

Peramalan Kualitas Udara di Surabaya Menggunakan Model *Vector Autoregressive Integrated Moving Average* (VARIMA)

Safira Azizah dan Mike Prastuti

Departemen Statistika Bisnis, Institut Teknologi Sepuluh Nopember (ITS)

e-mail: mike.prastuti@gmail.com

Abstrak—Pada tahun 2021, Surabaya meraih penghargaan di ASEAN ESC pada kategori Udara Terbersih Kota Besar, namun pada Juni 2022, Surabaya menduduki 10 kota dengan kualitas udara terburuk di Indonesia. Oleh karena itu, diduga kualitas udara di Surabaya mengalami fluktuasi sehingga perlu dilakukan analisis peramalan. Pada penelitian ini akan digunakan model VARIMA untuk meramalkan kualitas udara di Surabaya menggunakan PM₁₀, NO₂, dan O₃. Hasil yang didapatkan adalah pada tanggal 10 April hingga 10 Oktober, konsentrasi PM₁₀ berada di kategori baik hingga sedang, namun beberapa kali masuk pada kategori tidak sehat, yaitu pada tanggal 12 April, 31 Mei, 2 Juni dan tertinggi berada pada tanggal 3 Juni. Konsentrasi NO₂ berada pada kategori baik hingga sedang, namun pada tanggal 31 Mei berada pada kategori tidak sehat. Konsentrasi O₃ berada pada kategori baik. Peramalan kualitas udara di Surabaya pada tanggal 11 Oktober hingga 25 Oktober 2022 menggunakan model VARIMA (3,1,0) menunjukkan bahwa konsentrasi PM₁₀ berada pada kategori sedang sedangkan konsentrasi NO₂ dan O₃ berada pada kategori baik, sehingga tingkat kualitas udara masih dapat diterima pada kesehatan manusia, hewan, dan tumbuhan.

Kata Kunci—Kualitas Udara, Peramalan, Surabaya, VARIMA.

I. PENDAHULUAN

SURABAYA sebagai kota metropolitan kedua di Indonesia yang mempunyai banyak kegiatan di bidang industri dan transportasi. Hal tersebut membuat udara di Surabaya menjadi rawan tercemar. Pada tahun 2021, Surabaya meraih penghargaan ASEAN *Environmentally Sustainable City* (ESC) pada kategori Udara Terbersih Kota Besar, sehingga Surabaya diakui sebagai kota yang memiliki udara terbersih se-Asia Tenggara, namun pada Juni 2022 Surabaya menduduki peringkat ke-4 pada 10 kota dengan kualitas udara terburuk di Indonesia.

Pada tahun 2022, Organisasi Kesehatan Dunia (WHO) mengatakan bahwa terdapat 7 juta kematian yang disebabkan oleh polusi udara. Selain itu, polusi udara juga dapat menyebabkan berbagai jenis penyakit pernapasan. Oleh karena itu, perlu dilakukan pencegahan dan perlindungan diri. Salah satu hal yang dapat dilakukan oleh masyarakat dalam melakukan perlindungan diri adalah mengetahui kualitas udara yang ada di sekitarnya. Peraturan Menteri Lingkungan Hidup dan Kehutanan nomor 14 tahun 2020 menyebutkan bahwa kualitas udara dapat diketahui menggunakan Indeks Standar Pencemar Udara (ISPU). ISPU terdiri dari parameter partikular (PM₁₀), partikular (PM_{2.5}), karbon monoksida (CO), nitrogen dioksida (NO₂), sulfur dioksida (SO₂), ozon (O₃), dan hidrokarbon (HC).

Putri (2021) dalam penelitiannya dengan judul “Kajian Hubungan Dampak Paparan NO₂, O₃, dan PM₁₀ Terhadap Kejadian Penyakit ISPA Pneumonia dan Non-Pneumonia di Ke-

bonsari, Surabaya dan Sekitarnya” menyebutkan bahwa konsentrasi NO₂ dan O₃ memiliki hubungan dengan penyakit ISPA Pneumonia [1] dan Prasetyo (2021) dalam tugas akhirnya dengan judul “Kajian Konsentrasi NO₂, O₃, dan PM₁₀ di Udara Terhadap Kejadian Penyakit ISPA Pneumonia dan Non-Pneumonia di Wonorejo, Surabaya dan Sekitarnya” menyebutkan bahwa konsentrasi O₃ dan PM₁₀ memiliki hubungan dengan penyakit ISPA Pneumonia sehingga dilakukan peramalan terhadap PM₁₀, NO₂, dan O₃ menggunakan analisis *time series* [2].

Tujuan dari penelitian adalah mengetahui model peramalan terbaik menggunakan model VARIMA pada kualitas udara di Surabaya berdasarkan konsentrasi PM₁₀, NO₂, dan O₃ dan mengetahui hasil peramalan kualitas udara di Surabaya pada periode selanjutnya menggunakan model peramalan terbaik yang didapat.

