

Forecasting Jumlah Kasus Harian Covid-19 di Provinsi Jawa Barat Menggunakan Model *Vector Autoregressive* (VAR)

Reshiva Yuriska¹, Aniq Atiqi Rohmawati², Annisa Aditsania³

^{1,2,3} Universitas Telkom, Bandung

¹reshivayrsk@students.telkomuniversity.ac.id,

²aniqatiqi@telkomuniversity.ac.id,³aaditsania@telkomuniversity.ac.id

Abstrak

Mengembangkan perencanaan strategis untuk pengendalian penyebaran Covid-19 masih menjadi tantangan saat ini. Meskipun demikian, membangun model memungkinkan peramalan besarnya kasus pada masa yang akan datang, karena perencanaan sistem kesehatan masyarakat dapat mencegah kematian. Dalam konteks ini, deret waktu multivariat dapat digunakan untuk memprediksi kasus aktif dan sembuh Covid-19 dalam mengembangkan model deret waktu yang efektif. Menurut WHO, Jawa Barat melaporkan 340.455 kasus kumulatif yang menempati posisi kedua di Indonesia dalam jumlah total kasus Covid-19. Penelitian ini menunjukkan proses *Vector Autoregressive* (VAR) untuk pemodelan data deret waktu kasus aktif dan sembuh di Jawa Barat. Dalam penelitian ini menunjukkan model VAR mampu menangkap tren data dan memprediksi kasus aktif Covid-19 berdasarkan kasus sembuh harian dengan MAPE 4,7%. Model ini mungkin dapat membantu pemerintah dan pemangku kepentingan kesehatan masyarakat dalam pengambilan keputusan.

Kata kunci : Autoregressive, Covid-19, Pengambilan Keputusan, Time Series

Abstract

Developing strategic planning to control the spread of Covid-19 remains a challenge nowadays. Nonetheless, constructing models allows forecasting future cases' magnitude due to the public health system planning to prevent deaths. In this context, multivariate time series can be used to predict active and cured cases of Covid-19 in developing an effective time series model. Also, according to WHO, West Java reported 340.455 cumulative cases, holding the second position in Indonesia in the number of Covid-19 total cases. This work demonstrates the parsimonious of *Vector Autoregressive* (VAR) process for modeling time series data of active and recovery cases in West Java. We show how this differenced- VAR model capable to capture trends and measure the future of Covid-19 active cases based on daily recovery cases with MAPE 4.7%. This model may be able to assist government and public health stakeholders in decision making.

Keywords: Autoregressive, Covid-19, Decision-making, Differencing Multivariate, Time Series

1. Pendahuluan

Latar Belakang

Virus Corona atau Covid-19 adalah virus baru yang menyerang hampir di setiap belahan bumi. Covid-19 pertama kali terkonfirmasi di Wuhan, China pada akhir 2019. Sampai saat ini belum ditemukan obat untuk menanggulangi penularan virus ini. Lebih dari 200 negara di seluruh dunia mengonfirmasi adanya kasus terinfeksi virus corona, termasuk Indonesia [1]. Menurut Djalante dkk (2020), berdasarkan data dari Gugus Tugas Penanganan Covid-19, di Indonesia kasus pertama kali dilaporkan pada 2 Maret 2020. Dimana Covid-19 terjadi pada dua warga Depok, Jawa Barat [2]. Untuk menekan penyebaran virus Covid-19 Indonesia menggunakan cara yang berbeda, dalam peraturan perundang-undangan di Indonesia istilah ini dikenal dengan karantina wilayah. Berbagai kebijakan telah dikeluarkan oleh pemerintah pusat dalam rangka menanggulangi Covid-19, tak terkecuali provinsi Jawa Barat yang merupakan provinsi ditemukannya kasus pertama terinfeksi Covid-19 serta provinsi terbesar ketiga setelah Jawa Timur [3].

Melihat permasalahan tersebut mendorong penulis untuk menerapkan model *time series* yang sangat berguna untuk pemodelan data yang dikumpulkan dan di indeks oleh waktu. Analisis *time series* telah efektif digunakan untuk memodelkan, memperkirakan, dan meramalkan masalah nyata yang sebenarnya [4]. Salah satu metode *time series* tersebut adalah metode *Vector Autoregressive* (VAR) yang digunakan untuk mengetahui prediksi kasus terkonfirmasi positif Covid-19 di provinsi Jawa Barat. Model VAR ini pernah digunakan oleh Hariyati H. Usman dalam penelitian yang berjudul "Pendekatan Model *Vector Autoregressive* (VAR) Untuk Meramalkan Faktor-Faktor yang Mempengaruhi Inflasi di Provinsi Gorontalo" [5]. Penelitian ini memberikan respon yang positif, yaitu berdasarkan model VAR dan pengujian signifikansi parameter diketahui bahwa sektor inflasi dipengaruhi oleh sektor makanan, minuman, dan tembakau. Model VAR, diusulkan oleh Christopher Sims (1980) yang bermanfaat untuk meramalkan dan menjelaskan perilaku dinamis antara pengamatan deret waktu

