

## Vector Autoregressive Integrated (VARI) Method for Forecasting the Number of Internasional Visitor in Batam and Jakarta

### Metode *Vector Autoregressive Integrated* (VARI) untuk Peramalan Jumlah Wisatawan Mancanegara di Batam dan Jakarta

Septie Wulandary\*

#### Abstract

Forecasting methods that are often used are time series analysis, the Autoregressive (AR) method. The AR method only carries out univariate analysis, meaning that it carries out a separate model between the number of international visitor coming to Indonesia through Batam and Jakarta. Though there is a possibility, the number of international visitor arriving through Jakarta affects the number of international visitor arriving through Batam. Therefore, in this study the Vector Autoregressive Integrated (VARI) method is used. The VARI model is used on the number of international visitor arrivals per month at Batam and Jakarta for the period Januari 2014 – December 2019. VARI model formation through several stages, namely stationarity test, autoregressive order determination, VARI model formation, and diagnostic checking of the model. With the VARI model, VARI(5,1), the two significant simultaneously equation results are obtained. The Mean Absolute Percentage Error (MAPE) in this model are as follows 11,85% and 14,86% in predicting the number of international visitor arrivals in Batam and Jakarta. In this study also forecasting the number of international visitor arrivals in Batam and Jakarta in January – December 2020.

**Keywords:** forecasting, international visitor, Vector Autoregressive Integrated (VARI), Mean Absolute Percentage Error (MAPE)

#### Abstrak

Metode peramalan yang sering digunakan adalah analisis deret waktu, yaitu metode *Autoregressive* (AR). Metode AR hanya melakukan analisis secara univariat, artinya melakukan pemodelan terpisah antara jumlah wisatawan mancanegara (wisman) yang datang ke Indonesia melalui Batam dan Jakarta. Padahal ada kemungkinan, jumlah wisman yang datang di Jakarta mempengaruhi jumlah wisman yang datang di Batam. Oleh karena itu, pada penelitian ini digunakan metode *Vector Autoregressive Integrated* (VARI). Model VARI digunakan pada jumlah kedatangan wisatawan mancanegara per bulan di Batam dan Jakarta periode Januari 2014 – Desember 2019. Pembentukan model VARI melalui beberapa tahapan, yaitu uji stasioneritas, penentuan ordo *autoregressive*, pembentukan model VARI, dan pengecekan diagnostik model. Melalui model VARI, yaitu VARI(5,1), didapatkan kedua hasil persamaan signifikan secara simultan. Nilai *Mean Absolute Percentage Error* (MAPE) dalam model adalah 11,85% dan 14,86% dalam memprediksi jumlah kedatangan wisatawan mancanegara di Batam dan Jakarta. Pada penelitian ini juga

dilakukan peramalan untuk jumlah kedatangan wisatawan mancanegara di Batam dan Jakarta pada Januari – Desember 2020.

**Kata kunci:** peramalan, wisatawan mancanegara, *Vector Autoregressive Integrated* (VARI), *Mean Absolute Percentage Error* (MAPE)

## 1. PENDAHULUAN

Kontribusi sektor pariwisata terhadap PDB dari tahun 2010 hingga 2015 selalu mengalami peningkatan. Menurut BPS dan Kementerian Pariwisata, tahun 2010 kontribusi sektor pariwisata terhadap PDB sebesar 261,05 triliun menjadi 461,36 triliun rupiah. Kontribusi sektor pariwisata terhadap Devisa sebesar 7.603,45 juta dollar pada tahun 2010 menjadi 12.225,89 juta dollar di tahun 2015. Kontribusi terhadap tenaga kerja sebesar 4 juta orang tahun 2010 menjadi 12,1 juta orang di tahun 2015 atau 10,6% dari total tenaga kerja nasional [6].

Pada tahun 2017, sektor pariwisata secara konsisten menjadi program prioritas, yaitu pembangunan pariwisata Indonesia “*Wonderful Indonesia*” [6]. Pariwisata merupakan salah satu dari 5 (lima) sektor prioritas pembangunan 2017, yaitu pangan, energi, maritim, pariwisata, kawasan industri dan Kawasan Ekonomi Khusus (KEK). Tiga kawasan prioritas pembangunan pariwisata nasional dengan sebutan *Great Jakarta*, *Great Bali* dan *Great Batam* merupakan pusat destinasi unggulan. Hal ini dibuktikan dari total seluruh wisatawan mancanegara (wisman) yang masuk ke Indonesia tahun 2019, 27% datang melalui Batam dan Jakarta.

Salah satu metode peramalan yang sering digunakan adalah analisis deret waktu [9]. Analisis deret waktu adalah analisis yang melakukan peramalan nilai-nilai masa depan yang didasarkan pada nilai-nilai masa lampau suatu variabel dan atau kesalahan masa lampau. Salah satu analisis deret waktu yang sering digunakan adalah metode *Autoregressive* (AR). Metode AR hanya melakukan analisis secara univariat, artinya melakukan pemodelan terpisah antara jumlah wisatawan mancanegara (wisman) yang datang ke Indonesia melalui Batam dan Jakarta. Padahal ada kemungkinan, jumlah wisman yang datang di Batam mempengaruhi jumlah wisman yang datang di Jakarta, begitupun sebaliknya.

