selanjutnya didiagnosa untuk mengetahui kecocokan model dalam mengestimasi produksi lada.

Diagnosa model VAR dilakukan melalui pengujian asumsi model serta tingkat kesalahan model yang ditunjukkan oleh MAPE. Terdapat tiga asumsi pada model VAR yang perlu diuji yaitu sisaan terdistribusi normal, homoskedastisitas, dan tidak terdapat autokorelasi pada sisaan. Pengujian asumsi normalitas pada sisaan dilakukan dengan uji *JB test*, *Skewness*, dan *Kurtosis*. Hasil pengujian tersebut menunjukkan bahwa sisaan tidak terdistribusi secara normal yang ditunjukkan dengan nilai  $p\text{-value}=0,00$  atau lebih kecil dari  $\alpha=5\%$ . Meskipun asumsi normalitas tidak terpenuhi, model VAR(1) dapat dipertimbangkan untuk digunakan dalam mengestimasi produksi lada mengingat fokus pemodelan pada penelitian ini adalah peramalan dengan series data yang cukup banyak. Pengujian asumsi berikutnya yaitu homoskedastisitas, yang dilakukan dengan uji ARCH. Hasil pengujian homoskedastisitas menunjukkan bahwa  $p\text{-value}$  pada ARCH sebesar 1 (lebih besar dari  $\alpha$ ), artinya asumsi homoskedastisitas terpenuhi. Pengujian asumsi model VAR yang ketiga adalah uji autokorelasi sisaan. Pengujian ini dilakukan dengan *Portmanteu Test*. Hasil pengujian tersebut menunjukkan bahwa nilai  $p\text{-value}$  dari *Portmanteu Test* sebesar 0,53 atau lebih besar dari  $\alpha=5\%$ . Hal ini menunjukkan bahwa tidak terdapat autokorelasi pada sisaan.

Selain ketiga asumsi tersebut, diagnosa kecocokan model VAR juga didasarkan pada MAPE yang dihasilkan baik untuk data training maupun data testing. Model VAR(1) type *none* menghasilkan MAPE data training sebesar 9,09 dan MAPE data testing sebesar 4,81. Kedua MAPE tersebut lebih besar jika dibandingkan dengan MAPE yang dihasilkan oleh model ARIMA (1,1,1) dan fungsi transfer ARIMA (0,1,1). Meskipun semakin kecil MAPE maka model yang dihasilkan semakin baik, namun MAPE bukan satu-satunya kriteria dalam menentukan model estimasi terbaik.

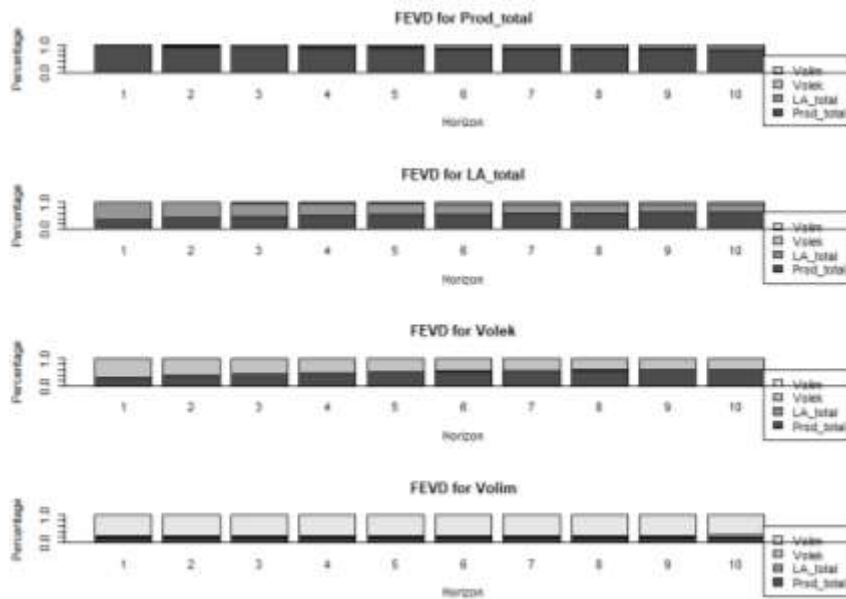
Salah satu kelebihan dari model VAR adalah dapat menampilkan *impulse respon* antar variabel. *Impulse respon* menggambarkan tingkat laju dari *shock* peubah yang satu terhadap peubah yang lainnya pada suatu rentang periode tertentu. Sehingga dapat dilihat lamanya

pengaruh dari *shock* suatu peubah terhadap peubah lain sampai pengaruhnya hilang atau kembali ke titik keseimbangan. *Impulse respon* produksi model VAR(1) type *none* ditunjukkan pada Gambar 15. Berdasarkan gambar tersebut terlihat bahwa perubahan produksi berdampak pada produksi itu sendiri, luas areal dan volume ekspor dalam jangka waktu yang cukup lama. Perubahan produksi lada juga berdampak pada volume impor lada namun hanya satu tahun ke depan. Setelah satu tahun, dampak perubahan produksi terhadap volume impor tidak lagi terasa.



Gambar 15. *Impulse Respon* Produksi pada Model VAR(1) *Type None*

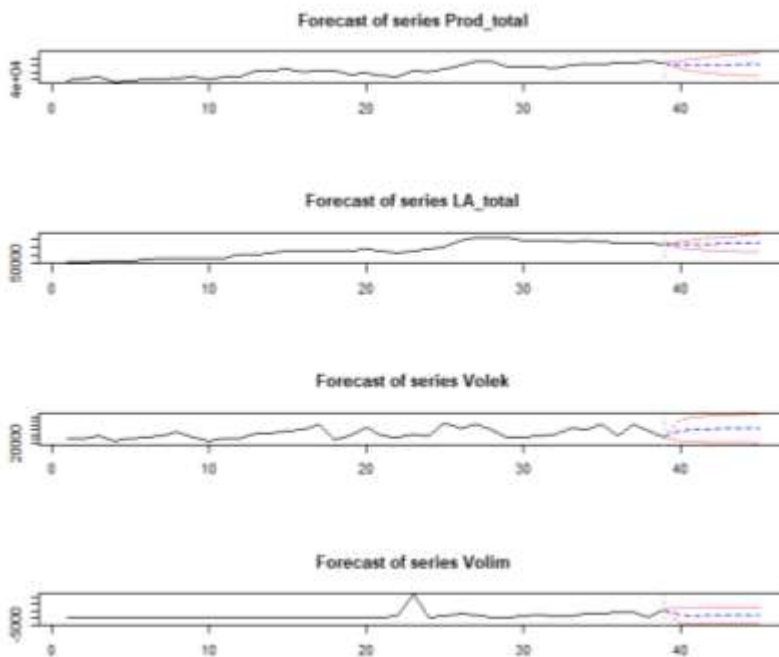
Selain *impulse respon*, model VAR juga menyajikan *variance decomposition* atau dekomposisi keragaman yang digunakan untuk melihat variabel apa saja yang mempengaruhi komposisi keragaman suatu variabel. Pada Gambar 16 terlihat bahwa komposisi produksi pada tahun pertama dipengaruhi sepenuhnya oleh produksi itu sendiri. Pada tahun kedua, komposisi produksi masih didominasi produksi itu sendiri serta dipengaruhi oleh luas areal. Semakin bertambah tahun, pengaruh produksi lada terhadap keragaman produksi lada sendiri semakin berkurang, namun diikuti dengan semakin tingginya pengaruh luas areal, volume ekspor dan volume impor lada.



Gambar 16. Dekomposisi Keragaman Model VAR(1) *Type None*

Model VAR(1) type none sebagai model terpilih, selanjutnya dilakukan pengepasan model pada data training dan data testing. Gambar 17 menunjukkan keseluruhan data training dan testing (pengepasan) beserta hasil ramalan dan plotnya. Setelah dipilih model VAR terbaik untuk peramalan, kemudian dilakukan peramalan produksi lada selama periode 2021-2025. Berdasarkan model VAR(1) type none, estimasi produksi lada selama lima tahun ke depan diramalkan naik 1,24% per tahun. Ramalan ini berbeda dengan perkembangan produksi lada selama lima tahun terakhir yang mengalami penurunan 0,06% per tahun. Pada tahun 2021 produksi lada diprediksi sebesar 88.103 ton dan meningkat hingga mencapai 89.276 ton pada tahun 2022. Tahun 2023 produksi lada kembali diperkirakan naik menjadi 90.351 ton. Peningkatan produksi juga terjadi pada tahun 2024 dan 2025, masing-masing diestimasi sebesar 91.443 ton dan 92.563 ton. Selain ramalan produksi lada, model VAR(1) *type none* ini juga mengestimasi variabel input lain yaitu luas areal, volume ekspor, dan volume impor

sebagaimana disajikan pada Gambar 17.



Gambar 17. Hasil Estimasi Produksi Lada dengan Model VAR(1) *Type None*

### Pemilihan Model Terbaik untuk Estimasi Produksi Lada

Kriteria pemilihan model estimasi terbaik didasarkan pada nilai terkecil dari MAPE data training dan data testing yaitu dengan memilih nilai MAPE terkecil, khususnya untuk MAPE data testing. Selain MAPE terkecil, kelogisan hasil ramalan juga perlu diperhatikan misalnya dengan membandingkan perkembangan produksi lada lima tahun terakhir dan hasil estimasi lima tahun ke depan. Selain itu, kelogisan hasil ramalan juga tercerminkan dari pola pergerakan hasil estimasinya. Plot data ramalan yang berhimpit/bersesuaian dengan data aktualnya memiliki performa hasil estimasi yang lebih baik.

Produksi lada selama lima tahun terakhir berfluktuasi dengan kecenderungan turun 0,06% per tahun. Tahun 2016 produksi lada tercatat 86.334 ton kemudian naik menjadi 87.991 ton pada tahun 2017. Tahun 2018 produksi lada kembali mengalami peningkatan menjadi 88.235 ton. Sejak tahun 2019 produksi lada turun menjadi 87.619 ton. Penurunan tersebut terus berlangsung hingga tahun 2020 dengan produksi tercatat sebesar 86.083 ton.

Dalam menyusun estimasi produksi lada selama periode 2021-2025, dilakukan uji coba menggunakan tiga model. Model pertama adalah ARIMA, dengan model terbaik yang dipilih yaitu ARIMA (1,1,1). Model ARIMA tersebut menghasilkan MAPE data training sebesar 8,98 dan MAPE data testing sebesar 1,98. MAPE data testing tersebut menunjukkan bahwa kemampuan model ARIMA (1,1,1) dalam melakukan estimasi produksi lada akan mengalami kesalahan sekitar 1,98% lebih tinggi atau 1,98% lebih rendah. Berdasarkan model ARIMA (1,1,1), hasil estimasi produksi lada selama lima tahun ke depan diramalkan naik 0,01% per tahun.

Model estimasi yang kedua adalah fungsi transfer, dengan model terbaik yang dipilih adalah fungsi transfer ARIMA (0,1,1). Dalam melakukan estimasi dengan fungsi transfer, output variabel yang digunakan adalah produksi lada sedangkan input variabelnya yaitu volume ekspor lada. Model fungsi transfer tersebut menghasilkan MAPE data training sebesar 8,06 dan MAPE data testing sebesar 1,88. MAPE data testing tersebut menunjukkan bahwa kemampuan model fungsi transfer ARIMA (0,1,1) dalam melakukan estimasi produksi lada akan mengalami kesalahan sekitar 1,88% lebih tinggi atau 1,88% lebih rendah. Berdasarkan model fungsi transfer ARIMA (0,1,1), hasil estimasi produksi lada selama lima tahun ke depan diramalkan naik 0,06% per tahun. Jika dibandingkan dengan model ARIMA (1,1,1), model fungsi transfer ARIMA (0,1,1) memiliki MAPE data training dan MAPE data testing yang lebih kecil.

Model estimasi yang ketiga adalah VAR, dengan model terbaik yang dipilih adalah VAR(1) *type none*. Dalam melakukan estimasi dengan model VAR(1) *type none*, output variabel yang digunakan adalah produksi lada sedangkan input variabelnya antara lain luas areal, volume ekspor dan volume impor. Model VAR tersebut menghasilkan MAPE data training sebesar 9,09 dan MAPE data testing sebesar 4,81. MAPE data testing tersebut menunjukkan bahwa kemampuan model VAR(1) dalam melakukan estimasi produksi lada akan mengalami kesalahan sekitar 4,81% lebih tinggi atau 4,81% lebih rendah. Berdasarkan model VAR(1), hasil estimasi produksi lada selama lima tahun ke depan diramalkan naik 1,24% per tahun. Jika dibandingkan dengan model ARIMA (1,1,1) dan model fungsi transfer

ARIMA (0,1,1), model VAR(1) *type none* memiliki MAPE data training dan MAPE data testing yang lebih besar.

Dari ketiga model estimasi yang dibandingkan, diperoleh informasi bahwa model fungsi transfer memberikan MAPE data training dan MAPE data testing paling kecil dibandingkan model lain (Tabel 1). Secara statistik, model fungsi transfer ARIMA (0,1,1) memiliki performa ramalan yang lebih baik dengan persentase kesalahan ramalan paling kecil dibandingkan model yang lain. Namun pada model fungsi transfer tersebut, produksi lada hanya dipengaruhi oleh volume ekspor lada. Dengan kata lain, model ini mengabaikan pengaruh dari variabel lain seperti intervensi pemerintah melalui program kebijakan intensifikasi maupun ekstensifikasi.