[6]. Menurut Gujarati, metode VAR memiliki beberapa keunggulan antara lain: tidak perlu membedakan antara peubah bebas dan peubah terikat, menggunakan metode *Ordinary Least Square* (OLS) dalam mengestimasi tiap persamaan, dan peramalan menggunakan metode VAR dalam beberapa kasus lebih baik [7]. Selain itu, VAR adalah proses stabilitas yang menyediakan *invariant* waktu pada *mean* momen pertama dan *varians* momen kedua. Penelitian ini disusun dengan cara sebagai berikut: hal pertama, penulis menentukan order VAR yang memungkinkan plot autokorelasi dan *Akaike Information Criteria* (AIC). Selanjutnya, menjalankan uji *Dickey-Fuller* untuk kriteria stabilitas VAR. Terakhir, penulis mengukur akurasi *error* prediksi dengan MAPE. Keakuratan hasil prediksi dilihat dari nilai MAPE menunjukkan hasil kemampuan peramalan yang sangat baik yaitu rata-rata di bawah 10% [8].

Topik dan Batasannya

Topik dan batasan masalah dalam penelitian Tugas Akhir ini yaitu untuk mengetahui secara pasti apakah penerapan model VAR dapat memprediksi kasus positif harian Covid-19 di Jawa Barat. Data yang digunakan merupakan data kasus terkonfirmasi positif dan tingkat kesembuhan Covid-19 di Jawa Barat yang bersumber dari kawalcovid19.id.

Tujuan

Tujuan dari Tugas Akhir ini adalah untuk memodelkan, memprediksi kasus aktif serta mengukur akurasi *error* Covid-19 di Jawa Barat menggunakan model VAR.

2. Studi Terkait

2.1. Covid-19

Di awal tahun 2020, merebaknya virus jenis baru yang menggemparkan dunia. Penyakit tersebut dinamakan *Coronavirus Disease 2019* (Covid-19). Seperti kita ketahui bersama, sumber virusersebut berasal dari Wuhan, China. Ditemukan pada akhir Desember 2019. World Health Organization memberi nama virus tersebut *Severe Acute Respiratory Syndrome Coronavirus-2* (SARS-CoV-2) dan nama penyakitnya sebagai *Coronavirus Disease 2019* (Covid-19).

Virus ini dapat merusak saraf, saluran pencernaan, hati, dan sistem pernafasan [9]. Virus ini menyebar ke Indonesia dengan cepat. Provinsi Jawa Barat termasuk ke dalam daerah yang terinfeksi Covid-19, bahkan menjadi provinsi dengan kasus terkonfirmasi positif tertinggi ketiga setelah Jawa Timur. Sampai 22 November 2020 terdapat 48.064 kasus positif, sebanyak 866 jiwa dinyatakan meninggal, dan 39.011 jiwa dinyatakan sembuh.

2.2. Uji Kestasioneran Data

Menurut Ekananda, analisis *time series* merupakan serangkaian pengamatan terhadap suatu variabel yang diamati dari waktu ke waktu. Jika data *time series* terdiri dari beberapa variabel, maka digunakan *time series* multivariat yaitu model *Vector Autoregressive* (VAR) [7]. Hill, Griffiths, dan Lim memperkenalkan variabel stasioner adalah variabel yang tidak mengalami tren dan penyebarannya tidak beraturan tanpa dapat kembali ke nilai rata-ratanya atau dapat dikatakan variabel stasioner adalah variabel yang datanya menyebar di sekitar rata-rata nilai variabel [8]. Penulis melakukan uji *Augmented Dickey Fuller* (ADF) untuk mengevaluasi data stasioner dengan hipotesis H_0 menyatakan bahwa data tersebut non-stasioner. H_0 ditolak jika nilai statistik uji ADF lebih kecil dari nilai kritis ADF yaitu 5%. Apabila H_0 ditolak, maka data tersebut telah stasioner. Data yang tidak stasioner dapat diubah menjadi stasioner melalui proses *differencing* yang di definisikan sebagai berikut.

$$\nabla y = y_t - y_{t-1} \quad (1)$$

Dimana y_t adalah nilai yang diamati pada waktu t . Menurut Salmon, *differencing* adalah menormalkan nilai data pada suatu periode dengan nilai periode sebelumnya yang bertujuan agar data menjadi stasioner [10].

2.3. Model Vector Autoregressive (VAR)

Menurut Gujarati, VAR merupakan pemodelan persamaan yang memiliki beberapa variabel endogen secara bersamaan, namun masing-masing endogen dijelaskan oleh *lag* dari nilainya sendiri dan variabel endogen lainnya dalam model. *Lag* sendiri adalah selisih antara sampel ke-*i* dengan sampel ke-*i-1*. Model VAR digunakan jika data stasioner [7]. Secara umum, berikut ini adalah model VAR(*p*) untuk *m* -variabel yang dapat didefinisikan sebagai berikut.

$$y_t = A_0 + A_1y_{t-1} + A_2y_{t-2} + A_3y_{t-3} + \dots + A_p y_{t-p} + e_t \quad (2)$$