Model *Vector Autoregressive* (VAR) merupakan gabungan dari beberapa model *Autoregressive* (AR), dimana model membentuk sebuah vektor yang antara variabel-variabelnya saling mempengaruhi [8]. Pada pola data yang tidak stasioner, dilakukan proses *differencing*. Pemodelan data deret waktu multivariat berupa data *differencing* digunakan model *Vector Autoregressive Integrated* (VARI).

Beberapa penelitian sebelumnya yang menggunakan model VAR dengan *differencing* antara lain Hardani, Hoyyi, dan Sudarno [3] pada data inflasi, suku bunga, dan harga saham. Sedangkan Ranangga, Sumarjaya, dan Srinadi [7] yang menggunakan metode *Vector Autoregressive* (VAR) pada data yang ditransformasi ke bentuk logaritma natural. Iknas, Salam, dan Agustina [4] menggunakan metode VAR dalam meramalkan jumlah penduduk di Kabupaten Gowa.

Penelitian ini bertujuan untuk memodelkan dan melakukan peramalan jumlah wisman yang datang di Batam dan Jakarta menggunakan metode VARI. Kemudian membandingkan ketepatan peramalan dari model VARI menggunakan kriteria *Mean Absolut Percentage Error* (MAPE).

## 2. METODE PENELITIAN

### 2.1. Data Penelitian

Data penelitian ini menggunakan data sekunder yang dikeluarkan oleh Badan Pusat Statistik (BPS), yaitu data jumlah wisman yang masuk melalui pelabuhan di Batam dan Bandara

Soekarno-Hatta (Jakarta). Data yang digunakan merupakan data bulanan periode Januari 2014 hingga Desember 2019. Pengolahan data dilakukan menggunakan program R versi 3.5.1.

## 2.2. Korelasi Pearson

Untuk menerapkan model VAR terlebih dahulu akan dilihat korelasi variabel jumlah kedatangan wisman di Bali dan Jakarta dengan menggunakan rumus korelasi *pearson product moment* sebagai berikut:

$$r = \frac{n \cdot \sum XY - \sum X \cdot \sum Y}{\sqrt{n \cdot \sum X^2 - (\sum X)^2} \cdot \sqrt{n \cdot \sum Y^2 - (\sum Y)^2}} \quad (2.1)$$

## 2.3. Model Vector Autoregressive (VAR)

Menurut Enders [2], model VAR merupakan suatu sistem persamaan dinamis dimana pendugaan suatu peubah pada periode tertentu tergantung pada pergerakan peubah tersebut dan peubah-peubah lain yang terlibat dalam sistem pada periode-periode sebelumnya. Model VAR merupakan gabungan dari beberapa model *Autoregressive* (AR).

Bentuk umum dari model time series VAR(1) untuk  $n = 2$  adalah:

$$Z_{1,t} = \phi_{11}Z_{1,t-1} + \phi_{12}Z_{2,t-1} + \varepsilon_{1,t} \quad (2.2)$$

$$Z_{2,t} = \phi_{21}Z_{1,t-1} + \phi_{22}Z_{2,t-1} + \varepsilon_{2,t} \quad (2.3)$$

Model (2.1) dan (2.2) dapat dituliskan:

$$\begin{bmatrix} Z_{1,t} \\ Z_{2,t} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix} \quad (2.4)$$

Menurut [9], VAR dengan ordo  $p$  untuk  $n$  buah variabel pada waktu ke- $t$  dapat dimodelkan:

$$\mathbf{Z}_t = \phi_1 \mathbf{Z}_{t-1} + \phi_2 \mathbf{Z}_{t-2} + \dots + \phi_p \mathbf{Z}_{t-p} + \boldsymbol{\varepsilon}_t \quad (2.5)$$

dengan:

$\mathbf{Z}_t$  = vektor variabel tak bebas berukuran  $m \times 1$

$\phi_i$  = matriks parameter berukuran  $m \times m$  untuk setiap  $i = 1, 2, \dots, p$

$\boldsymbol{\varepsilon}_t$  = vektor sisaan berukuran  $m \times 1$

## 2.4. Model Vector Autoregressive Integrated (VARI)

Apabila data mengalami proses *first differencing* untuk menghasilkan data yang stasioner, maka bentuk model VARI(1,1) menjadi sebagai berikut:

$$Z_{1,t} - Z_{1,t-1} = \phi_{11}(Z_{1,t-1} - Z_{1,t-2}) + \phi_{12}(Z_{2,t-1} - Z_{2,t-2}) + \dots + \phi_{1n}(Z_{n,t-1} - Z_{n,t-2}) + \varepsilon_{1,t}$$

$$Z_{2,t} - Z_{2,t-1} = \phi_{21}(Z_{1,t-1} - Z_{1,t-2}) + \phi_{22}(Z_{2,t-1} - Z_{2,t-2}) + \dots + \phi_{2n}(Z_{n,t-1} - Z_{n,t-2}) + \varepsilon_{2,t}$$

$\vdots$

$$Z_{n,t} - Z_{n,t-1} = \phi_{n1}(Z_{1,t-1} - Z_{1,t-2}) + \phi_{n2}(Z_{2,t-1} - Z_{2,t-2}) + \dots + \phi_{nn}(Z_{n,t-1} - Z_{n,t-2}) + \varepsilon_{n,t}$$