Tabel 1. Perbandingan Hasil Estimasi Produksi Lada dengan Model ARIMA, Fungsi Transfer dan VAR

Keterangan	Pengujian MAPE	Model ARIMA		Fungsi Transfer		Model VAR	
		ARIMA (1,1,1)	(%)	Arima (0,1,1) Xreg=Volume Ekspor	(%)	VAR (1) type=none	(%)
	MAPE Training	8,98		8,06		9,09	
	MAPE Testing	1,98		1,88		4,81	
ATAP	2016	86.334		86.334		86.334	
	2017	87.991	1,92	87.991	1,92	87.991	1,92
	2018	88.235	0,28	88.235	0,28	88.235	0,28
	2019	87.619	-0,70	87.619	-0,70	87.619	-0,70
	2020	86.083	-1,75	86.083	-1,75	86.083	-1,75
Angka Estimasi (AESTI)	2021	86.104		87.086		88.103	
	2022	86.119	0,02	87.126	0,05	89.276	1,33
	2023	86.130	0,01	87.386	0,30	90.351	1,20
	2024	86.137	0,01	87.360	-0,03	91.443	1,21
	2025	86.143	0,01	87.303	-0,07	92.563	1,22
Rata-rata Pertumbuhan (%)	ATAP 2016 - 2020		-0,06		-0,06		-0,06
	AESTI 2021 - 2025		0,01		0,06		1,24

Pada tahun 2021 Kementerian Pertanian melalui Direktorat Jenderal Perkebunan telah melakukan program perluasan tanam sebesar 475 ha dan rehabilitasi seluas 370 ha. Pada tahun 2022 perluasan lada ditargetkan sebesar 1.500 ha sedangkan rehabilitasi ditargetkan seluas 1200 ha. Untuk program intensifikasi seperti rehabilitasi tanaman, maka dampaknya dapat langsung dirasakan di tahun yang sama karena dilakukan untuk tanaman eksisting. Program ekstensifikasi seperti perluasan tanam, dampaknya terhadap peningkatan produksi lada baru terasa beberapa tahun ke depan. Idealnya, intervensi kebijakan pemerintah baik melalui ekstensifikasi maupun intensifikasi tidak boleh diabaikan ketika menyusun model estimasi produksi lada. Bagaimanapun juga, secara teknis intervensi tersebut berdampak langsung pada produksi lada di tahun 2021, tahun 2022 serta beberapa tahun ke depan. Oleh karena itu, analisis ini mempertimbangkan kombinasi antara model terbaik secara statistik serta intervensi pemerintah. Secara statistik, model terbaik yang direkomendasikan untuk estimasi produksi lada adalah fungsi transfer ARIMA (0,1,1). Di sisi lain, intervensi pemerintah yang dampaknya langsung terasa di tahun 2021 dan 2022 adalah rehabilitasi. Rehabilitasi tanaman lada bertujuan untuk meningkatkan produktivitas, sehingga produksi juga diharapkan akan naik.

Produktivitas lada Indonesia selama lima tahun terakhir (2016-2020) berkisar antara 0,7 - 0,8 ton/ha. Dengan mengasumsikan produktivitas lada tahun 2021-2022 sebesar 0,8 ton/ha, rehabilitasi tanaman seluas 370 ha pada tahun 2021 akan berkontribusi memberikan tambahan produksi sebanyak 296 ton. Tambahan produksi tersebut belum termasuk dampak dari perluasan tanam yang dilakukan sebelum tahun 2021. Pada tahun 2022 rehabilitasi tanaman lada sebesar 1.200 ha sehingga ditargetkan ada penambahan produksi sebanyak 960 ton yang berasal dari program rehabilitasi. Dengan menambahkan hasil estimasi model fungsi transfer dengan produksi yang berasal dari rehabilitasi, diperoleh estimasi produksi lada tahun 2021 dan 2022 kurang lebih sekitar 87.466 ton dan 88.326 ton. Hasil estimasi produksi di tahun 2021-2020 ini lebih besar jika dibandingkan dengan hasil estimasi dengan model ARIMA (1,1,1) dan fungsi transfer ARIMA (0,1,1), namun lebih kecil dibandingkan dengan hasil estimasi dengan model VAR(1). Mengingat hasil estimasi tersebut hanya mempertimbangkan dampak volume ekspor dan rehabilitasi namun belum mempertimbangkan penambahan dari



program pemerintah lainnya seperti perluasan, maka produksi lada di tahun 2021-2022 diprediksikan bisa lebih tinggi lagi dan kemungkinan mendekati hasil estimasi model VAR(1) *type none*. Oleh karena itu, model terbaik yang dipilih untuk estimasi produksi lada pada analisis ini adalah model VAR(1) *type none*. Hasil estimasi model VAR tersebut dianggap dapat mengakomodir intervensi pemerintah melalui program perluasan dan rehabilitasi, serta pengaruh variabel lain seperti volume ekspor. Selain itu, pembentukan model VAR(1) *type none* pada analisis ini telah mempertimbangkan dampak dari luas areal, volume ekspor dan volume impor terhadap produksi lada. Lada Indonesia sendiri merupakan komoditas yang paling banyak diekspor. Meskipun Indonesia juga mengimpor lada beberapa tahun terakhir, volume impor tersebut sangat kecil sehingga diasumsikan dampaknya tidak begitu signifikan terhadap produksi. Dasar pemilihan VAR(1) *type none* sebagai model terbaik untuk estimasi lada beserta perbandingan hasil estimasi dengan model ARIMA (1,1,1) dan fungsi transfer ARIMA (0,1,1) disajikan pada Tabel 2.

Tabel 2. Dasar Pemilihan Model Terbaik untuk Estimasi Produksi Lada

Tahun	Intervensi Pemerintah		Asumsi Produktivitas (Ton/Ha)	Estimasi Produksi (Ton)				
	Perluasan (Ha)	Rehabilitasi (Ha)		Tambahan Produksi dari Rehabilitasi	Fungsi Transfer ARIMA (0,1,1)	Fungsi Transfer + Rehabilitas	ARIMA (1,1,1)	VAR(1) Type None
(1)	(2)	(3)	(4)	(5)=(2)*(4)	(6)	(7)=(5)+(6)	(8)	(9)
2021	370	475	0,8	380	87.086	87.466	86.104	88.103
2022	1.200	1.500	0,8	1.200	87.126	88.326	86.119	89.276
2023	Data belum tersedia	Data belum tersedia	0,8	Data belum tersedia	87.386	Data belum tersedia	86.130	90.351
2024	Data belum tersedia	Data belum tersedia	0,8	Data belum tersedia	87.360	Data belum tersedia	86.137	91.443
2025	Data belum tersedia	Data belum tersedia	0,8	Data belum tersedia	87.303	Data belum tersedia	86.143	92.563
MAPE Training (%)					8,06	-	8,98	9,09
MAPE Testing (%)					1,88	-	1,98	4,81

## KESIMPULAN DAN SARAN

### Kesimpulan

Model terbaik yang terpilih untuk mengestimasi produksi lada adalah VAR(1) *type none* dengan produksi lada sebagai output variabel dan input variabelnya yaitu luas areal, volume ekspor, dan volume impor. Pemilihan model VAR(1) *type none* sebagai model terbaik untuk estimasi produksi lada tidak semata-mata didasarkan pada nilai MAPE data training dan MAPE data testing terkecil dibandingkan model lain. Namun juga dengan mempertimbangkan adanya intervensi pemerintah melalui program perluasan dan rehabilitasi tanaman lada. Dampak dari program tersebut direpresentasikan pada variabel luas areal yang menjadi salah satu input variabel pada model ini. Hasil estimasi produksi lada selama lima tahun ke depan dengan model VAR(1) *type none* cenderung meningkat dengan rata-rata pertumbuhan 1,24% per tahun. Pada tahun 2020 produksi lada di Indonesia sebesar 86.083 ton. Produksi tersebut diproyeksikan meningkat menjadi 88.103 ton di tahun 2021. Tahun 2022 produksi lada diperkirakan terus meingkat menjadi 89.276 ton. Peningkatan produksi lada di Indonesia diramalkan terus terjadi hingga tahun 2025.

### Saran

Penelitian ini membatasi estimasi produksi lada menggunakan tiga model yaitu ARIMA, fungsi transfer dan VAR. Meskipun berdasarkan metode statistik model fungsi transfer memiliki performa estimasi yang lebih baik dibandingkan dua model lainnya, model VAR dipilih menjadi model terbaik untuk mengestimasi produksi lada. Dasar pemilihan model VAR tersebut adalah adanya intervensi pemerintah pada tahun berjalan dalam peningkatan produksi lada yang direpresentasikan oleh salah satu input variabel pada model VAR yaitu luas areal. Intervensi tersebut dilakukan melalui program perluasan areal dan rehabilitasi tanaman yang dampaknya dirasakan secara langsung pada tahun berjalan. Oleh karena itu, pada penelitian berikutnya disarankan untuk mencoba model lain yang dapat mengakomodir dinamika program kebijakan pemerintah ke dalam model estimasi. Selain itu, variabel lain

selain yang telah digunakan pada penelitian ini dapat juga dicobakan untuk mengestimasi produksi lada.

## DAFTAR PUSTAKA

- Damanik, Sabarman. 2001. Analisis Penawaran dan Permintaan Lada Indonesia di Pasar Internasional. *Jurnal Litri*. 7(4):113-119.
- Ditjenbun.pertanian.go.id. (2019, 13 Agustus). IPC Pintu Masuk Negosiasi Perdagangan Lada Indonesia. Diakses pada 2 Desember 2020, dari <http://ditjenbun.pertanian.go.id/ipc-pintu-masuk-negosiasi-perdagangan-lada-indonesia/>
- Enders, W. 2004. *Applied Econometric Time Series*. USA: University of Alabama.
- Fatma, Hikmah, N., & Usman. 2020. Faktor-faktor yang Memengaruhi Produksi Lada di Desa Kongkomas Kecamatan Basidondo Kabupaten Tolitoli. *Jurnal Agrotech*. 10(1):35-40.
- Gujarati, D.N., & Porter, D.C. 2010. *Dasar-dasar Ekonometrika*. Jakarta: Salemba Empat, Buku 2, Edisi 5.
- Hamdani, Tety, E., & Eliza. 2015. Faktor-faktor yang Mempengaruhi Produksi Lada di Indonesia. *Jom Faperta*. 7(2):1-7.
- Kardinan, A., Laba, I.W., & Rismayani. 2018. Peningkatan Daya Saing Lada Melalui Budidaya Organik. *Perspektif*. 7(1):26-39.
- Kemala, Syarif. 2006. Strategi Pengembangan Sistem Agribisnis Lada untuk Meningkatkan Pendapatan Petani. *Perspekif*. 5(10): 48-54.

Kementerian Pertanian. 2020. Statistik Perkebunan Unggulan Nasional 2019-2021. Jakarta: Direktorat Jenderal Perkebunan Kementerian Pertanian.

Nurdjannah, Nanan. 2006. Perbaikan Mutu Lada Dalam Rangka Meningkatkan Daya Saing di Pasar Dunia. *Perspektif*. 5(1): 13-25.

Nursalam. 2020. Analisis Produksi dan Efisiensi Alokatif Usahatani Lada di Desa Ameroro Kecamatan Tinondo Kabupaten Kolaka Timur. *Agrimor*. 5(3):57-59.

Zikria, Roydatul. 2020. Outlook Lada. Jakarta: Pusat Data dan Sistem Informasi Pertanian.

# ANALISIS ESTIMASI PRODUKSI KAKAO INDONESIA MELALUI PENDEKATAN MODEL ARIMA, FUNGSI TRANSFER DAN VAR

## *Estimation Analysis of Indonesian Cocoa Production Through Model of ARIMA, Transfer Function And VAR Approach*

Yuliawati Rohmah, S.P., M.S.E

*Pusat Data dan Sistem Informasi Pertanian, Kementerian Pertanian  
Jalan Harsono RM No. 3 Gd D Lt 4, Ragunan-Jakarta Selatan, Indonesia  
Telp. (021) 7816384 Fax. (021) 7816385  
E-mail: yuliawati@pertanian.go.id*

### ABSTRAK

Kakao merupakan salah satu komoditas perkebunan strategis di Indonesia. Data produksi kakao yang dirilis oleh Direktorat Jenderal Perkebunan bersama BPS berupa data tahunan yang disajikan untuk kondisi 2 tahun yang lalu merupakan Angka Tetap (ATAP), satu tahun yang lalu merupakan Angka Sementara (ASEM) dan untuk tahun yang berjalan merupakan Angka Estimasi (AESTI). Metode yang digunakan untuk menyusun AESTI selama ini adalah Metode *Single Exponential Smoothing* (SES) dan *Double Exponential Smoothing* (DES). Metode untuk menghasilkan AESTI data perkebunan perlu dikaji kembali agar didapatkan metode yang lebih akurat, lebih objektif dan lebih baik secara statistik. Sehingga kajian ini bertujuan untuk melakukan pemodelan, membandingkan efektivitas dari pemodelan tersebut dan menentukan metode terbaik dalam mengestimasi produksi kakao. Metode yang diterapkan adalah metode *Autoregressive Integrated Moving Average* (ARIMA), fungsi transfer dan *Vector Auto Regression* (VAR) dengan menggunakan *software* program *RStudio*. Pemilihan model terbaik dilakukan dengan membandingkan besaran nilai *Mean Absolute Percentage Error* (MAPE) dan kerealistisan hasil estimasi dibandingkan dengan data series sebelumnya. Berdasarkan dari hasil estimasi dan nilai MAPE disimpulkan bahwa FT ARIMA (2,2,1) xreg = harga produsen adalah model terbaik untuk estimasi produksi kakao.