II. TINJAUAN PUSTAKA

A. *Multivariate Time Series*

Data *time series multivariate* merupakan pengamatan lebih dari satu variabel yang memiliki autokorelasi dan juga memiliki hubungan dengan variabel lainnya. Pada *multivariate time series* dilakukan berdasarkan *Matrix Cross Correlation Function* (MCCF) dan *Matrix Partial Cross Correlation Function* (MPCCF) setelah data stasioner. Salah satu model yang dapat digunakan untuk melakukan analisis pada *multivariate time series* adalah model *Vector Autoregressive Integrated Moving Average* (VARIMA) [3].

1) *Identifikasi Stasioneritas*

Pemodelan VARIMA dapat dilakukan jika data yang digunakan telah stasioner terhadap mean. Stasioneritas data *multivariate time series* terhadap *mean* secara visual dapat dilihat berdasarkan hasil plot MCCF [3]. Persamaan matriks korelasi sampel pada *vector time series* dengan banyak observasi sebesar n yang dituliskan sebagai $\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n$ ditunjukkan pada persamaan berikut.

$$\hat{\rho}(k) = [\hat{\rho}_{ij}(k)] \quad (1)$$

Matriks $\hat{\rho}(k)$ merupakan korelasi silang sampel pada komponen ke- i dan ke- j yang dijelaskan pada persamaan berikut.

$$\hat{\rho}_{ij}(k) = \frac{\sum_{t=1}^{n-k} (z_{i,t-k} - \bar{z}_i)(z_{j,t} - \bar{z}_j)}{[\sum_{t=1}^{n-k} (z_{i,t-k} - \bar{z}_i)^2 (z_{j,t} - \bar{z}_j)^2]^{\frac{1}{2}}} \quad (2)$$

keterangan: $\hat{\rho}_{ij}(k)$ merupakan korelasi silang sampel acak variabel ke- i dan ke- j pada lag ke- k , \bar{z}_i merupakan rata-rata sampel variabel komponen ke- i , dan \bar{z}_j merupakan rata-rata sampel variabel komponen ke- j .

Stasioneritas vektor menunjukkan $\hat{\rho}(k)$ merupakan esti-

mator yang konstan. Apabila dimensi vektor bertambah, maka bentuk matriks semakin kompleks, sehingga Box dan Tiao (1981) memperkenalkan metode meringkas hasil sampel korelasi dengan menggunakan simbol (+), (-), dan (.) pada baris ke- i dan kolom ke- j pada matriks sampel korelasi sebagai berikut: (1) Simbol (+) menotasikan nilai $\hat{\rho}_{ij}(k)$ lebih besar dari 2 kali nilai estimasi *standard error* dan menunjukkan adanya hubungan korelasi positif antara komponen ke- i dan ke- j . (2) Simbol (-) menotasikan nilai $\hat{\rho}_{ij}(k)$ lebih kecil dari -2 kali nilai estimasi *standard error* atau adanya hubungan korelasi negatif antara komponen ke- i dan ke- j . (3) Simbol (.) menotasikan nilai $\hat{\rho}_{ij}(k)$ berada di antara -2 sampai 2 kali nilai estimasi *standard error* yang artinya tidak terdapat hubungan korelasi antara komponen ke- i dan ke- j [4]. Nilai estimasi *standard error* diperoleh pada persamaan berikut,

$$SE = \frac{1}{\sqrt{n-k}} \tag{3}$$

dengan n adalah jumlah data pengamatan, dan k adalah lag pada data. Apabila data tidak stasioner dalam *mean*, maka perlu dilakukan *differencing* data dengan cara sebagai berikut,

$$Y_t = Z_t - Z_{t-1} \tag{4}$$

dengan Y_t adalah data hasil *differencing*, Z_t adalah data pada periode ke- t , dan Z_{t-1} adalah data pada periode ke- $(t - 1)$.

2) Model VARIMA

Dalam analisis *time series*, sangat umum untuk mengamati deret yang menunjukkan perilaku tidak stasioner. Salah satu cara yang berguna untuk mereduksi deret non stasioner menjadi deret stasioner adalah dengan cara melakukan *differencing* [3]. Model VARIMA merupakan model VARMA yang tidak stasioner. Model VARIMA secara umum untuk Z_t adalah sebagai berikut,

$$\Phi_p(B)D(B)Z_t = \Theta_q(B)a_t \tag{5}$$

dengan

$$\Phi_p(B) = \Phi_0 - \Phi_1 B - \Phi_2 B^2 - \dots - \Phi_p B^p \tag{6}$$

$$\Theta_q(B) = \theta_0 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q \tag{7}$$