Sehingga dalam bentuk matriks menjadi :

$$\begin{bmatrix} Z_{1,t} - Z_{1,t-1} \\ Z_{2,t} - Z_{2,t-1} \\ \vdots \\ Z_{n,t} - Z_{n,t-1} \end{bmatrix} = \begin{bmatrix} \phi_{11} & \phi_{12} & \dots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \dots & \phi_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{p1} & \phi_{p2} & \dots & \phi_{pn} \end{bmatrix} \begin{bmatrix} Z_{1,t-1} - Z_{1,t-2} \\ Z_{2,t-1} - Z_{2,t-2} \\ \vdots \\ Z_{n,t-1} - Z_{n,t-2} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{n,t} \end{bmatrix}$$

atau dapat ditulis :

$$\mathbf{Z}_t - \mathbf{Z}_{t-1} = \Phi(\mathbf{Z}_{t-1} - \mathbf{Z}_{t-2}) + \boldsymbol{\varepsilon}_t \quad (2.6)$$

dengan asumsi bahwa *error* berdistribusi normal,  $\boldsymbol{\varepsilon}_t \sim N(0, \sigma^2)$  dan  $\mathbf{Y}_t = \mathbf{Z}_t - \mathbf{Z}_{t-1}$

## 2.5. Tahapan Analisis Model VARI

### Stasioneritas Data

Data stasioner adalah data deret waktu yang tidak memiliki tren, pola musiman, serta rata-rata dan ragamnya konstan atau homogen dari waktu ke waktu. Kestasioneran data deret waktu dapat diperiksa dengan melihat plot deret waktu. Plot deret waktu yang berfluktuasi dengan ragam yang konstan disekitar rata-rata yang konstan menunjukkan bahwa data deret waktu tersebut stasioner.

Kestasioneran data dapat diuji dengan uji *Augmented Dickey Fuller*. Hipotesis yang diuji adalah:

$H_0: \gamma = 0$  (data bersifat tidak stasioner)

$H_0: \gamma < 0$  (data bersifat stasioner)

Nilai  $\gamma$  diduga melalui metode kuadrat terkecil dan pengujian dilakukan dengan uji t. Statistik uji-nya yaitu:

$$t_{hit} = \frac{\hat{\gamma}}{se(\hat{\gamma})}$$

dengan :

$\hat{\gamma}$  = nilai dugaan  $\gamma$

$se(\hat{\gamma})$  = simpangan baku dari  $\gamma$

Jika  $t_{hit} <$  nilai kritis *MacKinnon* ( $\alpha$ ), maka keputusan yang diambil adalah menolak  $H_0$  yang berarti data bersifat stasioner.

### Pemilihan Ordo $p$

Berdasarkan [9], sebelum melakukan analisis VARI perlu dilakukan pemeriksaan lag yang optimal terlebih dahulu. Penentuan panjang lag optimal untuk model VARI menggunakan nilai *Akaike Information Criterion* (AIC) terkecil dengan rumus:

$$AIC(M) = n \ln(\hat{\sigma}_\alpha^2) + 2M \quad (2.7)$$

dengan M adalah banyaknya parameter dalam model. Orde optimal dalam model dipilih berdasarkan nilai dari M, yang merupakan fungsi dari  $p$  dan  $q$ , sehingga nilai AIC (M) menjadi minimum.

### Differencing

Menurut [5], proses *differencing* pada data pengamatan dengan mengurangi nilai-nilai yang berurutan dengan satu periode sebelumnya, yang akan didapat pembeda pertama (*first difference*).

## 2.6. Pengujian Diagnostik Model

### Uji Normalitas Residual

Uji kenormalan residual pada model VAR, umumnya menggunakan uji normalitas *Jarque-Berra*.

Hipotesis untuk uji *Jarque-Berra* adalah:

$H_0$  : Residual berdistribusi normal

$H_1$  : Residual tidak berdistribusi normal

Penghitungan nilai Multivariat *Jarque-Berra* (MJB) menggunakan rumus:

$$JB = n \left( \frac{(\sqrt{b_1})^2}{6} + \frac{(b_2-3)^2}{24} \right) \quad (2.8)$$

dengan:  $b_1$  : kemencengan dan  $b_2$  : kurtosis

Tolak  $H_0$  jika  $P(\chi_{(2)}^2 > JB)$  atau  $p$ -value  $< \alpha$  yang berarti bahwa data residual tidak menyebar normal.

### Uji Asumsi Autokorelasi

Autokorelasi adalah suatu kondisi dimana ada tidaknya korelasi antara variabel gangguan (*error*) pada periode tertentu dengan variabel gangguan pada periode sebelumnya. Uji asumsi autokorelasi menggunakan uji *Portmanteau*.

$H_0 = \hat{r}_1 = \hat{r}_2 = \dots = \hat{r}_h = 0$  (tidak ada autokorelasi residual sampai lag ke-h)

$H_1: \exists r_i \neq 0$  (ada autokorelasi dalam residual sampai lag ke-h).

Statistik uji yang digunakan adalah :

$$Q = T(T + 2) \sum_{h=1}^h \frac{\hat{r}_h^2}{T-h} \quad (2.9)$$

dengan:

$T$  : banyaknya residual

$\hat{r}_h$  : autokorelasi antar residual

$h$  : lag

Jika nilai  $p - value > \alpha$  maka terima  $H_0$  atau tidak ada komponen autokorelasi yang signifikan hingga lag ke-h.