**Kata-Kata Kunci:** ARIMA, fungsi transfer, VAR, produksi, kakao

### Abstract

*Cocoa is one of the strategic estate crop commodities in Indonesia. Cocoa production data released by the Directorate General of Estate Crops together with BPS in the form of annual data presented for the conditions 2 years ago were Fixed Figures (ATAP), one year ago were Provisional Figures (ASEM) and for the current year were Estimated Figures (AESTI). . The methods used to develop AESTI so far are Single Exponential Smoothing (SES) and Double Exponential Smoothing (DES) methods. The method for producing AESTI estate crops data needs to be reviewed in order to obtain a method that is more accurate, more objective and statistically better. Therefore, this study aims to conduct modeling, compare the effectiveness of the modeling and determine the best*

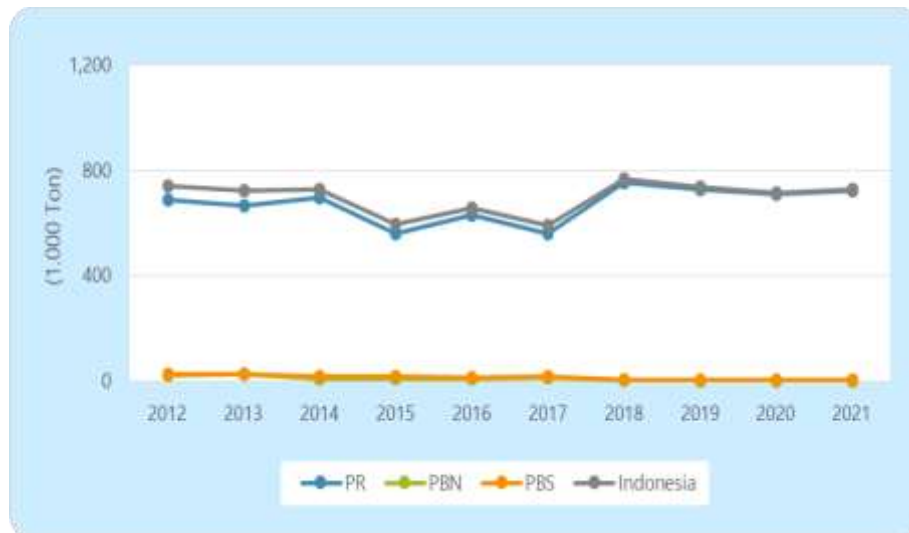
*method for estimating cocoa production. The method applied is the Autoregressive Integrated Moving Average (ARIMA), transfer function and Vector Auto Regression (VAR) using the RStudio software program. The selection of the best model is done by comparing the value of the Mean Absolute Percentage Error (MAPE) and the realism of the estimation results compared to the previous data series. Based on the estimation results and MAPE values, it is concluded that FT ARIMA (2,2,1) xreg = producer price is the best model for cocoa production estimation.*

*Key Words: ARIMA, transfer function, VAR, production, cocoa*

## **PENDAHULUAN**

Kakao merupakan salah satu komoditas ekspor perkebunan strategis yang ditetapkan pemerintah dan menempati posisi ketiga sebagai penyumbang devisa terbesar pada tahun 2020 setelah kelapa sawit dan karet. Berdasarkan data FAO selama tahun 2015-2019, Indonesia menempati posisi ketiga sebagai negara produsen kakao dunia dengan kontribusi sebesar 13,09% atau rata-rata produksi 0,68 juta ton per tahun terhadap total produksi dunia setelah Pantai Gading dan Ghana dengan masing-masing kontribusi mencapai 37,81% dan 16,99%.

Perkebunan kakao di Indonesia pada periode 10 tahun terakhir (2012-2021), didominasi oleh Perkebunan Rakyat (PR) sebesar 97,60% dan sisanya dimiliki oleh Perkebunan Besar Nasional (PBN) sebanyak 0,99% dan Perkebunan Besar Swasta (PBS) mencapai 1,41%. Daerah sentra produksi kakao di Indonesia berdasarkan rata-rata produksi dalam periode 5 tahun terakhir (2015-2019) pada urutan 3 teratas yang memberikan kontribusi diatas 10% terhadap total produksi nasional adalah Provinsi Sulawesi Tengah (17,34%), Sulawesi Selatan (16,51%) dan Sulawesi Tenggara (15,66%). Perkembangan produksi kakao Indonesia dalam 10 tahun terakhir berfluktuatif dengan rata-rata pertumbuhan hanya sebesar 0,59% per tahun (Gambar 1).



Gambar 1. Perkembangan Produksi Kakao Indonesia Menurut Status Pengusahaan Tahun 2012-2021

Saat ini, rilis resmi data produksi kakao dikeluarkan oleh Direktorat Jenderal Perkebunan bersama BPS berupa data tahunan yang disajikan untuk kondisi 2 tahun yang lalu (*lag n-2*) merupakan Angka Tetap (ATAP), satu tahun yang lalu (*lag n-1*) merupakan Angka Sementara (ASEM) dan untuk tahun yang berjalan merupakan Angka Estimasi (AESTI). Data statistik perkebunan yang diperoleh merupakan hasil sinkronisasi dan validasi yang dilaksanakan secara berjenjang dari tingkat kabupaten sampai tingkat nasional dengan mengacu pada Pedoman Pengelolaan Data Komoditas Perkebunan (PDKP) yang diterbitkan oleh Direktorat Jenderal Perkebunan. Saat ini dibutuhkan data yang terkini atau data *near real time* untuk perumusan kebijakan dan sebagai sarana peringatan dini atau *Early Warning System* (EWS) bagi para pengambil kebijakan agar dapat menghasilkan keputusan yang lebih efektif dan efisien.

Selama ini metode yang digunakan untuk menyusun AESTI adalah Metode Pemulusan Eksponensial Tunggal atau *Single Exponential Smoothing* (SES) dan Pemulusan Eksponensial Ganda atau *Double Exponential Smoothing* (DES). Pemilihan model terbaik dilakukan dengan membandingkan besaran nilai *Mean Absolute Percentage Error* (MAPE) dan kerealistisan hasil estimasi dibandingkan dengan data series sebelumnya. Dalam rangka melengkapi atau

menyempurnakan estimasi yang telah dihasilkan oleh Ditjen Perkebunan serta untuk memenuhi kebutuhan penyediaan data estimasi maka metode untuk menghasilkan AESTI data perkebunan perlu dikaji kembali agar didapatkan metode yang lebih akurat, lebih objektif dan lebih baik secara statistik dibandingkan metode yang telah dilakukan selama ini. Sejak tahun 2021 angka estimasi khususnya AESTI 2022 yang dihasilkan dari kerjasama antara Pusdatin, Ditjen Perkebunan dan BPS akan digunakan dalam publikasi Statistik Perkebunan Unggulan Nasional.

Berdasarkan hal di atas, maka kajian ini bertujuan untuk:

- d. Melakukan analisis dan estimasi data produksi kakao menggunakan model *Auto Regressive Integrated Moving Average* (ARIMA), Fungsi Transfer dan *Vector Auto Regression* (VAR).
- e. Membandingkan metode tersebut dalam memperoleh model estimasi data produksi kakao yang memiliki tingkat akurasi tertinggi.
- f. Menentukan metode terbaik dalam mengestimasi data produksi kakao.

## MATERI DAN METODE

Data yang digunakan dalam kajian ini adalah data sekunder *time series* tahunan. Wujud produksi kakao yang dibahas berupa biji kakao kering. Variabel, satuan, level, periode dan sumber data yang digunakan dalam kajian ini terdapat dalam Tabel 1.

Tabel 1. Jenis Variabel, Periode dan Sumber Data

No	Variabel Data	Satuan	Periode	Sumber
1	Produksi nasional	Ton	1980-2020	Ditjen Perkebunan, BPS
2	Luas areal nasional	Ha	1980-2020	Ditjen Perkebunan, BPS
3	Harga dunia internasional	US\$/Kg	1980-2020	World Bank
4	Harga produsen nasional	Rp/Kg	1996-2020	Ditjen Perkebunan



5	Volume ekspor nasional	Ton	1980-2020	BPS
6	Volume impor nasional	Ton	1980-2020	BPS

---

Variabel yang digunakan dalam metode ARIMA adalah produksi, sedangkan variabel harga produsen kakao nasional digunakan pada metode fungsi transfer sebagai variabel yang diduga dapat mempengaruhi produksi kakao nasional. Adapun pada metode VAR, variabel yang digunakan adalah produksi, luas areal, harga kakao dunia, volume ekspor dan volume impor karena kakao merupakan komoditas ekspor andalan. Volume ekspor dan volume impor yang digunakan pada proses pengolahan telah dilakukan konversi kedalam wujud produksi yang setara dengan biji kakao kering berdasarkan perhitungan konversi yang dikeluarkan oleh *International Cocoa Organization* (ICCO). Pertimbangan lain dalam pemilihan dan penggunaan variabel data dalam model adalah ketersediaan series data dalam jangka waktu yang cukup panjang.

Pada proses pengolahan dan analisis, data *time series* dibagi menjadi dua bagian yakni data *training* untuk penyusunan model periode tahun 1980-2014 dan sisanya sebagai data *testing* untuk validasi model periode tahun 2015-2020. Kemudian dari hasil data *training* disusun model dan dilakukan estimasi sesuai periode data *testing*, setelah itu dilakukan evaluasi kesesuaian ramalannya. Model terbaik dipilih dari berbagai alternatif metode estimasi yang dicoba dengan melihat nilai MAPE dan kesesuaian hasil estimasi dengan historis data aktualnya. Model estimasi terbaik yang terpilih kemudian dilakukan untuk estimasi 5 tahun ke depan yakni tahun 2021 – 2025 dengan menggabungkan seluruh data (*training* dan *testing*). Metode estimasi produksi kakao nasional yang digunakan dalam kajian ini terdiri dari model ARIMA, model Fungsi Transfer dan model VAR menggunakan *software* program *RStudio*.

### **1. Model Autoregressive Integrated Moving Average (ARIMA)**

ARIMA atau biasa disebut juga dengan metode time series Box Jenkins, sangat sesuai digunakan untuk melakukan peramalan jangka pendek, sementara untuk peramalan jangka

panjang kurang baik ketepatannya. Metode ARIMA merupakan metode yang hanya menggunakan variabel dependen dan mengabaikan variabel independen sewaktu melakukan peramalan.

Metode ARIMA dibagi kedalam tiga kelompok model, yaitu model *Auto Regressive* (AR), model *Moving Average* (MA) dan model campuran yang memiliki karakteristik kedua model di atas yaitu *Autoregressive Integrated Moving Average* (ARIMA).

### **Model *Auto Regressive* (AR)**

AR adalah suatu model yang menjelaskan pergerakan suatu variabel melalui variabel itu sendiri di masa lalu. Model *autoregressive* orde ke- $p$  dapat ditulis AR ( $p$ ) atau model ARIMA ( $p, d, 0$ ).

$$Y_t = \mu + \theta_1 Y_{t-1} + \theta_2 Y_{t-2} + \dots + \theta_p Y_{t-p} + \varepsilon_t$$

dimana:

- $Y_t$  = data *time series* sebagai variable dependen pada waktu ke-t
- $Y_{t-p}$  = data *time series* pada kurun waktu ke ( $t-P$ )
- $\mu$  = suatu konstanta
- $\theta_1 \dots \theta_p$  = parameter *autoregressive* ke-p
- $\varepsilon_t$  = nilai kesalahan pada waktu ke t

### **Model *Moving Average* (MA)**

MA adalah suatu model yang melihat pergerakan variabelnya melalui sisaannya di masa lalu. Bentuk model MA dengan ordo  $q$  dapat ditulis MA ( $q$ ) atau model ARIMA ( $0, d, q$ ).

$$Y_t = \mu - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$Y_t$  = data *time series* sebagai variable dependen pada waktu ke- $t$

$\phi_1 \dots \phi_q$  = parameter-parameter *moving average*

$\varepsilon_{t-q}$  = nilai kesalahan pada waktu ke- $(t-q)$

### ***Autoregressive Intergrated Moving Average (ARIMA)***

ARIMA merupakan model dari fungsi linier nilai lampau beserta nilai sekarang dan sisaan lampainya.

$$y_t = \mu + \theta_1 y_{t-1} + \theta_2 y_{t-2} + \dots + \theta_p y_{t-p} - \phi_1 \varepsilon_{t-1} - \phi_2 \varepsilon_{t-2} - \dots - \phi_q \varepsilon_{t-q} + \varepsilon_t$$

dimana:

$y_t$  = data *time series* sebagai variable dependen pada waktu ke- $t$

$y_{t-p}$  = data *time series* pada kurun waktu ke- $(t-p)$

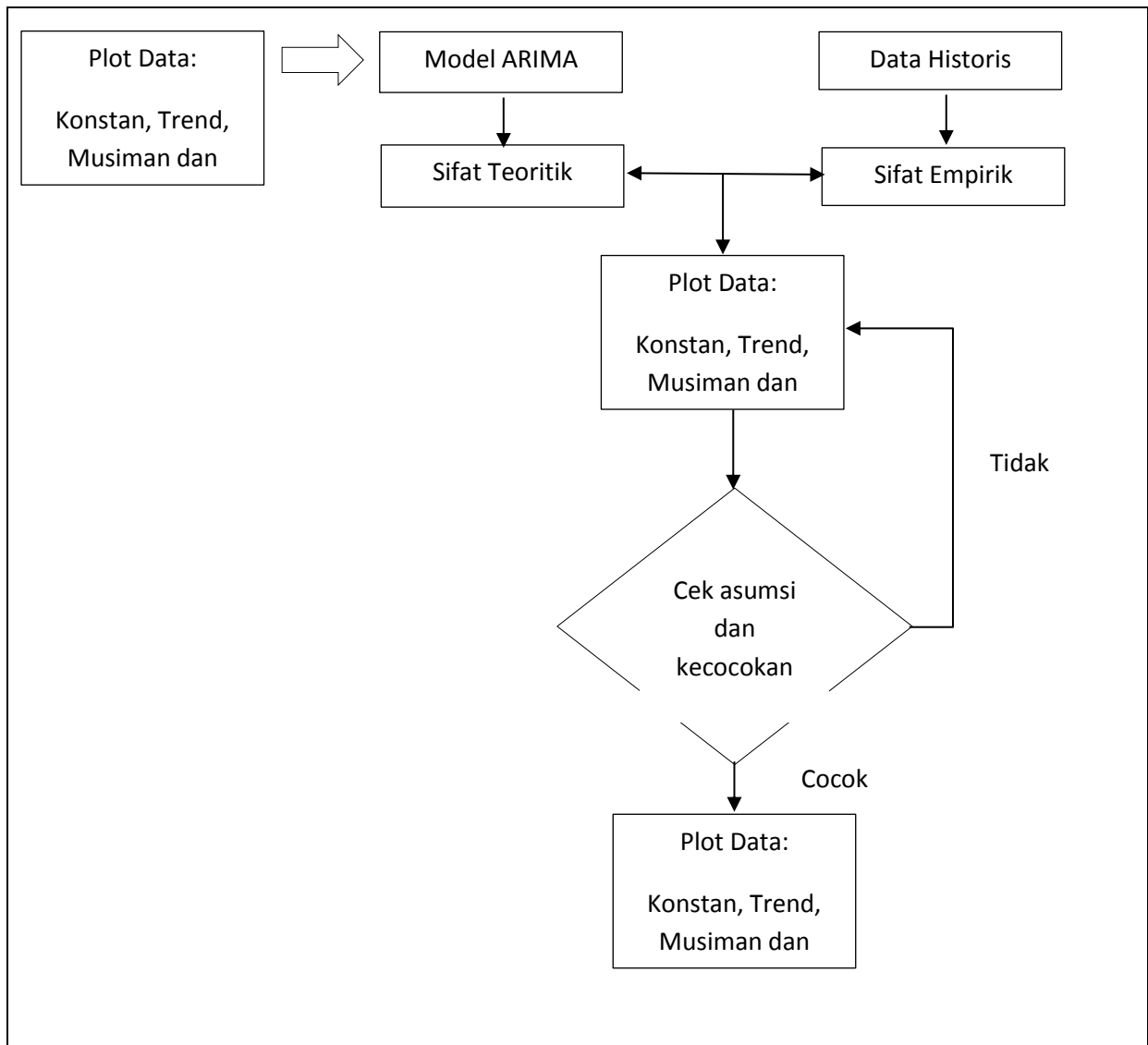
$\mu$  = suatu konstanta

$\theta_1 \theta_2 \phi_1 \phi_n$  = parameter-parameter model

$\varepsilon_{t-q}$  = nilai sisaan pada waktu ke- $(t-q)$

Tahapan estimasi pada model ARIMA dimulai dari uji kestasioneran data. Apabila data belum stasioner maka harus dilakukan proses *differencing* sampai diperoleh data yang stasioner. Tahapan selanjutnya adalah identifikasi model ARIMA, baik dengan *autoarima* maupun *armaselect*. Kemudian diikuti oleh serangkaian pengujian asumsi dan kecocokan, apabila telah memenuhi semua syarat pengujian maka estimasi dapat dilakukan, tetapi apabila

belum memenuhi syarat pengujian maka harus kembali ke tahapan sebelumnya yakni mengidentifikasi model ARIMA tentatif (Gambar 2).