dan

$$D(B) = \begin{bmatrix} (1-B)^{d_1} & 0 & \dots & 0 & 0 \\ 0 & (1-B)^{d_2} & \ddots & 0 & \vdots \\ \vdots & \ddots & \ddots & 0 & 0 \\ 0 & \dots & 0 & 0 & (1-B)^{d_m} \end{bmatrix} \tag{8}$$

dengan d_1, d_2, \dots, d_m adalah himpunan bilangan bulat tidak-negatif. $\Phi_p(B)$ dan $\Theta_q(B)$ berturut-turut adalah suatu matriks polinomial *autoregressive* dan *moving average* pada orde p dan q , dan Φ_0 dan Θ_0 adalah matriks $m \times m$ non singular dan a_t adalah vektor *error white noise* [3].

a) Identifikasi Model

Identifikasi model pada *multivariate time series* dilakukan berdasarkan MCCF, MPCCF, dan nilai *Akaike Information Criterion* (AIC) terkecil setelah data stasioner. MCCF digunakan untuk mengidentifikasi orde model VMA, MPCCF digunakan untuk mengidentifikasi orde model VAR [3] dan AIC terkecil digunakan untuk mengidentifikasi orde model keduanya. Penjelasan dari MCCF telah dijabarkan pada subbab di atas. Heyse dan Wei (1985) memperluas definisi autokorelasi parsial univariat menjadi deret waktu vektor dan mendapatkan matriks korelasi antara Z_t dan Z_{t+s} setelah

menghapus dependensi linier dari masing-masing pada vektor intervensi Z_{t+1}, Z_{t+2}, \dots , dan Z_{t+s-1} [5]. Matriks korelasi ini didefinisikan sebagai korelasi antara vektor *residual*,

$$\begin{aligned} \mathbf{u}_{(s-1,t+s)} &= \mathbf{Z}_{t+s} - \alpha_{s-1,1}\mathbf{Z}_{t+s-1} - \dots - \alpha_{s-1,s-1}\mathbf{Z}_{t+1} \\ &= \begin{cases} \mathbf{Z}_{t+s} - \sum_{k=1}^{s-1} \alpha_{s-1,k}\mathbf{Z}_{t+k} & , s \geq 2 \\ \mathbf{Z}_{t+s} & , s = 1 \end{cases} \end{aligned} \tag{9}$$

dan

$$\begin{aligned} \mathbf{v}_{s-1,t} &= \mathbf{Z}_t - \beta_{s-1,1}\mathbf{Z}_{t+1} - \dots - \beta_{s-1,s-1}\mathbf{Z}_{t+s-1} \\ &= \begin{cases} \mathbf{Z}_t - \sum_{k=1}^{s-1} \beta_{s-1,k}\mathbf{Z}_{t+k} & , s \geq 2 \\ \mathbf{Z}_t & , s = 1 \end{cases} \end{aligned} \tag{10}$$

dengan $\mathbf{u}_{s-1,t+s}$ adalah *residual* dari regresi antara \mathbf{Z}_{t+s} dengan $\mathbf{Z}_{t+s-1}, \dots, \mathbf{Z}_{t+1}$ dan $\mathbf{v}_{s-1,t}$ adalah *residual* dari regresi antara \mathbf{Z}_t dengan $\mathbf{Z}_{t+s-1}, \dots, \mathbf{Z}_{t+1}$. Heyse dan Wei (1985) mendefinisikan MPCCF pada lag s sebagai berikut [5],

$$P(s) = [D_v(s)]^{-1}V_{vu}(s)[D_u(s)]^{-1} \tag{11}$$

dengan $\mathbf{V}_u(s)$ adalah definisi dari $\text{var}(\mathbf{u}_{s-1,t+s})$, $\mathbf{V}_v(s)$ adalah definisi dari $\text{var}(\mathbf{v}_{s-1,t})$, dan $\mathbf{V}_{vu}(s)$ adalah definisi dari $\text{cov}(\mathbf{u}_{s-1,t+s}, \mathbf{v}_{s-1,t})$. $\mathbf{D}_v(s)$ adalah matriks diagonal dengan elemen diagonal ke- i adalah akar kuadrat dari elemen diagonal ke- i pada $\mathbf{V}_v(s)$ dan $\mathbf{D}_u(s)$ didefinisikan sama untuk $\mathbf{V}_u(s)$ [3]. Box dan Tiao (1981) memperkenalkan metode meringkas hasil dengan menggunakan simbol (+), (-), dan (.) pada baris ke- i dan kolom ke- j pada matriks sampel korelasi sebagai berikut [4]. (1) Simbol (+) menotasikan nilai $\hat{\rho}_{ij}(s)$ lebih besar dari 2 kali nilai estimasi *standard error* dan menunjukkan adanya hubungan korelasi positif antara komponen ke- i dan ke- j . (2) Simbol (-) menotasikan nilai $\hat{\rho}_{ij}(s)$ lebih kecil dari -2 kali nilai estimasi *standard error* atau adanya hubungan korelasi negatif antara komponen ke- i dan ke- j . (3) Simbol (.) menotasikan nilai $\hat{\rho}_{ij}(s)$ berada di antara -2 sampai 2 kali nilai estimasi *standard error* yang artinya tidak terdapat hubungan korelasi antara komponen ke- i dan ke- j . Nilai estimasi *standard error* diperoleh pada persamaan berikut,