Dimana,

$y_t = (m \times 1)$ vektor nilai observasi pada waktu *t* mengikuti model VAR(*p*)

p = Order VAR

$A_0 = (m \times 1)$ vektor konstanta

$A_i = (m \times m)$ matrix parameter VAR, $i = 1, \dots, p$

$e_t = (m \times 1)$ vektor residual dengan matrix kovarians, Σ_e

Untuk menentukan order-*p*, kita dapat menilai dengan menggunakan konsep autokorelasi deret waktu univariat, jumlah *lag* sebelum *cut-off* pertama mempertimbangkan menjadi order *autoregresif*. Alternatifnya, model VAR terbaik dari beberapa yang sesuai adalah dengan mengamati nilai minimum *Akaike Information Criteria* (AIC) yang merupakan kriteria untuk memilih model terbaik. Semakin kecil nilai AIC dalam suatu model, maka semakin baik model tersebut [11]. Kuantifikasi nilai AIC adalah sebagai berikut.

$$AIC = \ln |\Sigma p| + \frac{2k^2 p}{n} \quad (3)$$

Dimana, $\ln |\Sigma p|$ adalah matriks prediksi variansi-kovariansi residual, *k* menyajikan jumlah parameter yang diprediksi dalam model, dan *n* sesuai dengan jumlah pengamatan.

2.4. Prediksi Error dengan MAPE

Menurut Singgih, *Mean Absoulte Percentage Error* (MAPE) dihitung dengan menggunakan kesalahan absolut pada tiap periode dibagi dengan nilai observasi untuk periode tersebut [12]. Menurut Heizer dan Render, kemampuan peramalan sangat baik jika nilai MAPE kurang dari 10 [12]. Secara sistematis, rumusnya ditulis sebagai berikut.