Gambar 2. Tahapan Estimasi Model ARIMA

## 2. Model Fungsi Transfer

Model fungsi transfer adalah suatu model yang menggambarkan nilai prediksi masa depan dari suatu deret berkala (disebut deret *output* atau  $y_t$ ) didasarkan pada nilai-

nilai masa lalu dari deret itu sendiri ( $y_t$ ) dan didasarkan pula pada satu atau lebih deret berkala yang berhubungan (disebut deret input atau  $x_t$ ) dengan deret *output* tersebut. Model fungsi transfer merupakan fungsi dinamis yang pengaruhnya tidak hanya pada hubungan linier antara deret input dengan deret output pada waktu ke- $t$ , tetapi juga pada waktu  $t+1, t+2, \dots, t+k$ . Hubungan seperti ini pada fungsi transfer dapat menimbulkan delay (waktu senjang) antara variabel *input* dan variabel *output*.

Tujuan pemodelan Fungsi Transfer adalah untuk menetapkan model yang sederhana, yang menghubungkan deret output ( $Y_i$ ) dengan deret input ( $X_i$ ) dan gangguan/noise( $n_i$ ). Wei (1994) juga menjelaskan bahwa di dalam fungsi transfer terdapat rangkaian *output* yang mungkin dipengaruhi oleh rangkaian *multiple input*. Pada kasus *single input* variabel, dapat menggunakan metode korelasi silang yang dianjurkan oleh Box and Jenkins (1976). Teknik ini juga dapat digunakan ketika terdapat *single input* variabel yang lebih dari satu selama antar variable *input* tidak berkorelasi silang. Jika beberapa atau semua variabel input berkorelasi silang maka teknik *prewhitening* atau metode korelasi silang tidak dapat digunakan secara langsung. Alasan utama bagi perlunya suatu perencanaan atau peramalan adalah adanya tenggang waktu pengambilan keputusan yang dapat berkisar dari beberapa hari atau sampai beberapa tahun.

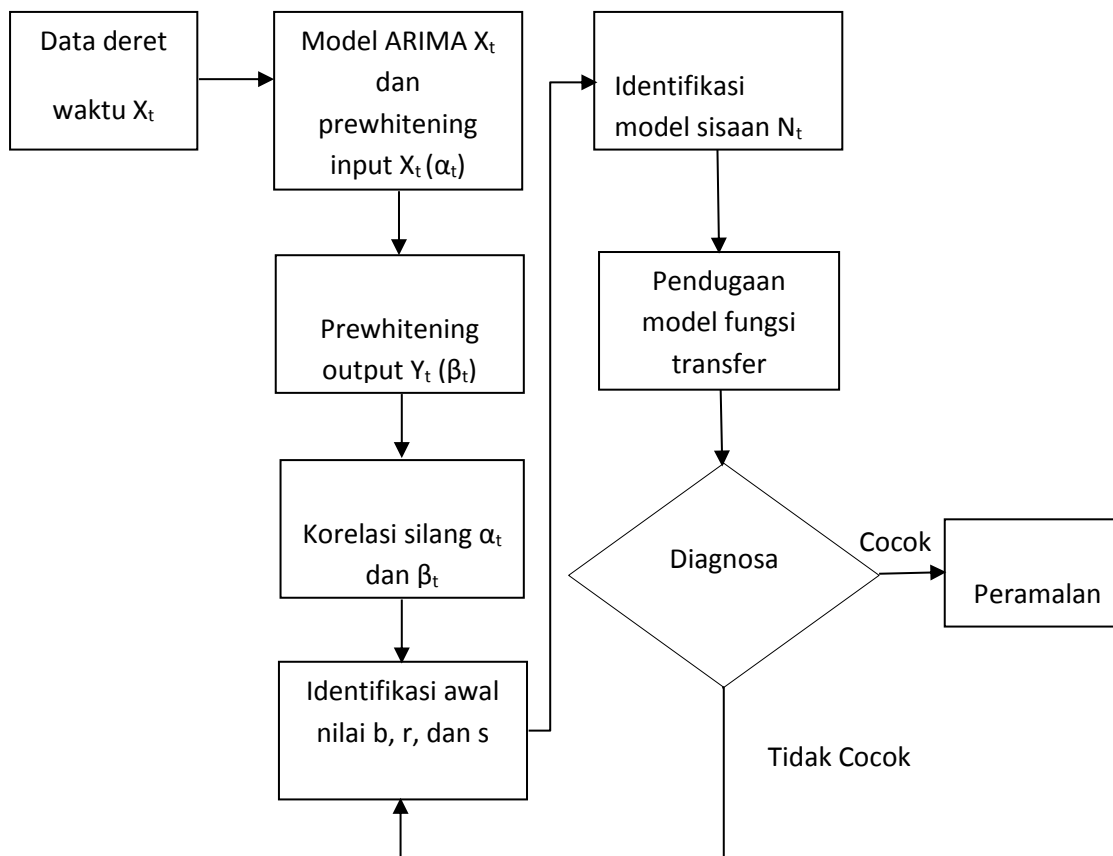
Pada analisis fungsi transfer untuk peramalan deret berkala *univariate*, terdapat deret berkala output yang diperkirakan dipengaruhi oleh deret berkala input dan input-input lain yang digabungkan dalam satu kelompok yang disebut gangguan (*noise*). Deret input mempengaruhi deret output melalui sebuah fungsi transfer yang mendistribusikan pengaruhnya secara dinamis melalui beberapa periode waktu yang akan datang dengan persentase tertentu yang disebut sebagai bobot *respons inputs* atau bobot Fungsi Transfer.

$$y_t = u(B)x_t + N_t \quad \Rightarrow \quad y_t = \frac{\omega_r(B)}{\delta_r(B)} x_{t-b} + \frac{\theta_q(B)}{\phi_p(B)} \varepsilon_t$$

dimana :

- $b$  = panjang jeda pengaruh  $x_t$  terhadap  $y_t$
- $r$  = panjang *lag*  $y$  periode sebelumnya yang masih mempengaruhi  $y_t$
- $s$  = panjang *lag*  $x$  periode sebelumnya yang masih mempengaruhi  $y_t$
- $p$  = ordo AR bagi *noise*  $N_t$
- $q$  = ordo MA bagi *noise*  $N_t$

Estimasi dengan menggunakan model fungsi transfer melalui serangkaian tahapan, mulai dari pemeriksaan kestasioneran input data dan pencarian model untuk variable input. Kemudian melakukan proses prewhitening dan korelasi silang antara data input dengan data output, pengepasan model awal, mengidentifikasi model sisaan atau *noise*, pengepasan model dengan *noise* sampai melakukan estimasi berbasis fungsi transfer (Gambar 3).



Gambar 3. Tahapan Estimasi Model Fungsi Transfer

### 3. Model VAR

*Vector Autoregression* (VAR) dikemukakan pertama kali oleh Christopher Sims (1980). Sims mengembangkan model ekonometri dengan mengabaikan pengujian asumsi secara apriori. Metode VAR memperlakukan seluruh variabel secara simetris tanpa memperlakukan variabel dependen dan independen atau dengan kata lain model ini memperlakukan seluruh variabel sebagai variabel endogen., karena pada kenyataannya suatu variabel dapat bertindak sebagai endogen atau dependen dari suatu variabel yang lainnya (Gujarati, 2010).

Kelebihan dalam penggunaan metode VAR (Gujarati, 2010):

- r. Kemudahan dalam penggunaan, tidak perlu mengkhawatirkan tentang penentuan variabel endogen dan variabel eksogen.
- s. Kemudahan dalam estimasi, metode *Ordinary Least Square* (OLS) dapat diaplikasikan pada tiap persamaan secara terpisah.
- t. *Forecast* atau peramalan yang dihasilkan pada beberapa kasus ditemukan lebih baik daripada yang dihasilkan oleh model persamaan simultan yang kompleks.
- u. Perangkat estimasi yang digunakan adalah *Impulse Respon Function* (IRF) untuk melacak respon dari variabel dependen dalam sistem VAR terhadap *shock* dari *error term* dan *Variance Decomposition* yang memberikan informasi mengenai pentingnya masing-masing *error term* dalam mempengaruhi variabel-variabel dalam VAR.

Di sisi lain, terdapat beberapa kelemahan model VAR (Gujarati, 2010) :

- a. Model VAR merupakan model yang *atheoritic* atau tidak berdasarkan teori, hal ini tidak seperti pada persamaan simultan.
- b. Pada model VAR penekanannya terletak pada *forecasting* atau peramalan sehingga model ini kurang cocok digunakan dalam menganalisis kebijakan.
- c. Permasalahan yang besar dalam model VAR adalah pada pemilihan *lag length* atau panjang lag yang tepat. Karena semakin panjang lag, maka akan menambah jumlah parameter yang akan bermasalah pada *degrees of freedom*.

- d. Variabel yang tergabung pada model VAR harus stasioner. Apabila tidak stasioner, perlu dilakukan transformasi bentuk data, misalnya melalui *first difference*.
- e. Sering ditemui kesulitan dalam menginterpretasi tiap koefisien pada estimasi model VAR, sehingga interpretasi dilakukan pada estimasi fungsi *impulse respon*.

Enders (2004) menjelaskan, ketika peneliti tidak memiliki kepastian untuk menentukan bahwa suatu variabel adalah eksogen, maka suatu perluasan analisis fungsi perpindahan alami akan memperlakukan masing-masing variabel secara simetris.

$$\begin{array}{c}
 \begin{array}{ccc}
 p & p & p \\
 x_t = \beta_{10} + \sum_{i=1}^p \alpha_{1i} x_{t-i} + \sum_{i=1}^p \beta_{1i} y_{t-i} + \sum_{i=1}^p \gamma_{1i} z_{t-i} + e_{1t} \\
 \\
 p & p & p \\
 y_t = \beta_{20} + \sum_{i=1}^p \alpha_{2i} x_{t-i} + \sum_{i=1}^p \beta_{2i} y_{t-i} + \sum_{i=1}^p \gamma_{2i} z_{t-i} + e_{2t}
 \end{array}
 \end{array}$$

dimana :

- $x_t, y_t, z_t$  = variabel endogen
- $\beta_0$  = vektor konstanta  $n \times 1$
- $\alpha, \beta, \gamma$  = parameter dari  $x, y,$  dan  $z$
- $p$  = panjang *lag*
- $t$  = waktu
- $\varepsilon$  = vektor dari *shock* masing-masing variabel

Untuk menguji kebaikan pada model VAR menggunakan kriteria  $R^2$  dan  $R^2$  *Adjusted*.  $R^2$  *squared* merupakan angka yang berkisar antara 0 sampai 1 yang mengindikasikan besarnya kombinasi variabel independen secara bersama–sama mempengaruhi nilai variabel dependen. Semakin mendekati angka satu, model yang dikeluarkan oleh regresi tersebut akan semakin baik. Secara manual,  $R^2$  merupakan rumus pembagian antara *Sum Squared Regression* dengan *Sum Squared Total*.



$$R^2 = \frac{SSR}{SST}$$

dimana :

SSR = Kuadrat dari selisih nilai Y prediksi dengan nilai rata-rata

$$Y = \sum (Y_{pred} - Y_{rata-rata})^2$$

SST = Kuadrat dari selisih nilai Y aktual dengan nilai rata-rata

$$Y = \sum (Y_{aktual} - Y_{rata-rata})^2$$

Sedangkan  $R^2$ -adjusted sudah mempertimbangkan jumlah sampel data dan jumlah variabel yang digunakan.  $R^2$ -adjusted merupakan  $R^2$  yang sudah dilengkapi.

$$R^2_{adj} = 1 - \left[ \frac{(1 - R^2)(n - 1)}{n - k - 1} \right]$$

dimana :

n = jumlah observasi

k = jumlah variabel

$R^2$ -adjusted akan menghitung setiap penambahan variabel dan mengestimasi nilai  $R^2$  dari penambahan variabel tersebut. Apabila penambahan pola baru tersebut ternyata memperbaiki model hasil regresi lebih baik dari pada estimasi, maka penambahan variabel tersebut akan meningkatkan nilai  $R^2$ -adjusted. Namun, jika pola baru dari penambahan variabel tersebut menunjukkan hasil yang kurang dari estimasinya, maka  $R^2$ -adjusted akan berkurang nilainya. Sehingga nilai  $R^2$ -adjusted tidak selalu bertambah apabila dilakukan penambahan variabel. Jika melihat dari rumus diatas, nilai  $R^2$ -adjusted memungkinkan untuk bernilai negatif, jika MSE-nya lebih besar dibandingkan (SST/p-1). Jika melihat rumus diatas, nilai  $R^2$ -adjusted pasti lebih kecil dibandingkan nilai  $R$  squared.