### Uji Homoskedastisitas Residual

Uji Homoskedastisitas residual menggunakan uji ARCH-LM. Hipotesis untuk uji ARCH adalah sebagai berikut:

$H_0$  : Kuadrat residual tidak menunjukkan heteroskedastisitas

$H_1$  : Kuadrat residual menunjukkan heteroskedastisitas

Statistik Uji

$$LM = T \times R^2 \quad (2.10)$$

dengan:

$T$  = banyaknya pengamatan

$R^2$  = koefisien determinasi model regresi  $\varepsilon_t^2$

$q$  = banyaknya pengamatan yang mempengaruhi

Jika  $LM$  lebih besar dari  $\chi_q^2$ , mengindikasikan adanya pengaruh ARCH.

## 2.7. Kriteria Peramalan

Kriteria peramalan yang digunakan untuk melihat seberapa akurat model memprediksi kejadian yang sebenarnya adalah *Mean Absolut Percentage Error* (MAPE). Semakin kecil nilai MAPE, maka hasil peramalan semakin mendekati hasil yang aktual. MAPE dirumuskan sebagai berikut:

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{Z_{i,t} - \hat{Z}_{i,t}}{Z_{i,t}} \right| \times 100\% \quad (2.11)$$

dengan :

$Z_{i,t}$  : nilai aktual periode ke-t, lokasi ke-i

$\hat{Z}_{i,t}$  : Nilai ramalan periode ke-t, lokasi ke-i

$n$  : jumlah observasi data deret waktu

Menurut Zhang dkk (2015) dalam Ranangga [7] kriteria MAPE adalah:

**Tabel 2.1** Kriteria Nilai MAPE

MAPE	Keterangan
< 10%	Peramalan sangat baik
10% - 20%	Peramalan baik
20% - 50%	Peramalan cukup baik
> 50%	Peramalan kurang baik

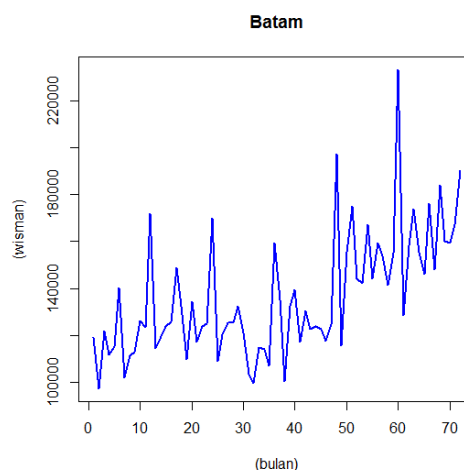
## 2.8. Langkah-langkah Penelitian

Langkah-langkah penelitian yang dilakukan adalah sebagai berikut:

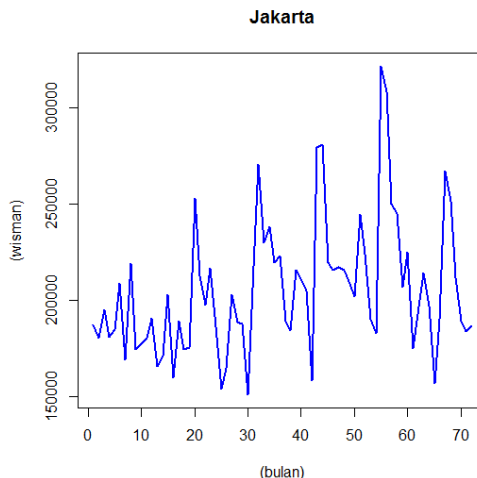
1. Plot data *time series* jumlah kedatangan wisman di Batam dan Jakarta dari bulan Januari 2014 – Desember 2019.
2. Penghitungan menggunakan model VARI, maka:
  1. Melakukan uji korelasi untuk melihat apakah terdapat hubungan diantara kedua variabel.
  2. Melakukan uji kestasioneran data. Apabila data belum stasioner, maka data distasionerkan terlebih dahulu dengan proses *differencing*.
  3. Menentukan *lag* VARI.
  4. Melakukan estimasi parameter.
  5. Menentukan model VARI.
3. Melakukan diagnostik model:
  - a. Uji Asumsi *Multivariate* Normal
  - b. Uji Asumsi Homoskedastisitas
  - c. Uji Asumsi *Non Autocorrelation*
5. Membandingkan ketepatan peramalan dari model VARI menggunakan kriteria *Mean Absolute Percentage Error (MAPE)*.
6. Melakukan peramalan menggunakan model VARI.

## 3. HASIL DAN PEMBAHASAN

Plot jumlah wisman di Batam dan Jakarta disajikan melalui gambar berikut:



**Gambar 3.1.** Plot Jumlah Wisman di Batam



**Gambar 3.2.** Plot Jumlah Wisman di Jakarta

Nilai korelasi Pearson yang diperoleh sebesar 0,18 menunjukkan bahwa jumlah kedatangan wisman di Bali dan Jakarta mempunyai hubungan yang positif pada orde waktu yang sama. Nilai ini memungkinkan untuk dilakukan analisis selanjutnya untuk melihat hubungan kedua sektor tersebut secara simultan dengan model *time series*.