$$MAPE = \frac{1}{n} \sum_{t=1}^n \frac{|A_t - F_t|}{A_t} \quad (4)$$

Dimana,

A_t : Nilai aktual pada data *t*

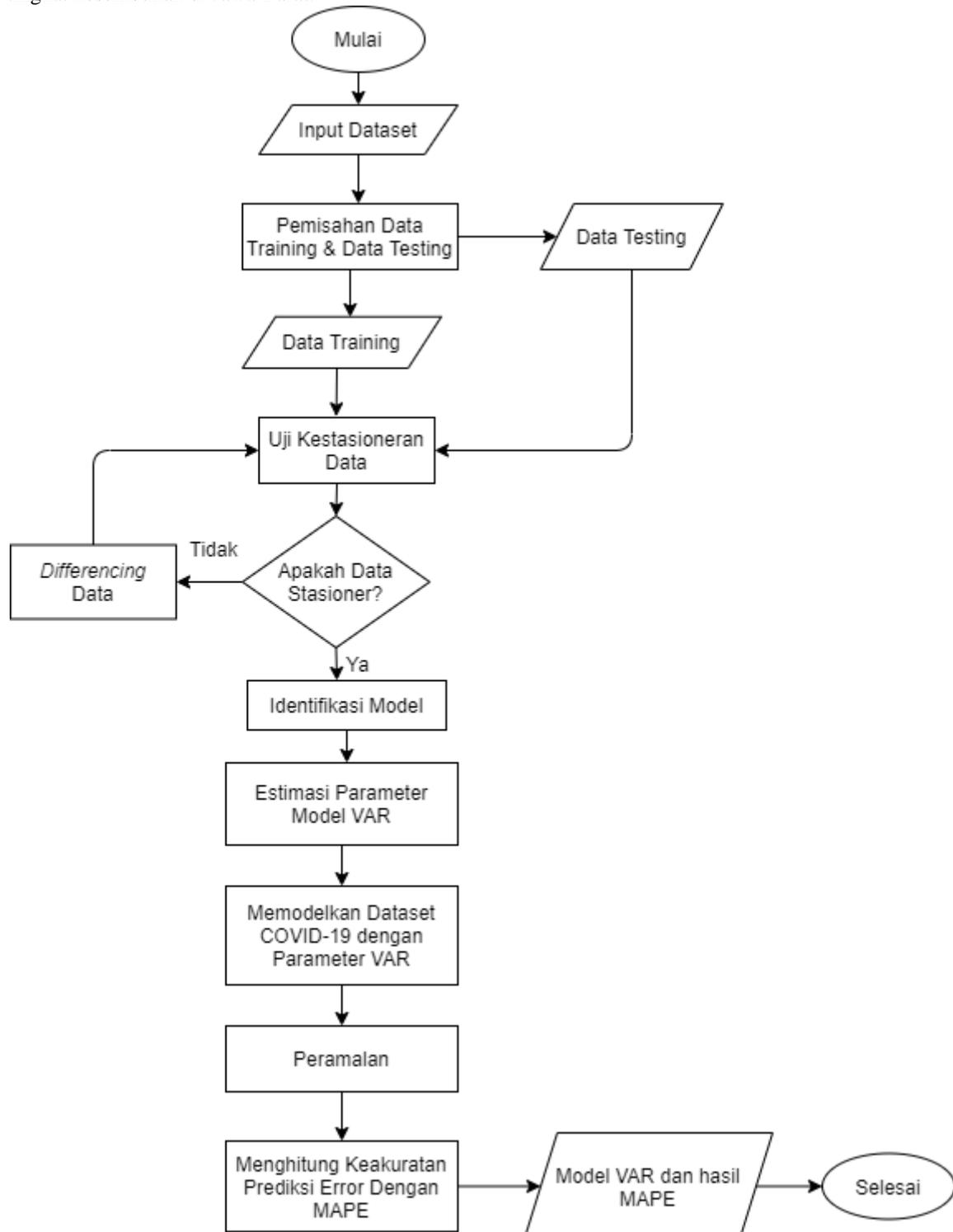
F_t : Nilai peramalan pada data *t*

n : Jumlah periode data

3. Sistem yang Dibangun

3.1. Rancangan Sistem

Secara umum, sistem yang akan dibangun dalam penelitian akan dimodelkan menggunakan model *Vector Autoregressive* (VAR). Data yang digunakan merupakan data kasus terkonfirmasi positif Covid-19 serta data tingkat kesembuhan di Jawa Barat.



Gambar 1. Flowchart Rancangan Sistem

3.2. Proses Perancangan Sistem

Berikut merupakan proses perancangan sistem:

1. Data yang digunakan adalah data kasus terkonfirmasi positif Covid-19 serta tingkat kesembuhan di Jawa Barat yang diperoleh dari kawalcovid19.id pada tanggal 1 Juli 2020 hingga 31 Januari 2021.
2. Dataset dipisah menjadi data *training* dan data *testing*, hal ini bertujuan untuk digunakan dalam *forecasting* observasi berikutnya. Pada pemisahan data ini menggunakan perbandingan 90:10. Data *training* digunakan untuk membuat model, sedangkan data *testing* untuk menguji performa berdasarkan data *training*.
3. Pada tahapan ini dilakukan uji kestasioneran data yang telah dijelaskan pada bab sebelumnya. Data dikatakan stasioner apabila jika rata-ratanya cenderung konstan dari waktu ke waktu dan bersifat stabil. Apabila belum stasioner, maka dapat dilakukan *differencing* data.
4. Nilai selisih yang diperoleh dicek lagi apakah stasioner atau tidak. Jika belum stasioner maka dilakukan *differencing* lagi.
5. Identifikasi order pada model VAR menggunakan metode AIC.
6. Pada tahapan estimasi parameter perlu dilakukan penentuan parameter pada model, sehingga memperoleh nilai yang optimal dengan menggunakan metode AIC.
7. Melakukan pemodelan dataset Covid-19 di provinsi Jawa Barat dengan menggunakan parameter VAR.
8. Melakukan peramalan terhadap kasus terkonfirmasi positif dan tingkat kesembuhan dengan model VAR. Setelah itu, membandingkan hasil peramalan dengan data aslinya.
9. Menggunakan MAPE untuk mengukur akurasi *error* prediksi pada model.
10. Penelitian akan menghasilkan model VAR dan keakuratan hasil prediksi dengan MAPE.

4. Evaluasi

4.1 Uji Kestasioneran Data dengan Augmented Dickey Fuller Test

Model VAR menggambarkan hubungan dinamis antara variabel stasioner. Data stasioner merupakan persyaratan penting untuk analisis deret waktu. Tabel 1 menunjukkan hasil pengujian *Augmented Dickey-Fuller* (ADF), dimana dataset aktif dan sembuh belum stasioner. Hal ini ditunjukkan dengan nilai *P-Value* lebih besar dari 5%. Oleh karena itu, perlu dilakukan proses *differencing*.

Table 1. Hasil Uji Akar Unit dengan ADF-Test

| <i>Variable</i> | <i>t-Statistic</i> | <u><i>Critical Value 5%</i></u> | <i>P-Value</i> |
|-----------------|--------------------|---------------------------------|----------------|
| Aktif | 2.004 | -2.877 | 0.9987 |
| Sembuh | 2.0263 | -2.877 | 0.9987 |

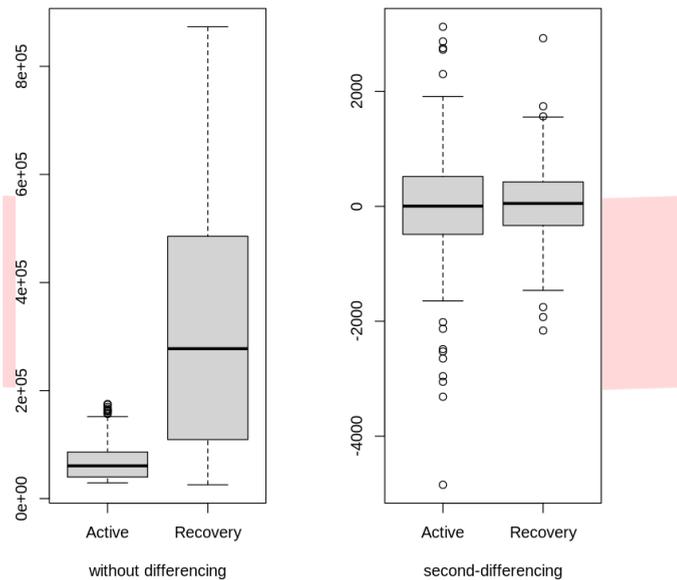
Table 2. Hasil Uji Akar Unit Data Diferensi 1 dengan ADF-Test

| <i>Variable</i> | <i>t-Statistic</i> | <u><i>Critical Value 5%</i></u> | <i>P-Value</i> |
|-----------------|--------------------|---------------------------------|----------------|
| Aktif | -2.2238 | -2.877 | 0.1977 |
| Sembuh | 0.0291 | -2.877 | 0.1977 |