Tahapan dalam penyusunan model VAR diawali dari pembagian data series menjadi data *training* dan data *testing*. Tahapan berikutnya berupa pemilihan *lag* dan *type*, dilanjutkan dengan serangkaian pengujian asumsi. Kemudian melakukan estimasi untuk data *training*, data *testing*, dan penghitungan MAPE. Tahapan akhir melakukan pemilihan model terbaik, pengepasan model untuk seluruh data dan estimasinya serta interpretasi dari hasil *Impulse Response Function* (IRF) dan *Variance Decomposition*.

### **Kriteria Pemilihan Model Terbaik**

Untuk menguji kebaikan dan kelayakan suatu model yang digunakan dalam kajian ini berdasarkan nilai kesalahan dengan menggunakan statistik *Mean Absolute Percentage Error* (MAPE) atau kesalahan persentase absolut rata-rata yang diformulasikan sebagai berikut:

$$\text{MAPE} = \frac{1}{n} \sum_{t=1}^n \left| \frac{X_t - F_t}{X_t} \right| \cdot 100$$

dimana :

$X_t$  = data aktual

$F_t$  = nilai ramalan

Kriteria MAPE untuk membandingkan keseluruhan model menggunakan kriteria MAPE terkecil. Semakin kecil nilai MAPE maka model yang diperoleh semakin baik, karena makin mendekati nilai aktual.

### **Pengolahan Data**

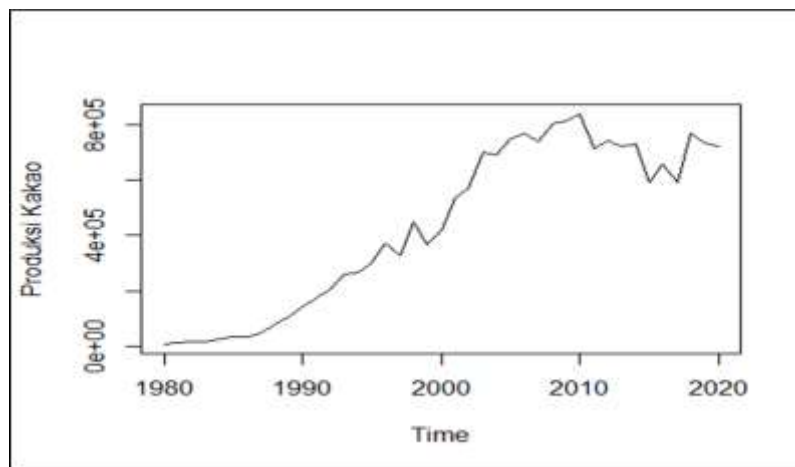
Pengolahan data yang dilakukan dalam kajian ini baik model ARIMA, model fungsi transfer maupun model VAR menggunakan Program R dan RStudio yang merupakan sebuah program komputasi statistika dan grafis.

## HASIL DAN PEMBAHASAN

### 1. Model ARIMA

#### Eksplorasi Data Produksi Kakao Indonesia

Produksi kakao dalam periode 41 tahun terakhir (1980-2020) berfluktuatif dengan rata-rata pertumbuhan sebesar 12,33% per tahun (Gambar 4). Berdasarkan uji kestasioneran data menggunakan *Augmented Dickey Fuller* (ADF) pada Tabel 2 memperlihatkan data produksi kakao tidak stasioner, karena nilai *test-statistic* (-0,59) lebih besar dibandingkan *critical value* pada tau3 (alpha 1%: -4,15; alpha 5%: -3,50; alpha 10%: -3,18) sehingga perlu dilakukan proses *differencing* 1. Hal ini diperkuat dengan plot produksi kakao berdasarkan sebaran datanya yang tidak konstan di sekitar rataaan (Gambar 4). Hasil *differencing* 1 produksi kakao telah bersifat stasioner karena nilai *test-statistic* (-3,13) lebih kecil dibandingkan *critical value* pada tau1 (alpha 1%: -2,62; alpha 5%: -1,95; alpha 10%: -1,61) seperti pada Tabel 3 dan sebaran datanya memiliki pola *single mean* atau konstan sekitar rataaan bukan nol (Gambar 5).



Gambar 4. Perkembangan Produksi Kakao Tahun 1980-2020

Tabel 2. Hasil Augmented Dickey Fuller Awal Data Produksi Kakao

---

Value of test-statistic is: -0.597 2.5991 0.6619			
Critical values for test statistics:			
	1pct	5pct	10pct
tau3	-4.15	-3.50	-3.18
phi2	7.02	5.13	4.31
phi3	9.31	6.73	5.61

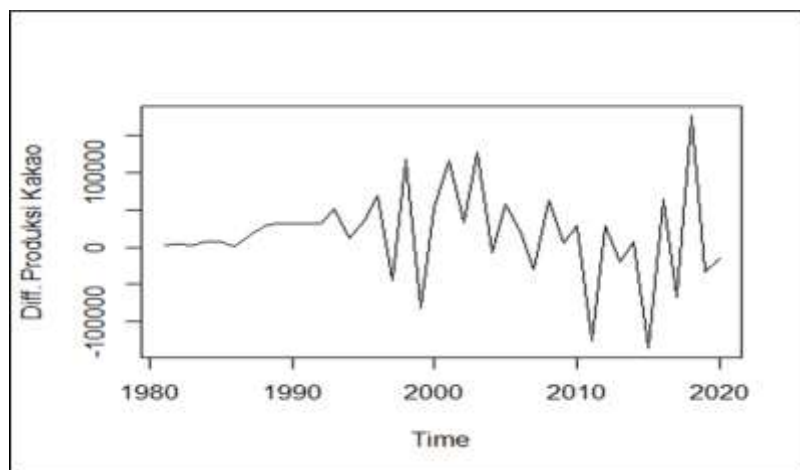
---

Tabel 3. Hasil Augmented Dickey Fuller Differencing1 Data Produksi Kakao

---

Value of test-statistic is: -3.1298			
Critical values for test statistics:			
	1pct	5pct	10pct
tau1	-2.62	-1.95	-1.61

---

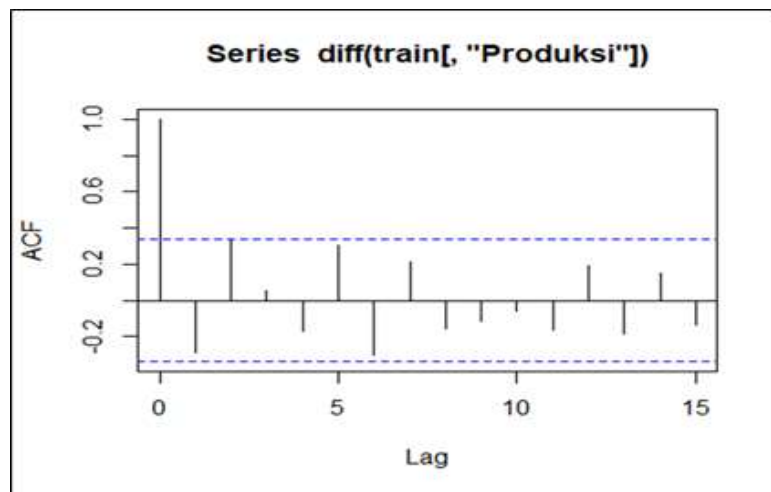


Gambar 5. Plot Data Produksi Kakao Differencing 1

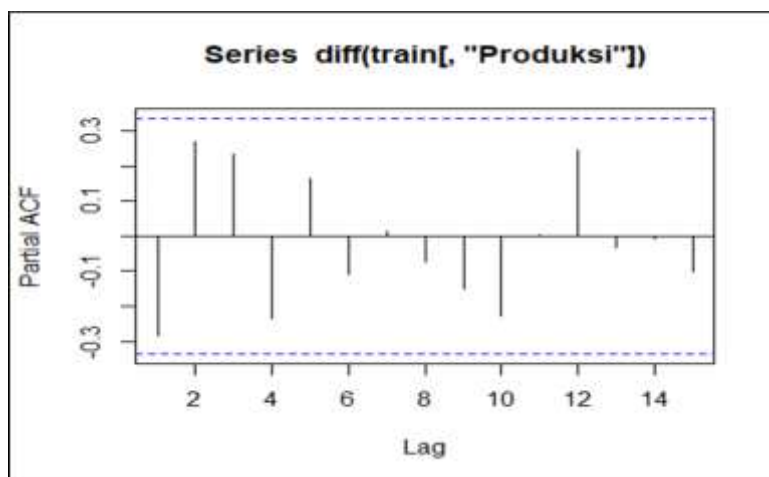
### Model ARIMA Produksi Kakao

Tahap awal pada metode estimasi dengan ARIMA, setelah dipastikan data bersifat stasioner, maka dilakukan identifikasi model ARIMA yang dapat diperoleh melalui 3 cara

yakni berdasarkan hasil plot *Auto Correlation Function* (ACF) dan *Partial Auto Correlation Function* (PACF), *Autoarima* maupun *Armaselect*. Dari hasil plot ACF bersifat *cut off* pada *lag* 1 (Gambar 6), sedangkan plot PACF, data sudah tidak memiliki pola khusus baik *tail off* ataupun *cut off* (Gambar 7) sehingga diperoleh dugaan awal untuk model ARIMA (1,1,0) atau ARIMA (0,1,1). Hal ini sesuai dengan hasil *autoarima* yakni ARIMA (1,1,0). Alternatif lain untuk mendapatkan model ARIMA dapat diperoleh dengan melakukan *overfitting* dari hasil *armaselect* dengan Uji *Minimum Information Criterion* (Minic) yang memberikan beberapa model alternatif (Tabel 5).



Gambar 6. Plot ACF Data Produksi Kakao Differencing 1



Gambar 7. Plot PACF Data Produksi Kakao Differencing 1

Setelah model dan hasil estimasi diperoleh, maka tahap berikutnya adalah mengevaluasi hasil estimasi baik dengan menggunakan Uji MAPE untuk data *training* dan data *testing* maupun dengan melihat kerealistisan hasil estimasi dengan data aktualnya. Model terbaik yang dipilih adalah ARIMA (5,1,0) karena hasil tes koefisiennya signifikan pada ar5 (alpha 10%) seperti tampak pada Tabel 4 serta memiliki hasil estimasi yang mendekati data historisnya. Pertimbangan lainnya adalah memiliki nilai MAPE *training* terkecil yakni 10,33% meskipun nilai MAPE *testing* bukan yang paling kecil yaitu 11,5%. Sedangkan hasil estimasi model lain dianggap terlalu rendah (*underestimate*) atau terlalu tinggi (*overestimate*) serta tidak selalu signifikan pada hasil tes koefisiennya meskipun nilai MAPE yang dihasilkan untuk data *training* maupun data *testing*-nya lebih kecil (Tabel 5).

Tabel 4. Hasil Test Coefficients Model ARIMA (5,1,0) Produksi Kakao

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )
ar1	-0.101951	0.159243	-0.6402	0.52203
ar2	0.384374	0.160056	2.4015	0.01633 *
ar3	0.126146	0.169187	0.7456	0.45591
ar4	-0.085807	0.168629	-0.5089	0.61085

ar5 0.322509 0.164582 1.9596 0.05005 .

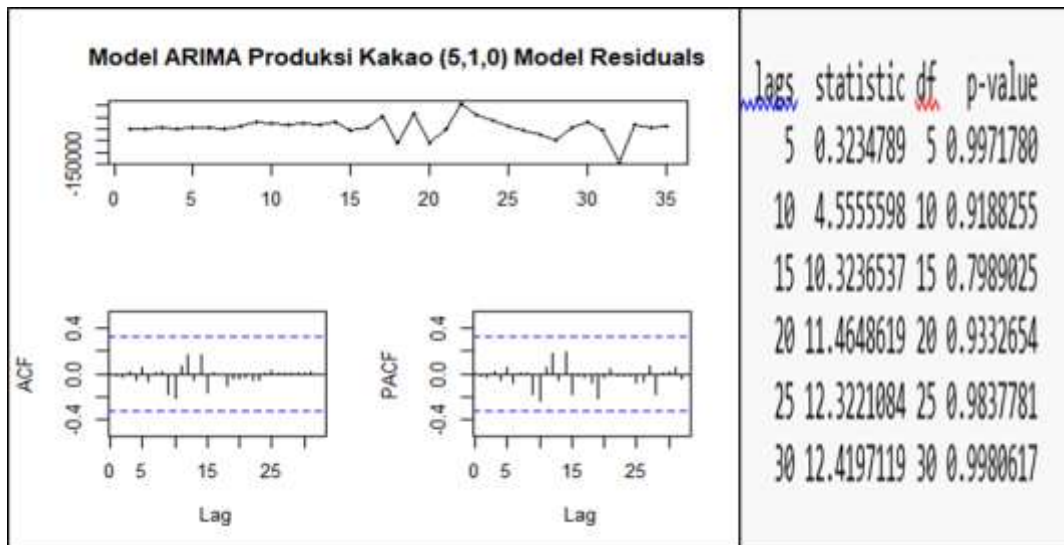
---

Signif. codes:0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

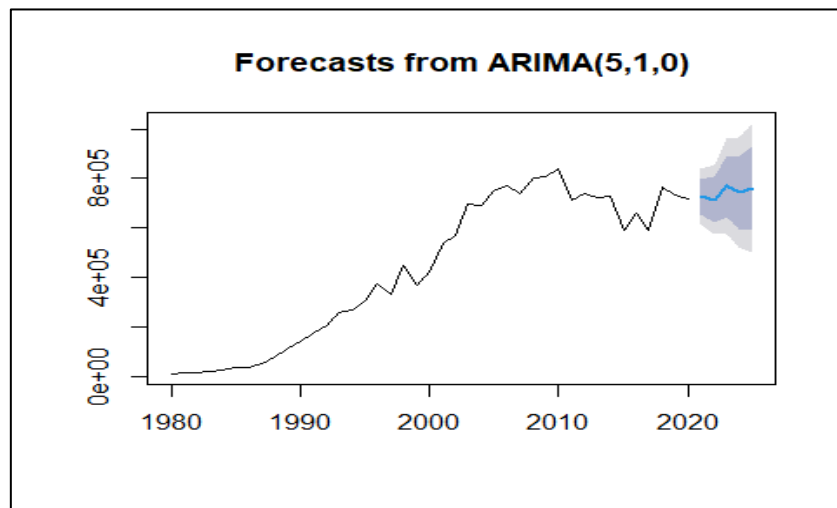
Tabel 5. Model ARIMA dan Estimasi Produksi Kakao Hasil Armaselect

No	Model	MAPE (%)		Hasil Estimasi Produksi Kakao (Ton)				
		Training	Testing	2021	2022	2023	2024	2025
1	ARIMA (1,1,0)	23.67	19.86	723,259	722,334	722,562	722,506	722,519
2	ARIMA (2,1,0)	11.37	10.2	709,787	705,588	702,628	701,520	700,593
3	ARIMA (3,1,0)	10.78	10.55	734,304	720,878	727,291	722,802	724,374
4	ARIMA (4,1,0)	10.62	11.88	733,067	727,301	729,986	728,588	729,617
5	ARIMA (5,1,0)	10.33	11.5	726,913	714,515	768,778	742,049	762,124
6	ARIMA (1,1,5)	10.41	10.84	760,119	716,131	800,121	772,033	758,548
7	ARIMA (2,1,5)	10.37	11.19	740,949	706,914	805,495	805,343	803,863
8	ARIMA (1,1,4)	10.82	10.73	737,799	703,602	802,271	804,403	805,497
9	ARIMA (0,1,2)	11.86	12.61	729,728	709,490	709,490	709,490	709,490

Langkah selanjutnya berupa pemeriksaan sisaan baik melalui plot sisaan serta plot ACF dan PACF sisaan. Hasil dari plot sisaan terdistribusi normal dan plot ACF serta PACF sisaan tidak nyata. Sedangkan dari hasil Uji *Ljung-Box*, autokorelasi sisaan tidak signifikan pada 30 lag (Gambar 8). Hasil estimasi produksi kakao dengan model ARIMA (5,1,0) untuk 5 tahun kedepan berfluktuatif dengan rata-rata pertumbuhan sebesar 1,23% per tahun (Gambar 9).