### 3.1. Pengujian kestasioneran data

Pengujian stasioneritas dengan uji ADF menunjukkan hasil bahwa jumlah wisman yang datang melalui Batam belum stasioner pada nilai  $\alpha = 5\%$ , dengan *p-value* sebesar 0,344. Jumlah wisman yang datang melalui Jakarta juga belum stasioner dengan *p-value* sebesar 0,066.

Setelah dilakukan *differencing* satu, dilakukan uji ADF untuk melihat apakah data sudah stasioner. Uji ADF menunjukkan bahwa jumlah wisman yang datang melalui Batam sudah stasioner dengan *p-value* sebesar 0,01. Begitu pula jumlah wisman di Jakarta juga sudah stasioner dengan *p-value* sebesar 0,01.

### 3.2. Penentuan Panjang Lag Model

Penentuan panjang *lag* model dalam penelitian ini menggunakan nilai AIC dan SC yang terkecil. Pada Tabel 3.1 terlihat hasil identifikasi untuk lima *lag* yang telah dilakukan. Nilai AIC dan SC terkecil terdapat pada AR(5), yaitu 40,65 dan 41,34. Sehingga model VARI yang digunakan adalah VARI ordo 5 atau VARI(5,1).

**Tabel 3.1.** Nilai Kriteria AIC dan SC Lag 1 – 5

Kriteria	AR (1)	AR (2)	AR (3)	AR (4)	AR (5)
AIC	41,61	41,18	41,01	41,09	40,65

### 3.3. Estimasi Model VARI(5,1)

Pada model VARI dengan ordo 5, VARI(5,1), suatu variabel dipengaruhi oleh masa lalunya dan masa lalu variabel lain pada lima satuan waktu. Estimasi parameter-parameter pada model VARI(5,1) dilakukan dengan metode *Ordinary Least Square* (OLS). Proses pertama yaitu melihat tingkat signifikansi parameternya. Signifikansi semua parameter secara serentak diuji dengan

statistik-F, sedangkan signifikansi parameter secara parsial diuji dengan statistik t. Hasil estimasi parameter model VARI(5,1) tersaji pada Tabel 3.2.

**Tabel 3.2.** Hasil Estimasi Model VARI(5,1)

Persamaan	Variabel	Koefisien	<i>p-value</i> parsial	<i>p-value</i> simultan
$\hat{Y}_{1,t}$	c	3027,7	0,1541	$1,153 \times 10^{-13}$
	$Y_{1,t-1}$	-1,0388	2,91e-14 *	
	$Y_{2,t-1}$	-0,0566	0,3885	
	$Y_{1,t-2}$	-0,8683	3,00e-07 *	
	$Y_{2,t-2}$	-0,0215	0,7668	
	$Y_{1,t-3}$	-0,6271	0,0006 *	
	$Y_{2,t-3}$	-0,1405	0,0716	
	$Y_{1,t-4}$	-0,4763	0,0027 *	
	$Y_{2,t-4}$	0,0787	0,2740	
	$Y_{1,t-5}$	-0,4781	1,92e-05 *	
$Y_{2,t-5}$	0,2676	0,0002 *		
$\hat{Y}_{2,t}$	c	324,9	0,9396	0,002208
	$Y_{1,t-1}$	-0,5081	0,0178 *	
	$Y_{2,t-1}$	-0,4395	0,0016 *	
	$Y_{1,t-2}$	-0,5919	0,0562	
	$Y_{2,t-2}$	-0,4666	0,0024 *	
	$Y_{1,t-3}$	-0,1450	0,6811	
	$Y_{2,t-3}$	-0,2315	0,1431	
	$Y_{1,t-4}$	0,3127	0,3160	
	$Y_{2,t-4}$	-0,1931	0,1876	
	$Y_{1,t-5}$	0,3676	0,0831	
$Y_{2,t-5}$	0,0022	0,9874		

Persamaan akhir VARI(5,1) untuk Batam ( $\hat{Y}_{1,t}$ ) dan Jakarta ( $\hat{Y}_{2,t}$ ) adalah dengan mengambil nilai parameter yang signifikan, dapat ditulis sebagai berikut:

$$\hat{Y}_{1,t} = -1,039 Y_{1,t-1} - 0,868 Y_{1,t-2} - 0,627 Y_{1,t-3} - 0,476 Y_{1,t-4} - 0,478 Y_{1,t-5} + 0,267 Y_{2,t-5}$$

$$\hat{Y}_{2,t} = -0,508 Y_{1,t-1} - 0,440 Y_{2,t-1} - 0,467 Y_{2,t-2}$$

dengan diketahui  $Y_{i,t} = Z_{i,t} - Z_{i,t-1}$ , sehingga persamaan di atas dapat diuraikan menjadi sebagai berikut :

$$\hat{Z}_{1,t} = -0,039 Z_{1,t-1} + 0,171 Z_{1,t-2} + 0,241 Z_{1,t-3} + 0,151 Z_{1,t-4} + 0,226 Z_{1,t-5} + 0,21 Z_{2,t-6}$$

$$\hat{Z}_{2,t} = -0,508 Z_{1,t-1} + 0,508 Z_{1,t-2} + 0,56 Z_{2,t-1} - 0,027 Z_{2,t-2} + 0,467 Z_{2,t-3}$$

Berdasarkan hasil olah data, diperoleh nilai *adjusted R*<sup>2</sup> persamaan parsial model VARI(5,1) pada jumlah kedatangan wisman di Batam cukup besar, yaitu 0,72. Sedangkan *adjusted R*<sup>2</sup> jumlah kedatangan wisman di Jakarta sebesar 0,26. Hal ini menunjukkan bahwa keragaman model pada persamaan parsial di Batam dapat dijelaskan oleh variabel jumlah wisman di Batam dan Jakarta secara simultan.