Table 3. Hasil Uji Akar Unit Data Diferensi 2 dengan ADF-Test

| <i>Variable</i> | <i>t-Statistic</i> | <u><i>Critical Value 5%</i></u> | <i>P-Value</i> |
|-----------------|--------------------|---------------------------------|----------------|
| Aktif | -9.7883 | -2.877 | 0.0 |
| Sembuh | -10.9835 | -2.877 | 0.0 |

Berdasarkan P -Value yang ditunjukkan pada Tabel 2, perbedaan pertama dari dataset aktif dan sembuh mengarah ke proses *non*-stasioner. Hal ini ditunjukkan dengan nilai P -Value lebih besar dari 5%. Oleh karena itu, perlu dilakukan *second differencing*. Berdasarkan hasil pada Tabel 3, bahwa P -Value kurang dari 5%, maka H_0 ditolak atau dapat dikatakan telah stasioner untuk kedua data: variabel aktif dan variabel sembuh. Selain itu, sesuai dengan hasil ADF, Gambar 2 menunjukkan jika nilai median sama dengan nilai rata-rata atau tepat di tengah, dan nilai maksimum jarak Q1 sama dengan nilai maksimum Q3 maka data stasioner. Pada gambar tanpa *differencing* ada perbedaan jarak antara Q1 dan Q3. Namun setelah dilakukan *second differencing* walaupun masih terdapat *outlier* terlihat jarak dari nilai minimum ke Q1 sama dengan jarak dari Q3 ke nilai maksimum.



Gambar 2. Karakteristik Boxplot antara Dataset tanpa *Differencing* dan dengan *Differencing* kedua

Tujuan dari model identifikasi adalah untuk menemukan order model *autoregresif*. Order tersebut diperoleh dengan identifikasi menggunakan AIC. *Lag* dengan nilai AIC terkecil diusulkan sebagai order model *autoregressive*.

4.2 Penentuan Panjang Lag Optimum

Table 4. Kriteria Pemilihan Order

| Lag | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----|--------|--------|--------|--------|---------------|--------|--------|--------|--------|
| AIC | 25.039 | 25.019 | 24.956 | 24.945 | 24.789 | 24.825 | 24.857 | 24.828 | 24.869 |

Berdasarkan Tabel 4, maka panjang lag optimum adalah lag 5, dimana lag 5 memiliki nilai AIC yang minimum dibandingkan dengan nilai AIC lag lain. Artinya, model VAR yang terbentuk adalah model VAR dengan order 5 atau VAR(5).

$$\nabla^2 y_t = A_0 + \sum_{i=1}^5 A_i \nabla^2 y_{t-i} + e_t \quad (5)$$

Table 5. Estimasi Paramter VAR (5) – Kasus Aktif

| | Coefficient | t-stat | Prob |
|------------|-------------|--------|--------|
| Const | 15.847478 | 0.286 | 0.775 |
| L1. Aktif | -0.278083 | -2.810 | 0.005* |
| L1. Sembuh | 0.508343 | 3.265 | 0.001* |
| L2. Aktif | -0.285364 | -2.702 | 0.007 |
| L2. Sembuh | 0.214206 | 1.194 | 0.232 |
| L3. Aktif | -0.360966 | -3.554 | 0.000* |
| L3. Sembuh | -0.023281 | -0.129 | 0.897 |
| L4. Aktif | -0.246349 | -2.271 | 0.023 |
| L4. Sembuh | -0.072074 | -0.393 | 0.694 |
| L5. Aktif | -0.451795 | -4.307 | 0.000* |
| L5. Sembuh | -0.494546 | -3.041 | 0.002* |

Table 6. Estimasi Parameter VAR(5) – Kasus Sembuh

| | Coefficient | t-stat | Prob |
|------------|-------------|--------|--------|
| Const | 93.605538 | 2.527 | 0.011 |
| L1. Aktif | -0.057235 | -0.866 | 0.386 |
| L1. Sembuh | -0.758458 | -7.295 | 0.000* |
| L2. Aktif | 0.001471 | 0.021 | 0.983 |
| L2. Sembuh | -0.416869 | -3.480 | 0.001* |
| L3. Aktif | -0.047782 | -0.704 | 0.481 |
| L3. Sembuh | -0.350568 | -2.913 | 0.004* |
| L4. Aktif | -0.028707 | -0.396 | 0.692 |
| L4. Sembuh | -0.267519 | -2.185 | 0.029 |
| L5. Aktif | -0.008401 | -0.120 | 0.905 |
| L5. Sembuh | -0.036203 | -0.333 | 0.739 |