Gambar 8. Hasil Uji Pemeriksaan Sisaan dan Hasil Uji Ljung-Box ARIMA (5,1,0)



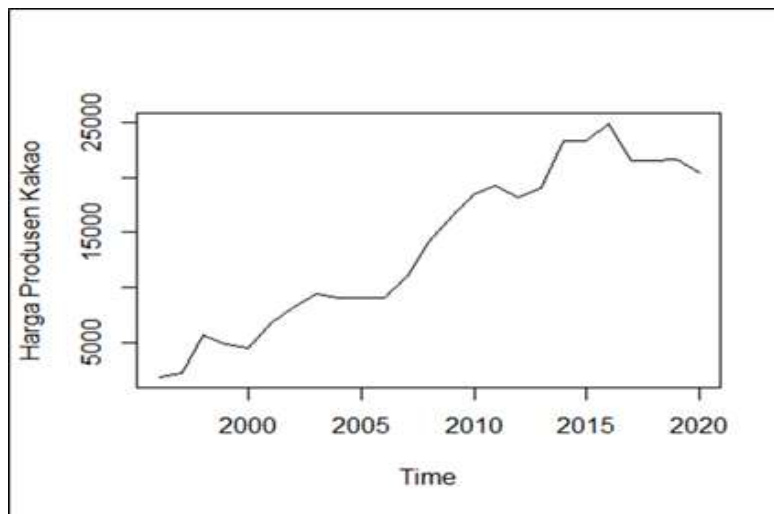
Gambar 9. Hasil Estimasi Produksi Kakao Model ARIMA (5,1,0) Tahun 2021-2025

## 2. Model Fungsi Transfer

Pada metode fungsi transfer, peubah input yang digunakan adalah harga kakao di tingkat produsen dengan pertimbangan naik turunnya harga kakao di tingkat produsen diduga sangat



mempengaruhi produksi kakao di dalam negeri. Perkembangan harga kakao di tingkat produsen dalam 25 tahun terakhir berfluktuatif dengan tren menaik seperti tampak pada Gambar 10.



Gambar 10. Perkembangan Harga Kakao Tingkat Produsen Tahun 1996-2020

Langkah pertama untuk proses analisis model adalah dengan mengidentifikasi model ARIMA peubah input berdasarkan hasil Uji ADF dan plot ACF serta PACF. Berdasarkan hasil Uji ADF pada Tabel 6 diketahui data peubah input non stasioner dimana nilai *test-statistic* (-2,38) lebih besar dibanding *critical values* pada tau3 (-4,38 pada alpha 1%; -3,60 pada alpha 5%; -3,24 pada alpha 10%), sehingga harus dilakukan proses *differencing*. Data harga kakao tingkat produsen stasioner setelah di-*differencing* 2 (Tabel 8) dengan nilai *critical value* lebih tinggi pada tau1 (alpha 1%: -2,66; alpha 5%: -1,95; alpha 10%: -1,6) dibandingkan nilai *test-statistic* (-3,51). Sedangkan pada *differencing* 1 data peubah input masih non stasioner karena nilai *test-statistic* (-1,98) masih lebih besar dibanding *critical values* pada tau1 (-2,66 pada alpha 1%) meskipun pada alpha 5% dan 10% sudah lebih kecil dari -1,95 dan -1,6 (Tabel 7).

Tabel 6. Hasil Augmented Dickey Fuller Awal Data Harga Kakao di Tingkat Produsen

---

Value of test-statistic is: -2.3765 4.3176 3.1017
---

---

---

Critical values for test statistics:

	1pct	5pct	10pct
tau3	-4.38	-3.60	-3.24
phi2	8.21	5.68	4.67
phi3	10.61	7.24	5.91

---

Tabel 7. Hasil Augmented Dickey Fuller Differencing 1 Data Harga Kakao di Tingkat Produsen

---

Value of test-statistic is: -1.9791

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.62	-1.95	-1.61

---

Tabel 8. Hasil Augmented Dickey Fuller Differencing 2 Data Harga Kakao di Tingkat Produsen

---

Value of test-statistic is: -3.5079

Critical values for test statistics:

	1pct	5pct	10pct
tau1	-2.62	-1.95	-1.61

---

Langkah kedua menduga model ARIMA peubah input baik dengan *autoarima* ataupun *armaselect*. Setelah melakukan *overfitting* dari berbagai kemungkinan model ARIMA peubah input, maka dipilih ARIMA (2,2,1) dengan pertimbangan hasil tes koefisiennya signifikan pada ar2 (alpha 10%) dan ma1 (alpha 0,1%) serta nilai MAPE 12,18%. Hasil tes koefisien model ARIMA peubah input terdapat pada Tabel 9.

Tabel 9. Hasil Test Coefficients Model ARIMA (2,2,1) Harga Kakao di Tingkat Produsen

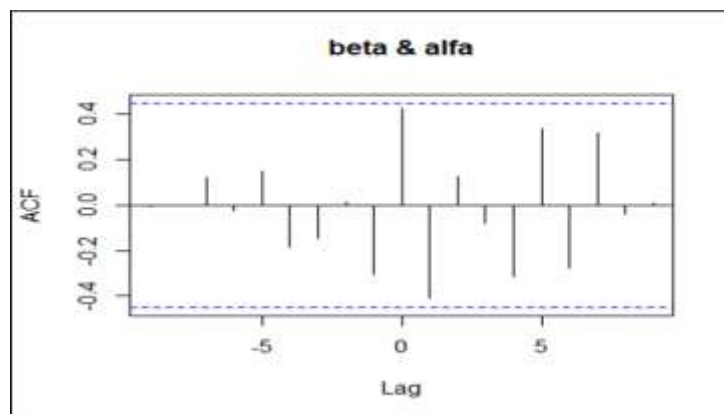
z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )
ar1	0.22661	0.25665	0.8829	0.3772727
ar2	-0.46239	0.26904	-1.7186	0.0856793
ma1	-0.90752	0.26164	-3.4686	0.0005231 ***

---

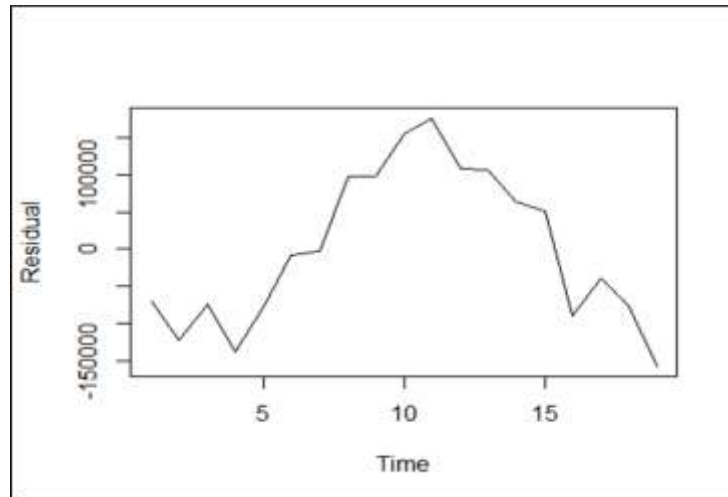
Signif. codes: 0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

Langkah ketiga yakni melakukan *prewhitening* dan korelasi silang antara deret input dengan produksi yang menghasilkan nilai r,s,b yakni (0,0,0) karena tidak ada yang nyata seperti tampak pada Gambar 11. Nilai b merupakan *lag* pertama kali dampak input berpengaruh terhadap output, s adalah *lag* berikutnya setelah b dimana input berdampak terhadap output, dan r merupakan pengaruh output terhadap dirinya sendiri. Pada Gambar 11 dapat dijelaskan bahwa nilai b dan s adalah 0 karena tidak ada yang nyata, sedangkan r dianggap 0 karena kakao merupakan tanaman tahunan.



Gambar 11. Plot Hasil Prewhitening dan Korelasi Silang Antara Deret Input dengan Output

Langkah keempat yaitu pengepasan model  $(r,s,b) = (0,0,0)$  yang menghasilkan nilai MAPE 15,53%. Identifikasi model *noise* atau residual dari peubah input merupakan langkah kelima yang dilakukan dengan model ARIMA seperti langkah kedua yang menghasilkan model ARIMA (2,2,1) sebagai model terpilih untuk residual. Plot dari nilai residual tampak pada Gambar 12.



Gambar 12. Residual Harga Kakao di Tingkat Produsen

Langkah selanjutnya melakukan pengepasan model  $(r,s,b) = (0,0,0)$  dan *noise* (2,2,1) dengan nilai MAPE 5,79% dan signifikansi pada ar2, ma1 dan xreg (Tabel 10).

Tabel 10. Hasil Test Coefficients Model Fungsi Transfer Produksi Kakao

z test of coefficients:

	Estimate	Std. Error	z value	Pr(> z )	
ar1	-1.57148	0.15762	-9.9702	< 2.2e-16	***
ar2	-0.85021	0.12912	-6.5849	4.553e-11	***
ma1	0.75236	0.25048	3.0036	0.002668	**
xreg	14.73448	6.19807	2.3773	0.017441	*
---					
Signif. codes:0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1					

Langkah ketujuh melakukan serangkaian estimasi dan membandingkan nilai MAPE data *testing* yang terdiri dari:

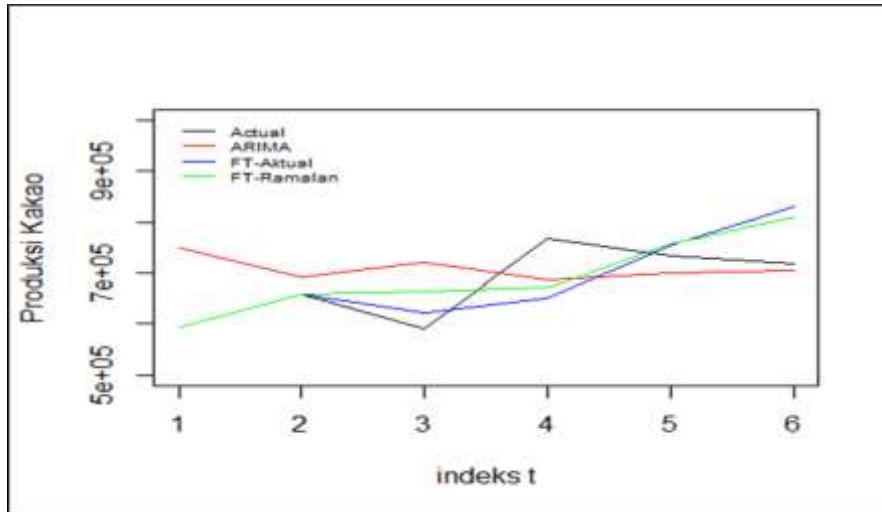
- a. Model fungsi transfer dengan data input nilai aktual.
- b. Model fungsi transfer dengan data input nilai estimasi.
- c. Model ARIMA output yang telah lebih dahulu dibahas pada Model ARIMA (5,1,0).

Nilai MAPE data *training* untuk model fungsi transfer yang menggunakan data input nilai aktual dan nilai ramalan sebesar 5,79% serta 10,33% nilai MAPE dari hasil model ARIMA (5,1,0) yang diperoleh dari model 1. Sedangkan Nilai MAPE data *testing* untuk model fungsi transfer yang menggunakan data input nilai aktual dan nilai ramalan hampir sama, yakni 6,43% untuk data input nilai aktual dan 6,74% untuk data input nilai ramalan. Sedangkan apabila dibandingkan dengan MAPE data *testing* model ARIMA (5,1,0) yang sebesar 11,50%, maka model fungsi transfer lebih baik (Tabel 11).

Tabel 11. Nilai MAPE Data Training Model Fungsi Transfer Produksi Kakao

No	Model Estimasi	MAPE (%)	
		Training	Testing
1	FT ARIMA (2,2,1) xreg=harga kakao produsen aktual	5,79	6,43
2	FT ARIMA (2,2,1) xreg=harga kakao produsen ARIMA (2,2,1)	5,79	6,74
3	ARIMA (5,1,0)	10,33	11,50

Berdasarkan plot hasil estimasi dengan fungsi transfer baik data input nilai ramalan maupun data input nilai aktual lebih mengikuti pola data aktual dibandingkan model ARIMA (Gambar 13).