### 3.4. Diagnostik Model

#### Uji Normalitas Multivariat Residual



## Septie Wulandary

Berdasarkan hasil uji *Jarque-Bera* (JB) diperoleh nilai *chi-square* 32,57 ( $p - value = 1,46 \times 10^{-6}$ ) yang lebih kecil dari  $\alpha = 0,05$ . Hal ini berarti tolak  $H_0$ , sehingga dapat disimpulkan bahwa data residual tidak berdistribusi normal multivariat.

### Uji Homoskedastisitas Residual

Uji Homokedastisitas Residual bertujuan untuk mengetahui apakah varians residual konstan atau tidak. Berdasarkan hasil pengujian menggunakan uji *Autoregressive Conditional Heteroskedasticity* (ARCH) diperoleh  $p-value$  0.763 yang lebih besar dari  $\alpha = 0,05$ , yang berarti menerima  $H_0$ . Sehingga dapat disimpulkan bahwa data residual memenuhi asumsi homoskedastisitas.

### Uji Autokorelasi Residual

Uji *Portmanteau* bertujuan untuk mengetahui apakah residual memenuhi syarat tidak ada autokorelasi dalam residual. Hasil pengujian menunjukkan dengan taraf signifikansi 10 %, diperoleh  $p-value=0,0926$ , yang lebih besar dari  $\alpha = 0,05$  (terima  $H_0$ ). Sehingga disimpulkan bahwa data residual memenuhi asumsi *non* autokorelasi.

### 3.5. Nilai Ketepatan Peramalan Model VARI(5,1)

Hasil olah data menunjukkan bahwa nilai MAPE untuk variabel jumlah wisman di Batam sebesar 11,85%. Sedangkan nilai MAPE untuk variabel jumlah wisman di Jakarta sebesar 14,86%. Nilai MAPE tersebut menunjukkan bahwa model VARI(5,1) di Batam dan Jakarta baik untuk peramalan karena memiliki MAPE di antara 10% - 20%.

### 3.6. Peramalan Jumlah Wisman dengan Model VARI(5,1)

Pada penelitian ini, dilakukan peramalan jumlah kedatangan wisman di Batam dan Jakarta sebanyak 12 bulan (Januari – Desember 2020).

**Tabel 3.3.** Peramalan Jumlah Wisman di Batam dan Jakarta

Bulan	Jumlah Wisman (orang)	
	Batam	Jakarta
Januari	181 482	176 319
Februari	178 883	183 957
Maret	175 636	186 776
April	173 747	183 618
Mei	175 354	184 651
Juni	172 438	184 855
Juli	169 109	185 764
Agustus	169 343	182 145
September	170 092	183 223
Oktober	170 855	184 323
November	171 378	182 752
Desember	173 089	182 943

## 4. KESIMPULAN

Berdasarkan hasil penelitian menggunakan model *Vector Autoregressive Integrated* (VARI), dapat disimpulkan sebagai berikut:

1. Jumlah kedatangan wisman di Batam dan Jakarta telah stasioner sehingga model *Vector Autoregressive Integrated* (VARI) dapat diterapkan. Model VARI yang cocok digunakan sesuai dengan pengujian panjang lag adalah model VARI dengan ordo lima.
2. Berdasarkan hasil diagnostik model dapat disimpulkan bahwa model VARI(5,1) telah memenuhi asumsi homoskedastisitas dan *non* autokorelasi residual. Akan tetapi, model VARI(5,1) tidak memenuhi asumsi normalitas residual.
3. Nilai MAPE untuk variabel jumlah wisman di Batam sebesar 11,85% dan di Jakarta 14,86%. Hal ini menunjukkan bahwa model VARI(5,1) memberikan hasil peramalan yang baik.