Tabel 5 dan Tabel 6 menunjukkan probabilitas kontribusi masing-masing prediktor terhadap variabel respon. Tanda (*) menunjukkan bahwa nilai signifikan kurang dari 5%. Model VAR(5) untuk kasus aktif lebih banyak memiliki variabel prediktor yang signifikan dibandingkan dengan VAR(5) untuk kasus sembuh. Pada kasus aktif, variabel yang signifikan mempengaruhi adalah kasus aktif pada *lag* 1, *lag* 3, dan *lag* 5, serta kasus sembuh pada *lag* 1 dan *lag* 5. Sedangkan, pada kasus sembuh, variabel yang signifikan mempengaruhi adalah kasus sembuh pada *lag* 1, *lag* 2, dan *lag* 3.

4.3. Pemodelan Vector Autoregressive (VAR)

Model VAR(5) yang terbentuk adalah:

$$\nabla^2 y_{1,t} = \text{const} + a_1 \nabla^2 y_{1,t-1} + a_2 \nabla^2 y_{1,t-2} + a_3 \nabla^2 y_{1,t-3} + a_4 \nabla^2 y_{1,t-4} + a_5 \nabla^2 y_{1,t-5} + \beta_1 \nabla^2 y_{2,t-1} + \beta_2 \nabla^2 y_{2,t-2} + \beta_3 \nabla^2 y_{2,t-3} + \beta_4 \nabla^2 y_{2,t-4} + \beta_5 \nabla^2 y_{2,t-5}$$

$$\nabla^2 y_{2,t} = \text{const} + a_1 \nabla^2 y_{1,t-1} + a_2 \nabla^2 y_{1,t-2} + a_3 \nabla^2 y_{1,t-3} + a_4 \nabla^2 y_{1,t-4} + a_5 \nabla^2 y_{1,t-5} + \beta_1 \nabla^2 y_{2,t-1} + \beta_2 \nabla^2 y_{2,t-2} + \beta_3 \nabla^2 y_{2,t-3} + \beta_4 \nabla^2 y_{2,t-4} + \beta_5 \nabla^2 y_{2,t-5}$$

Dari persamaan dan hasil yang diperoleh bahwa variabel pada persamaan VAR yaitu $\nabla^2 y_{1,t}$ adalah kasus aktif *differencing* dua dan $\nabla^2 y_{2,t}$ adalah kasus sembuh *differencing* dua. Berdasarkan pengolahan data, diperoleh model VAR(5) untuk variabel aktif dan sembuh sebagai berikut:

$$\begin{aligned} \nabla^2 y_{1,t} &= 15.847 - 0.278 \nabla^2 y_{1,t-1} + 0.508 \nabla^2 y_{1,t-2} - 0.285 \nabla^2 y_{1,t-3} + 0.214 \nabla^2 y_{1,t-4} - 0.360 \nabla^2 y_{1,t-5} \\ &\quad - 0.023 \nabla^2 y_{2,t-1} - 0.264 \nabla^2 y_{2,t-2} - 0.072 \nabla^2 y_{2,t-3} - 0.415 \nabla^2 y_{2,t-4} - 0.494 \nabla^2 y_{2,t-5} \\ \nabla^2 y_{2,t} &= 93.605 - 0.057 \nabla^2 y_{1,t-1} - 0.758 \nabla^2 y_{1,t-2} + 0.001 \nabla^2 y_{1,t-3} - 0.416 \nabla^2 y_{1,t-4} - 0.0047 \nabla^2 y_{1,t-5} \\ &\quad - 0.350 \nabla^2 y_{2,t-1} - 0.028 \nabla^2 y_{2,t-2} - 0.267 \nabla^2 y_{2,t-3} - 0.008 \nabla^2 y_{2,t-4} - 0.036 \nabla^2 y_{2,t-5} \end{aligned}$$

Dengan demikian, kita dapat menjalankan *forecasting* kasus aktif dan sembuh secara bersamaan selama 23 hari ke depan dari tanggal 10 hingga 31 Januari 2021.

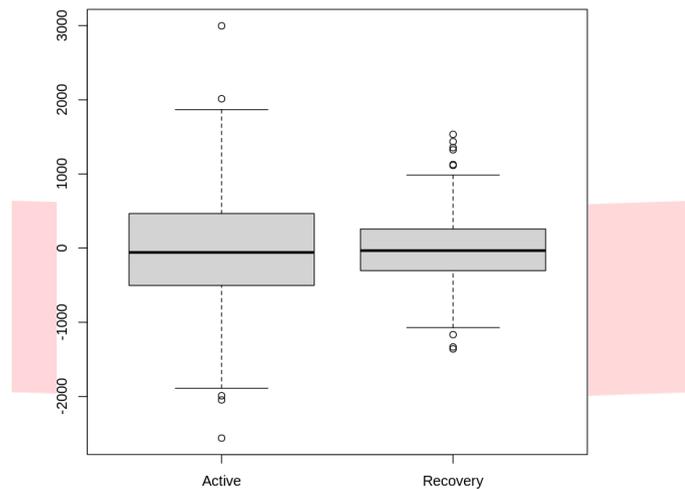
Table 7. Data Asli Covid-19 di Jawa Barat