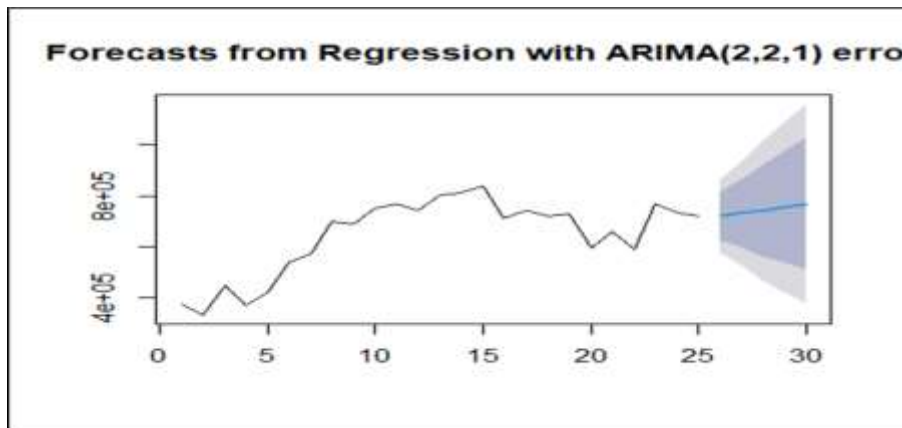


Gambar 13. Perbandingan Hasil Estimasi Data Testing Pada Model Fungsi Transfer

Langkah kedelapan yang merupakan langkah terakhir adalah menduga ulang model input harga kakao tingkat produsen dengan model ARIMA (2,2,1) dan menduga ulang fungsi transfer ARIMA (2,2,1) untuk melakukan estimasi produksi kakao 5 tahun kedepan yang menunjukkan pertumbuhan sekitar 1,35% per tahun secara rata-rata meskipun berfluktuatif setiap tahunnya (Tabel 12 dan Gambar 14).

Tabel 12. Hasil Estimasi Produksi Kakao Model FT ARIMA (2,2,1) xreg=Harga Produsen Tahun 2021-2025

Tahun	Produksi Kakao (Ton)
2021	719.782
2022	731.861
2023	742.262
2024	756.330
2025	769.297
Rata-rata Pertumbuhan (%)	1,35



Gambar 14. Hasil Estimasi Produksi Kakao Model FT ARIMA (2,2,1)  
 xreg = Harga Kakao di Tingkat Produsen Tahun 2021-2025

### 3. Model VAR

Pada model VAR, variabel yang digunakan terdiri dari produksi, luas areal, harga kakao dunia, volume ekspor dan volume impor. Volume ekspor dan volume impor yang digunakan pada proses pengolahan telah dilakukan konversi kedalam wujud produksi yang setara dengan biji kakao kering berdasarkan perhitungan konversi yang dikeluarkan oleh *International Cocoa Organization* (ICCO). Tahap awal dalam penentuan model VAR adalah melakukan penelusuran model dari *lag* atau  $p=1$  sampai dengan  $p=4$  dengan dan tanpa tren yang terdiri dari 3 tipe yakni *trend*, *const*, dan *both*. Dengan Melakukan *overfitting* dari semua kemungkinan model yang ada, hasil model terpilih yakni VAR (2) *type=const* dengan pertimbangan memenuhi serangkaian pengujian serta memiliki nilai MAPE untuk data *training* dan data *testing* yang paling kecil (Tabel 13).

Tabel 13. Model VAR Produksi Kakao

No.	Tipe	Variabel Produksi	Total Variabel	Tipe Signifikan	R <sup>2</sup> (%)	R <sup>2</sup> adjusted (%)
1	Both p=1	3	12	Trend	98.16	97.76
2	Both p=2	2	17	Trend dan Const	99.07	98.58
3	Both p=3	2	17	Trend dan Const	99.58	99.13
4	Both p=4	7	33	Trend dan Const	99.78	99.28
6	Const p=1	4	15	Const (2)	97.94	97.57
7	Const p=2	5	17	Const (3)	98.85	98.33
8	Const p=3	2	16	Const (2)	99.53	99.09
9	Const p=4	9	37	Const (4)	99.67	99.01
11	Trend p=1	4	17	Trend (3)	99.37	99.24
12	Trend p=2	2	12	Trend (2)	99.67	99.5
13	Trend p=3	3	15	Trend (2)	99.83	99.67
14	Trend p=4	4	19	Trend (2)	99.93	99.78

Tahap selanjutnya melakukan serangkaian pengujian terhadap model yakni normalitas sebaran, autokorelasi dan keragaman. Dari hasil Uji *Chi-squared*, Uji *Jarque-Bera* dan ARCH dapat disimpulkan asumsi normalitas dan non autokorelasi terpenuhi serta ragam homogen (Tabel 14).

Tabel 14. Hasil Uji Asumsi Model VAR Produksi Kakao

Portmanteau Test (asymptotic) data: Residuals of VAR object varkakao.c2 Chi-squared = 324.27, df = 350, p-value = 0.8345
JB-Test (multivariate) data: Residuals of VAR object varkakao.c2 Chi-squared = 47.607, df = 10, p-value = 7.319e-07
Skewness only (multivariate) data: Residuals of VAR object varkakao.c2 Chi-squared = 18.673, df = 5, p-value = 0.002211
Kurtosis only (multivariate) data: Residuals of VAR object varkakao.c2 Chi-squared = 28.934, df = 5, p-value = 2.389e-05
ARCH (multivariate)



---

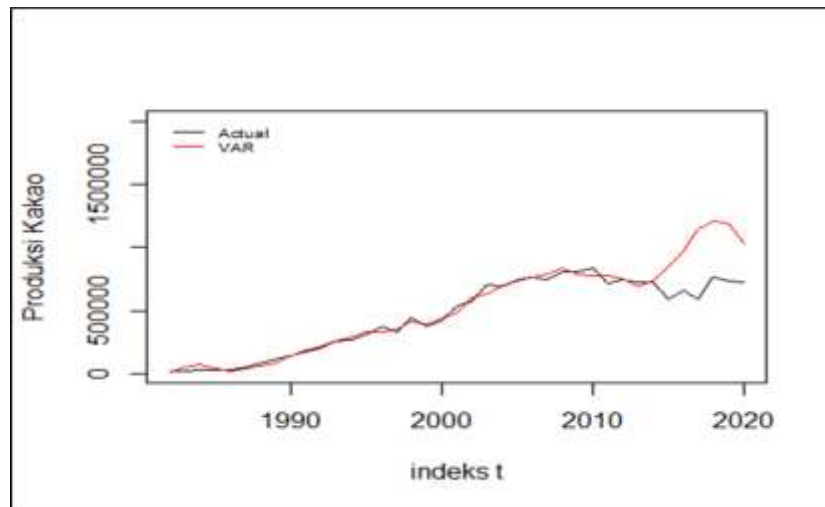
data: Residuals of VAR object varkakao.c2  
Chi-squared = 420, df = 1125, p-value = 1

---

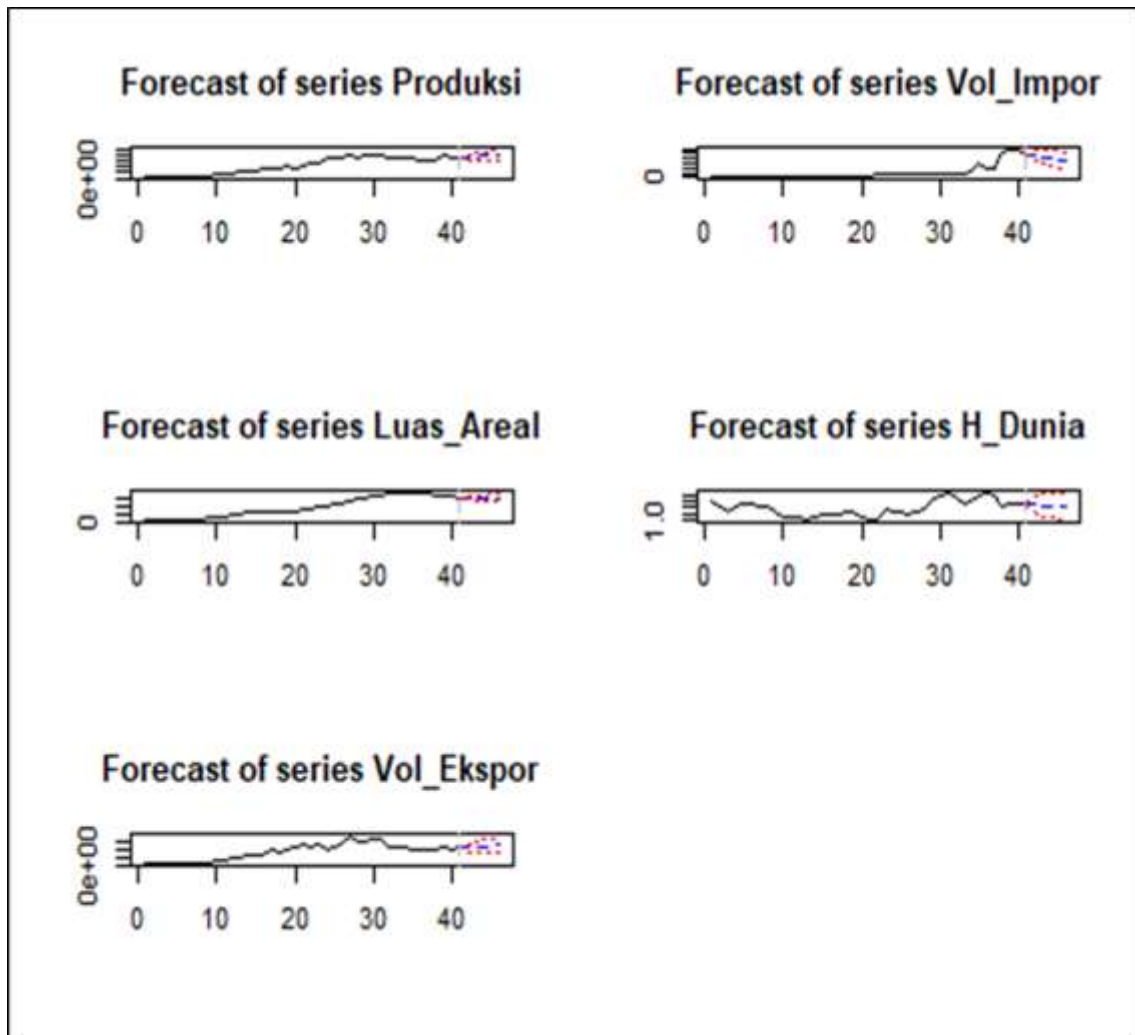
Tahap berikutnya dalam model VAR adalah menghitung nilai MAPE dari data *training* dan data *testing* seperti yang tampak di Tabel 15. Untuk MAPE data *training* lebih kecil 17,75% dibandingkan MAPE data *testing* yakni 57,75%. Pada Gambar 15 terlihat hasil plot estimasi data *training* dan data *testing* terhadap data aktual, dimana data *training* lebih mengikuti pola data aktual dibandingkan hasil estimasi data *testing*.

Tabel 15. Nilai MAPE Data Training dan Data Testing Model VAR (2) Type=Const Produksi Kakao

Data	MAPE (%)
Training	17,75
Testing	57,75



Gambar 15. Plot Data Ramalan Model VAR(2) Type=Const Terhadap Data Aktual Produksi Kakao Tahun 1980-2020

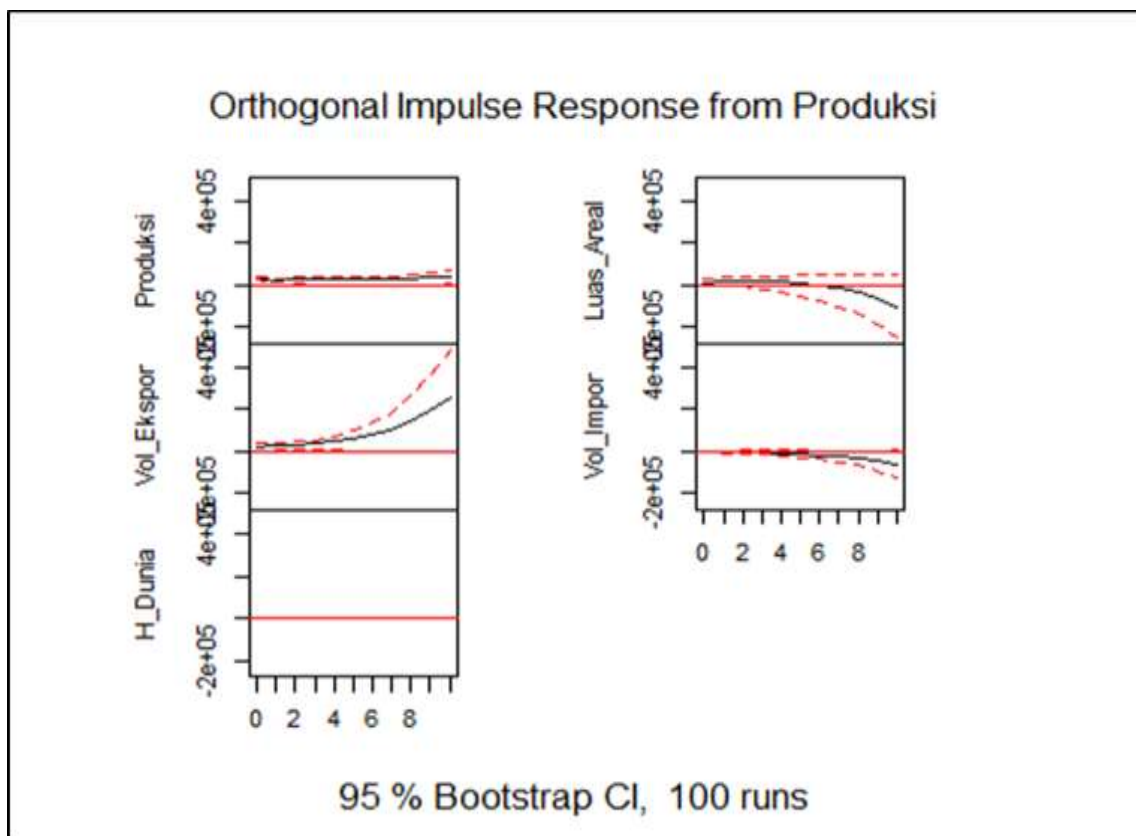


Gambar 16. Plot Estimasi Produksi Kakao  
Model VAR (2) Type=Const

Tahap akhir dari serangkaian tahapan pada proses pemodelan dengan metode VAR berupa estimasi produksi kakao untuk periode tahun 2021-2025 yang menduga akan terjadi kenaikan produksi kakao secara kontinyu pada 5 tahun mendatang dengan rata-rata bertumbuh sebesar 2,12% per tahun. Produksi kakao tahun 2021 diestimasi sebesar 741.650 ton, naik menjadi 750.229 ton di tahun 2022, naik kembali di tahun 2023 menjadi 767.732 ton. Sedangkan di tahun 2024 dan 2025, produksi naik menjadi 781.434 ton dan 799.025 ton (Tabel 16). Plot estimasi produksi kakao Model VAR (2) *Type = Const* memperlihatkan grafik produksi yang meningkat dalam lima tahun kedepan (Gambar 16).