## DAFTAR PUSTAKA

- [1] Badan Pusat Statistik (BPS). 2020. Tabel Dinamik Jumlah Kunjungan Wisman. <https://www.bps.go.id/> (diakses tanggal 20 Juni 2020).
- [2] Enders, Walter. 2004. *Applied Time Series Econometrics*. New York: John Wiley & Sons Inc.
- [3] Hardani, P. R., Hoyyi, A., Sudarno. 2016. Peramalan Laju Inflasi, Suku Bunga Indonesia dan Indeks Harga Saham Gabungan Menggunakan Metode *Vector Autoregressive* (VAR). *Jurnal Gaussian*, Vol. 6, No. 1, Tahun 2016, pp. 101-110.
- [4] Ibtnas, R., Salam, N., dan Agustina, N. 2019. Implementasi Metode *Vector Autoregressive* (VAR) dalam Meramalkan Jumlah Penduduk (Studi Kasus: Kabupaten Gowa). *Journal Matematika dan Statistika serta Aplikasinya*, Vol. 7 No. 2 (2019), pp. 56-61.
- [5] Makridakis, McGee, dan Wheelwright. 1999. *Metode dan Aplikasi Peramalan Edisi Kedua*. Diterjemahkan oleh Suminto, H. Jakarta: Binarupa Aksara.
- [6] Media Indonesia. 2018. <https://mediaindonesia.com/read/detail/192746-pariwisata-indonesia-peringkat-9-di-dunia> (diakses tanggal 10 Maret 2020).
- [7] Ranangga, T., Sumarjaya, I., dan Srinadi, I. 2018. Metode *Vector Autoregressive* (VAR) dalam Peramalan Jumlah Wisatawan Mancanegara ke Batam. *E-Jurnal Matematika* Vol. 7 (2), Mei 2018, pp. 157-164.
- [8] Sihombing, P. Robinson dan Susilowati, B. Endar. 2019. Aplikasi Model *Vector Autoregressive* (VAR) pada Data Tamu Mancanegara di Hotel Bintang dan Non Bintang di Daerah Istimewa Yogyakarta. *Jurnal Statistika dan Aplikasinya (JSA)*, Vol. 3, No. 2, Desember 2019.
- [9] Wei, William W.S. 2006. *Time Series Analysis Univariate and Multivariate Methods; Second Edition*. USA: Pearson Addison Wesley.

**LAMPIRAN SYNTAX R DAN OUTPUT**

```
library(vars)
library(astsa)
library (forecast)
library(TSA)
library(tseries)

#Membangkitkan Data
data<-read.table(file.choose(), header=T, sep=";")
data
> data
      Batam Jakarta
1  119054  187123
2   97445  180362
3  122019  194720
4  111929  180787
5  115323  184534
6  140218  208624
7  101996  169135
8  111455  218903
9  113090  174169
10 126169  177274
11 123505  180208
12 171907  190598
13 114478  165746
14 119642  170741
15 124019  203019
16 125816  159873
17 148920  189307
18 132569  174319
19 109931  175347
20 134540  252914
21 117089  212706
22 123750  197487
23 124983  216517
24 170081  186299
25 108888  153503
26 120351  164317
27 125324  202669
28 125573  188369
29 132410  187545
30 121253  150956
31 103647  217452
32  99724  270496
33 115052  229964
34 114020  237914
35 106953  219246
36 159277  222844
37 135046  189242
38 100329  183977
39 132551  215651
40 139361  211131
41 117372  204723
```

## Septie Wulandary

```

42 130541 158147
43 122869 279550
44 124044 280737
45 122835 219348
46 117589 215612
47 124833 217201
48 197347 215450
49 115590 209553
50 155671 201981
51 175194 244467
52 144104 223814
53 142157 190216
54 167398 182887
55 144235 321468
56 159218 308231
57 153944 250058
58 141337 244775
59 155283 206820
60 233153 224989
61 128675 174963
62 158088 196183
63 173947 214161
64 155619 196977
65 146075 156654
66 176187 190031
67 148205 267143
68 184077 251596
69 160293 211775
70 159292 189231
71 167288 183759
72 190232 186723
attach(data)
wisman<-as.matrix(data,ncol=2,nrow=72)

#Uji Korelasi
cor(Batam,Jakarta)
> cor(Batam,Jakarta)
[1] 0.1754234

#Uji Stasioneritas Data

##Plot Time Series
plot.ts(wisman[,1],          ylab="(wisman)",xlab="(bulan)",main="Batam",
col="blue",lwd=2)
plot.ts(wisman[,2],        ylab="(wisman)",      xlab="(bulan)",main="Jakarta",
col="blue",lwd=2)

#Differencing
dBatam<-diff(Batam)
dJakarta<-diff(Jakarta)
p1 <- cbind(dBatam,dJakarta)

```

## Septie Wulandary

```
##Unit Root Test
adf.test(dBatam,alternative = c("stationary"))
adf.test(dJakarta,alternative = c("stationary"))

> adf.test(dBatam,alternative = c("stationary"))

      Augmented Dickey-Fuller Test

data:  dBatam
Dickey-Fuller = -7.2783, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(dBatam, alternative = c("stationary")) :
  p-value smaller than printed p-value
> adf.test(dJakarta,alternative = c("stationary"))

      Augmented Dickey-Fuller Test

data:  dJakarta
Dickey-Fuller = -4.4837, Lag order = 4, p-value = 0.01
alternative hypothesis: stationary

Warning message:
In adf.test(dJakarta, alternative = c("stationary")) :
  p-value smaller than printed p-value

##Uji Lag Optimum
VARselect(p1, lag.max=10, type= "none") # Pilih sendiri

$`selection`
AIC(n)  HQ(n)  SC(n) FPE(n)
   10    10     6    10

$criteria
      1          2          3          4          5
AIC(n) 4.161025e+01 4.118221e+01 4.100950e+01 4.109169e+01 4.065204e+01
HQ(n)  4.166450e+01 4.129070e+01 4.117225e+01 4.130867e+01 4.092327e+01
SC(n)  4.174867e+01 4.145905e+01 4.142476e+01 4.164536e+01 4.134413e+01
FPE(n) 1.177941e+18 7.680153e+17 6.467789e+17 7.034157e+17 4.545151e+17
      6          7          8          9         10
AIC(n) 4.040072e+01 4.047456e+01 4.044474e+01 4.028198e+01 4.011217e+01
HQ(n)  4.072621e+01 4.085430e+01 4.087872e+01 4.077020e+01 4.065464e+01
SC(n)  4.123123e+01 4.144349e+01 4.155209e+01 4.152774e+01 4.149635e+01
FPE(n) 3.550790e+17 3.846894e+17 3.765581e+17 3.235543e+17 2.769032e+17

##Model VAR(1)
var5<-VAR(p1, p=5, type="const")
summary(var5)
```