| Tanggal | Kasus Aktif | Kasus Sembuh |
|------------|-------------|--------------|
| 10-01-2021 | 35349 | 33529 |
| 11-01-2021 | 35764 | 34719 |
| 12-01-2021 | 36455 | 35638 |
| 13-01-2021 | 36636 | 36689 |
| 14-01-2021 | 37226 | 37636 |
| 15-01-2021 | 37247 | 39050 |
| 16-01-2021 | 37450 | 40345 |
| 17-01-2021 | 37339 | 41834 |
| 18-01-2021 | 37598 | 43268 |
| 19-01-2021 | 36977 | 45401 |
| 20-01-2021 | 36998 | 46977 |
| 21-01-2021 | 37083 | 46977 |
| 22-01-2021 | 37083 | 48466 |
| 23-01-2021 | 37037 | 50255 |
| 24-01-2021 | 36808 | 53945 |
| 25-01-2021 | 37218 | 55354 |
| 26-01-2021 | 37342 | 56655 |
| 27-01-2021 | 37292 | 58173 |
| 28-01-2021 | 36611 | 60539 |

Table 8. Hasil forecasting untuk 23 hari berikutnya

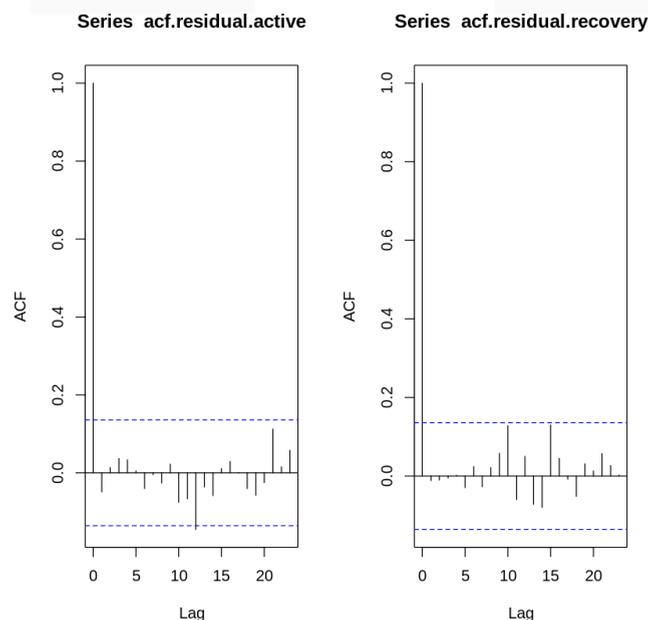
| Tanggal | Kasus Aktif | Kasus Sembuh |
|----------------|--------------------|---------------------|
| 10-01-2021 | 122854 | 680490 |
| 11-01-2021 | 124066 | 687428 |
| 12-01-2021 | 125405 | 694487 |
| 13-01-2021 | 126826 | 701680 |
| 14-01-2021 | 129127 | 708829 |
| 15-01-2021 | 131725 | 715981 |
| 16-01-2021 | 134267 | 723147 |
| 17-01-2021 | 136300 | 730328 |
| 18-01-2021 | 138093 | 737581 |
| 19-01-2021 | 139723 | 744862 |
| 20-01-2021 | 141575 | 752209 |
| 21-01-2021 | 143693 | 759572 |
| 22-01-2021 | 146047 | 766963 |
| 23-01-2021 | 148323 | 766963 |
| 24-01-2021 | 150486 | 781808 |
| 25-01-2021 | 152454 | 789279 |
| 26-01-2021 | 154386 | 796798 |
| 27-01-2021 | 156369 | 804352 |
| 28-01-2021 | 158515 | 811942 |
| 29-01-2021 | 160735 | 819557 |
| 30-01-2021 | 162982 | 827201 |
| 31-01-2021 | 165155 | 834872 |

Berdasarkan hasil *forecasting* Tabel 8, pemodelan dengan VAR(5) menyatakan bahwa jumlah kasus di Jawa Barat akan meningkat secara signifikan dalam 23 hari ke depan. Ketepatan peramalan diukur dengan menghitung statistik *Mean Absolute Percentage Error* (MAPE). Statistik MAPE untuk data pengujian masing-masing adalah 4,67% untuk kasus aktif dan 1,77% untuk kasus sembuh. Hal tersebut mungkin bisa diatasi untuk mengantisipasi potensi meningkatnya jumlah kasus harian Covid-19 dengan MAPE kurang dari 10%. Dapat dilihat dari peramalan tren bahwa semakin banyak prediksi, semakin besar kesalahannya.



Gambar 3. Boxplot mengungkapkan karakteristik peramalan residual antara kasus aktif dan sembuh.