Tabel 16. Hasil Estimasi Produksi Kakao Model VAR (2) Type=Const Tahun 2021-2025

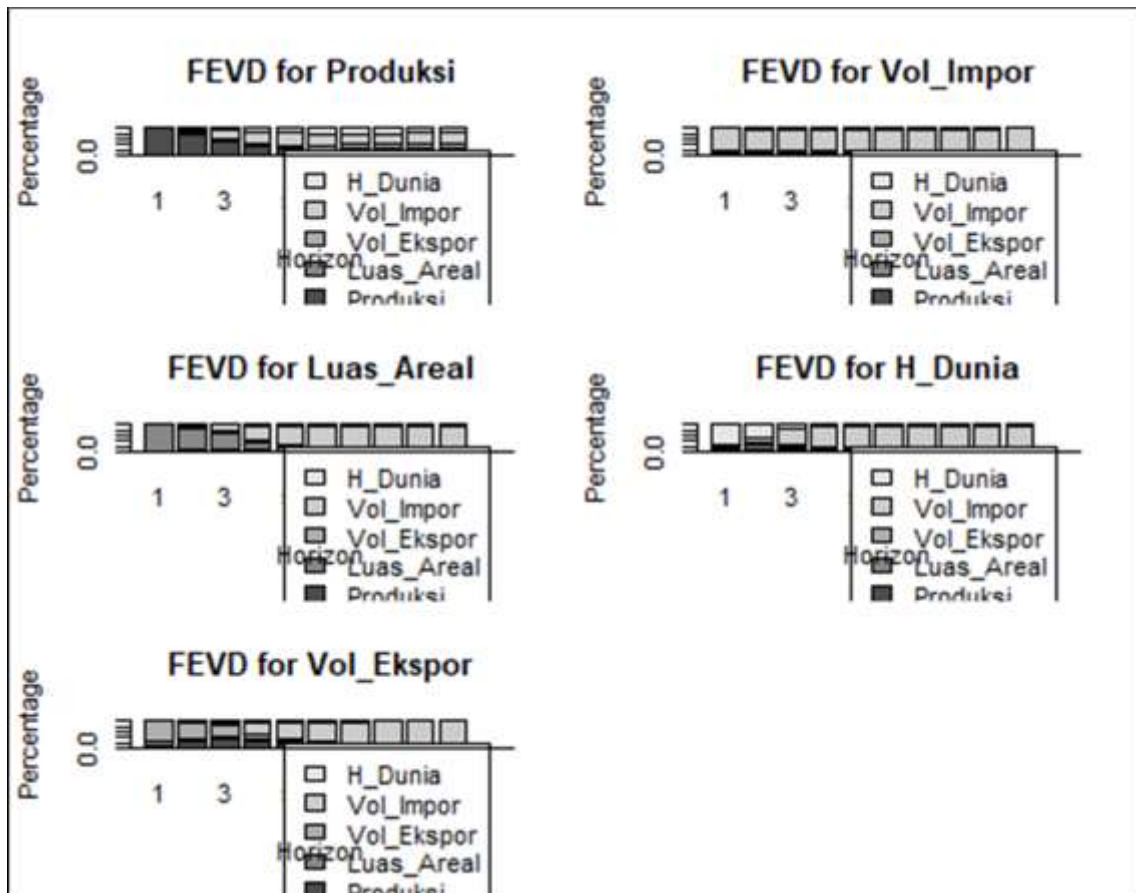
Tahun	Produksi Kakao (Ton)
2021	741.650
2022	750.229
2023	767.732
2024	781.434
2025	799.025
Rata-rata Pertumbuhan (%)	2,12



Gambar 17. Plot Orthogonal Impulse Response Function Produksi Kakao Dari hasil estimasi dengan Model VAR, juga diperoleh *Impulse Response Function* (IRF) dan *Variance Decomposition*. Berdasarkan hasil IRF produksi model terbaik VAR (2) *type = const*, terlihat bahwa jika terjadi perubahan pada produksi di tahun tertentu maka hanya akan

berdampak pada volume ekspor. Sedangkan dampak perubahan produksi tidak terjadi pada luas areal, harga dunia, volume impor dan produksi itu sendiri (Gambar 17).

Dari grafik dekomposisi keragaman model terbaik VAR (2) *type = const* dapat dilihat bahwa komposisi produksi pada tahun pertama dipengaruhi sepenuhnya oleh produksi itu sendiri. Pada tahun kedua, komposisi produksi secara mayoritas masih dipengaruhi oleh produksi itu sendiri serta sedikit dari luas areal. Pada tahun ketiga, komposisi produksi dipengaruhi 50% oleh produksi itu sendiri, 50% lainnya dipengaruhi oleh luas areal, volume ekspor dan volume impor. Semakin bertambahnya tahun, pengaruh produksi kakao terhadap keragaman produksi kakao sendiri semakin berkurang diikuti makin tingginya pengaruh variabel lain yakni luas areal, volume ekspor kakao, volume impor kakao dan harga kakao dunia (Gambar 18).

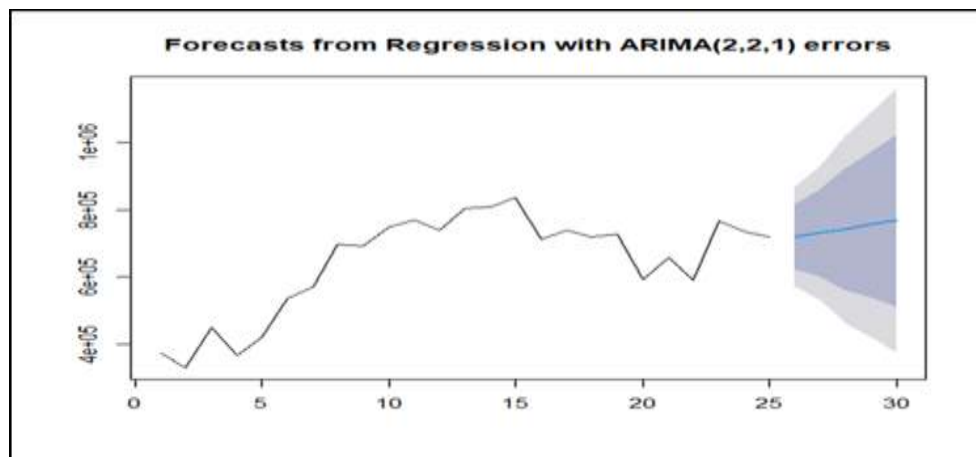


Gambar 18. Dekomposisi Keragaman Model VAR (2) Type=Cost Produksi Kakao Model Estimasi Terbaik

Berdasarkan Uji MAPE dari Tabel 17 maka model terbaik dan terpilih untuk estimasi produksi kakao adalah model FT ARIMA (2,2,1) xreg = harga produsen karena memiliki nilai MAPE terkecil untuk *training* 5,79% dan *testing* 6,74%. Model ARIMA menghasilkan nilai MAPE sebesar 10,33% untuk data *training* dan 11,5% untuk data *testing*. Sedangkan model VAR menghasilkan nilai MAPE paling tinggi dibandingkan model lainnya, yaitu 17,75% untuk data *training* dan 57,75% untuk data *testing*. Pertimbangan lainnya karena hasil estimasi dan rata-rata pertumbuhan yang dianggap lebih mendekati data aktualnya, dibandingkan dua model lainnya.

Tabel 17. Ringkasan Hasil Analisis Model Estimasi Produksi Kakao

	Pengujian MAPE	Model ARIMA		Fungsi Transfer		Model VAR		
		ARIMA (5,1,0)	(%)	Arima (2,2,1) Xreg=harga produsen	(%)	VAR (2) type=constant	(%)	
		MAPE Training	10.33		5.79		17.75	
		MAPE Testing	11.50		6.74		57.75	
ATAP	2016	658,399		658,399		658,399		
	2017	590,683	-10.28	590,683	#####	590,683	#####	
	2018	767,280	29.90	767,280	29.90	767,280	29.90	
	2019	734,797	-4.23	734,797	-4.23	734,797	-4.23	
	2020	719,498	-2.08	719,498	-2.08	719,498	-2.08	
Angka Estimasi (AESTI)	2021	726,913	1.03	719,782	0.04	741,650	3.08	
	2022	714,515	-1.71	731,861	1.68	750,229	1.16	
	2023	768,778	7.59	742,263	1.42	767,732	2.33	
	2024	742,049	-3.48	756,330	1.90	781,434	1.78	
	2025	762,124	2.71	769,297	1.71	799,025	2.25	
Rata-rata	ATAP 2016 - 2020		3.32		3.32		3.32	
Pertumbuhan	AESTI 2021 - 2025		1.23		1.35		2.12	



Gambar 19. Plot Hasil Ramalan Produksi Kakao Model FT ARIMA (2,2,1)  
xreg = Harga Produsen Tahun 2021-2025

Hasil estimasi dari model terbaik untuk produksi kakao tahun 2021-2025 akan terus meningkat meskipun dengan pertumbuhan yang berfluktuatif (Gambar 19). Tahun 2021, produksi meningkat sebesar 719.782 ton, naik 0,04% dibandingkan tahun 2020 yang mencapai 719.498 ton. Produksi kakao meningkat kembali sebesar 1,68% menjadi 731.861 ton di tahun 2022. Untuk tahun 2023-2025 estimasi produksi kakao terus meningkat meskipun dengan

pertumbuhan yang berfluktuatif yakni 742.263 ton (1,42%) di tahun 2023 kemudian naik 1,90% atau 756.330 ton di tahun 2024 dan turun kembali di tahun 2025 menjadi 1,71% atau 769.297 ton. Rata-rata pertumbuhan produksi kakao 5 tahun kedepan sebesar 1,35% (Tabel 12).

## KESIMPULAN

Dari ketiga metode estimasi yang digunakan dalam kajian ini yaitu ARIMA, Fungsi Transfer dan VAR, metode estimasi terbaik untuk estimasi produksi kakao berdasarkan pertimbangan statistik dan kerealistisan hasil estimasi dengan historis data aktualnya adalah Model FT ARIMA (2,2,1) xreg= harga produsen dengan MAPE *training* 5,79% dan *testing* 6,74%. Nilai ini dapat diartikan bahwa seluruh persentase kesalahan antara data aktual dengan data hasil ramalan pada data *training* adalah 5,79%, sedangkan rata-rata seluruh persentase kesalahan antara data aktual dengan data hasil ramalan pada data *testing* adalah 6,74%. Hasil estimasi produksi kakao tahun 2021-2025 akan terus meningkat dengan rata-rata pertumbuhan sebesar 1,35% per tahun.

## DAFTAR PUSTAKA

- Athif, Y.S. (2018). Pengaruh Kebijakan Bea Keluar Kakao Terhadap Harga Biji Kakao Domestik Indonesia. Skripsi. Bogor: Institut Pertanian Bogor.
- Enders, W. (2010). *Applied Econometric Time Series*. USA: University of Alabama. Wiley, Third Edition.
- Firdaus, M. (2019). Outlook Ekspor Kakao Indonesia. Bogor: Institut Pertanian Bogor.
- Gujarati, D.N. & Porter, D.C. (2010). Dasar-dasar Ekonometrika. Jakarta: Penerbit Salemba Empat, Buku 2, Edisi 5.

- Kementerian Pertanian. (2020). Statistik Perkebunan Indonesia 2019-2021. Jakarta: Direktorat Jenderal Perkebunan Kementerian Pertanian. Jakarta: Direktorat Jenderal Perkebunan, Kementerian Pertanian.
- Kementerian Pertanian. (2020). Buku Statistik Pertanian 2020. Jakarta: Pusat Data dan Sistem Informasi Pertanian, Kementerian Pertanian.
- Kementerian Pertanian & Badan Pusat Statistik. (2021). Petunjuk Teknis Metode Estimasi Data Komoditas Perkebunan. Jakarta: Pusat Data dan Sistem Informasi Pertanian, Kementerian Pertanian.
- Montgomery, D.C., Johnson, L.A. & Gardiner, J.S. (1990). Forecasting and Time Series Analysis. Singapore: Mc-Graw Hill.
- Rohmah, Y. (2020). Outlook Komoditas Perkebunan Kakao. Jakarta: Pusat Data dan Sistem Informasi Pertanian, Kementerian Pertanian.
- Sinuraya, J.F., Sinaga, B.M., Oktaviani, R., & Hutabarat, B. (2017). Dampak Kebijakan Pajak Ekspor dan Tarif Impor Terhadap Kesejahteraan Produsen dan Konsumen Kakao di Indonesia. *Jurnal Agro Ekonomi*, Vol. 35 No. 1 Mei 2017.
- Wei, William W.S. (2006). Time Series Analysis. Philadelphia: Department of Statistics The Fox School of Business and Management Temple University.
- Wooldridge, Jeffrey M. (2015). *Introductory Econometrics: A Modern Approach*. Boston: Cengage Learning.





**KUMPULAN ANALISIS MODEL ESTIMASI  
DATA KOMODITAS PERKEBUNAN**



**Pusat Data dan Sistem Informasi Pertanian  
Sekretariat Jenderal - Kementerian Pertanian  
TAHUN 2021**

**Jalan Harsono RM No. 3, Ragunan - Jakarta 12550  
Gedung D Lantai 4**