## Septie Wulandary

```

> summary(var5)

VAR Estimation Results:
=====
Endogenous variables: dBatam, dJakarta
Deterministic variables: const
Sample size: 66
Log Likelihood: -1503.994
Roots of the characteristic polynomial:
0.918 0.918 0.9117 0.8943 0.8943 0.8835 0.8835 0.7447 0.7447 0.3736
Call:
VAR(y = p1, p = 5, type = "const")

Estimation results for equation dBatam:
=====
dBatam = dBatam.l1 + dJakarta.l1 + dBatam.l2 + dJakarta.l2 + dBatam.l3 + dJakarta.l3 + dBatam.l4 + dJakarta.l4 + dBatam.l5 + dJakarta.l5 + const

      Estimate Std. Error t value Pr(>|t|)
dBatam.l1 -1.03881    0.10202 -10.182 2.91e-14 ***
dJakarta.l1 -0.05661    0.06513  -0.869 0.388526
dBatam.l2 -0.86830    0.14888  -5.832 3.00e-07 ***
dJakarta.l2 -0.02147    0.07203  -0.298 0.766786
dBatam.l3 -0.62709    0.17225  -3.641 0.000602 ***
dJakarta.l3 -0.14048    0.07647  -1.837 0.071612 .
dBatam.l4 -0.47626    0.15169  -3.140 0.002719 **
dJakarta.l4  0.07873    0.07125   1.105 0.274018
dBatam.l5 -0.47815    0.10218  -4.679 1.92e-05 ***
dJakarta.l5  0.26762    0.06731   3.976 0.000206 ***
const      3027.73858 2095.17458   1.445 0.154103
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 16490 on 55 degrees of freedom
Multiple R-Squared: 0.7592,    Adjusted R-squared: 0.7155
F-statistic: 17.34 on 10 and 55 DF,  p-value: 1.153e-13

Estimation results for equation dJakarta:
=====
dJakarta = dBatam.l1 + dJakarta.l1 + dBatam.l2 + dJakarta.l2 + dBatam.l3 + dJakarta.l3 + dBatam.l4 + dJakarta.l4 + dBatam.l5 + dJakarta.l5 + const

      Estimate Std. Error t value Pr(>|t|)
dBatam.l1 -5.081e-01  2.079e-01  -2.444 0.01776 *
dJakarta.l1 -4.395e-01  1.327e-01  -3.312 0.00164 **
dBatam.l2 -5.919e-01  3.034e-01  -1.951 0.05616 .
dJakarta.l2 -4.666e-01  1.468e-01  -3.179 0.00242 **
dBatam.l3 -1.450e-01  3.510e-01  -0.413 0.68115
dJakarta.l3 -2.315e-01  1.558e-01  -1.486 0.14311
dBatam.l4  3.127e-01  3.091e-01   1.012 0.31604
dJakarta.l4 -1.937e-01  1.452e-01  -1.334 0.18763
dBatam.l5  3.676e-01  2.082e-01   1.765 0.08306 .
dJakarta.l5  2.171e-03  1.372e-01   0.016 0.98743
const      3.249e+02  4.269e+03   0.076 0.93961
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 33600 on 55 degrees of freedom
Multiple R-Squared: 0.3732,    Adjusted R-squared: 0.2592
F-statistic: 3.274 on 10 and 55 DF,  p-value: 0.002208

Covariance matrix of residuals:
      dBatam dJakarta
dBatam 271929567 -19467793
dJakarta -19467793 1129064292

Correlation matrix of residuals:
      dBatam dJakarta
dBatam 1.00000 -0.03513
dJakarta -0.03513 1.00000

```

```

##Uji Asumsi Residual Berdistribusi Normal Multivariate
norm1 <- normality.test(var5)
norm1

```

## Septie Wulandary

```
$`JB`

      JB-Test (multivariate)

data:  Residuals of VAR object var5
Chi-squared = 32.573, df = 4, p-value = 1.461e-06

$Skewness

      Skewness only (multivariate)

data:  Residuals of VAR object var5
Chi-squared = 18.688, df = 2, p-value = 8.748e-05

$Kurtosis

      Kurtosis only (multivariate)

data:  Residuals of VAR object var5
Chi-squared = 13.885, df = 2, p-value = 0.000966

##Uji Heteroskedastisitas Residual
arch1 <- arch.test(var5,lags.multi = 10)
arch1
> arch1

      ARCH (multivariate)

data:  Residuals of VAR object var5
Chi-squared = 80.091, df = 90, p-value = 0.7634

##Uji Autokorelasi
ser1<-serial.test(var5, lags.pt = 10, type = "PT.asymptotic")
ser1
> ser1

      Portmanteau Test (asymptotic)

data:  Residuals of VAR object var5
Chi-squared = 28.761, df = 20, p-value = 0.09257
```