Selain itu, menarik untuk mengamati karakteristik residual model VAR(5), untuk mengamati residual yang memenuhi distribusi *white noise* yang *independent* dan identik dengan mean nol. Seperti yang ditunjukkan pada Gambar 3, nilai sisa memiliki nilai median yang sama dengan rata-rata yaitu, 0. Oleh karena itu, residual dianggap memenuhi elemen yang identik. Juga, plot autokorelasi pada Gambar 4 dari masing-masing residual menunjukkan bahwa tidak ada autokorelasi antara residual berdasarkan jeda waktu kecuali pada *lag* 0. Oleh karena itu, autokorelasi residual memenuhi persyaratan *independent*.



Gambar 4. Plot sisa autokorelasi dari peramalan aktif dan kasus sembuh

5. Kesimpulan

Saat ini, banyak negara yang masih bekerja keras untuk mengatasi dan mengendalikan wabah Covid-19 ini. Ada beberapa aspek, termasuk upaya publik untuk memperlambat penyebaran dan peneliti untuk mengamati lebih lanjut tentang virus ini. Karena sejumlah kasus positif telah mengakibatkan volatilitas yang belum pernah terjadi sebelumnya, peramalan kasus harian wabah Covid-19 menjadi hal yang krusial untuk mendukung informasi dan memelihara layanan kesehatan esensial. Simulasi dalam penelitian ini membahas model VAR(5) dengan *differencing* sebanyak dua kali yang dapat ditangkap tren dan memprediksi kasus Covid-19 harian terutama model kasus aktif berdasarkan kasus sembuh dengan MAPE sebesar 4,7% dan menggunakan nilai minimum AIC sebesar 24,789. Model VAR ini memiliki karakteristik residual yang memiliki nilai median, yaitu 0. Sehingga dianggap memenuhi syarat elemen identik. Model ini dapat membantu pemerintah, kesehatan masyarakat, dan pemangku kepentingan untuk mengambil keputusan dalam mempersiapkan sumber dayanya. Oleh karena itu, di bawah tingginya volatilitas tingkat kenaikan wabah ini, penulis menemukan model yang tepat untuk memprediksi Covid-19 di Indonesia.



REFERENSI

- [1] G. P. Nungsi and K. A. Gusa, "Kebijakan Rapid Test Drive Thru Oleh Pemerintah Daerah Provinsi Jawa Barat," *Jurnal Kebijakan Pemerintah*, vol. 3, no. 1, pp. 37-42, 2020.
- [2] R. N. Putri, "Indonesia dalam Menghadapi Pandemi Covid-19," *Jurnal Imiah Universitas Batanghari Jambi*, vol. 20, no. 2, pp. 705-709, Juli 2020.
- [3] S. Anggara, S. Afiyah, A. S. Farida and J. Muslim, "Inovasi Kebijakan Publik Tentang Pencegahan Dan Penanggulangan Corona Virus Disease 19 (Covid-19) Di Jawa Barat," *Digital Library Uin Sunan Gunung Djati*, 2020.
- [4] M. Maleki, M. R. Mahmoudi and M. H. Heydari, "Modeling and forecasting the spread and death rate of coronavirus," *Elsevier*, vol. 140, July 2020.
- [5] C. A. Sims, *Macroeconomics and Reality*, The Econometric Society, 1980.
- [6] D. R. Febrianti, M. A. Tiro and S. , "Metode Vector Autoregressive (VAR) dalam Menganalisis Pengaruh," *Journal of Statistics and Its Application on Teaching and Research*, vol. 3, no. 1, pp. 23-30, 2021.
- [7] H. H. Usman, I. Djakaria and M. R. F. Payu, "Pendekatan Model Vector Autoregressive (VAR) Untuk Meramalkan Faktor-Faktor Yang Mempengaruhi Inflasi Di Provinsi Gorontalo," *Jambura Journal of Probability and Statistics*, no. 1, Mei 2020.
- [8] Y. "Corona Virus Diseases (Covid-19)," *Wellness And Healthy Magazine*, vol. 2, no. 1, pp. 187-192, February 2020.
- [9] H. R. G. W. and L. G. , in *Principles of Econometrics*, Wiley, 2011.
- [10] K. Mendome, N. Nainggolan and J. Kekenusa, "Penerapan Model ARIMA dalam Memprediksi Jumlah Tindak Kriminalitas di Wilayah POLRESTA Manado," *Jurnal MIPA UNSRAT*, vol. 5, no. 2, pp. 113-116, Agustus 2016.
- [11] H. Lutkepohl, *New Introduction to Multiple Time Series Analysis*, Berlin: Springer, 2005.
- [12] A. Rahmawati, D. A. I. Maruddani and A. Hoyyi, "Structural Vector Autoregressive Untuk Analisis Dampak Shock Nilai Tukar Rupiah Terhadap Dolar Amerika Serikat Pada Indeks Harga Saham Gabungan," *Jurnal Gaussian*, vol. 6, no. 3, pp. 291-302, 2017.
- [13] H. D. E. Sinaga and N. Irawati2, "Perbandingan Double Moving Average Dengan Double Exponential Smoothing Pada Peramalan Bahan Medis Habis Pakai," *Jurnal Teknologi dan Sistem Informasi*, vol. 4, no. 2, pp. 197-204, 2018.
- [14] D. N. Gujarati, *Basic Econometrics Fourth Edition*, New York: Gary Burke, 2003.