$$SE = \frac{1}{\sqrt{n}} \tag{12}$$

dengan n adalah jumlah data pengamatan dan s merupakan lag pada data.

AIC adalah suatu kriteria pemilihan model terbaik yang diperkenalkan oleh Akaike (1973) dengan mempertimbangkan banyaknya parameter dalam model [6]. Dalam mengidentifikasi orde model VARIMA, dapat dilihat dari karakteristik orde spasial dan waktu dengan rumus sebagai berikut,

$$AIC = \ln(\bar{\Sigma}_i) + \frac{2k^2i}{T} \tag{13}$$

dengan $\bar{\Sigma}_i$ merupakan maksimum *likelihood* dari Σ_i , k adalah banyaknya parameter dalam model, T adalah banyaknya observasi, dan i adalah orde dari model VAR atau VAR(i) [7].

b) Estimasi Parameter Model

Metode yang bisa digunakan dalam menaksir parameter, yaitu metode *maximum likelihood*. Jika diketahui matriks parameter model VARIMA adalah $\Phi = (\Phi_1, \Phi_2, \dots, \Phi_p)$ dan $\Theta = (\Theta_1, \Theta_2, \dots, \Theta_q)$, $\mathbf{Z} = (\mathbf{Z}_1, \mathbf{Z}_2, \dots, \mathbf{Z}_n)$ adalah vektor deret waktu dengan matriks varian dan kovariansi Σ , maka

fungsi log *likelihood* adalah sebagai berikut [3].

$$f(\Phi, \theta, \Sigma | Z) = (2\pi)^{-\frac{1}{2}\Sigma^{-1}} \exp\left(a^{\Sigma^{-1}} a\right) \quad (14)$$

$$L(\Phi, \theta, \Sigma | Z) = (f(x_1), f(x_2), \dots, f(x_n)) \\ = (2\pi)^{-\frac{n}{2}} \left| \Sigma \right|^{-\frac{n}{2}} \exp\left(\frac{1}{2} \sum_{t=1}^n \mathbf{a}_t' \Sigma^{-1} \mathbf{a}_t\right) \quad (15)$$

$$\ln L(\Phi, \theta, \Sigma | Z) = -\frac{nm}{2} \ln 2\pi - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \sum_{t=1}^n \mathbf{a}_t' \Sigma^{-1} \mathbf{a}_t \\ = -\frac{nm}{2} \ln 2\pi - \frac{n}{2} \ln |\Sigma| - \frac{1}{2} \text{tr} \Sigma^{-1} s(\Phi, \theta) \quad (16)$$

dengan,

$$\mathbf{a}_t = Z_t - \Phi_1 Z_{t-1} - \Phi_2 Z_{t-2} - \dots - \Phi_p Z_{t-p} + \Theta_1 a_{t-1} + \Theta_2 a_{t-2} + \dots + \Theta_q a_{t-q} \quad (17)$$

$$s(\Phi, \theta) = \sum_{t=1}^n \mathbf{a}_t \mathbf{a}_t' \quad (18)$$

$$\hat{\Phi} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} \quad (19)$$

dengan,

$$X = \begin{bmatrix} 1 & Z'_p & \dots & Z'_1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & Z'_{n-1} & \dots & Z'_n \end{bmatrix}, Y = \begin{bmatrix} Z'_{p+1} \\ \vdots \\ Z'_n \end{bmatrix} \quad (20)$$

m adalah jumlah variabel, dan n adalah jumlah deret data. Sedangkan untuk estimasi parameter Θ_q , hanya bisa dicari menggunakan iterasi Newton-Raphson.

c) Uji Signifikansi Parameter Model

Uji signifikansi parameter model digunakan untuk mengetahui apakah parameter model telah layak masuk ke dalam model VAR dan VMA [3]. Model VAR dengan hipotesis:

$H_0: \Phi_p = 0$ (Parameter tidak signifikan), dan

$H_1: \Phi_p \neq 0$ (Parameter signifikan) dengan $p = 1, 2, \dots, i$, ditetapkan taraf signifikan (α), maka diperoleh daerah penolakan tolak H_0 jika nilai $|t_{hitung}|$ lebih besar dari $t_{\alpha/2, n-p-1}$ atau p -value kurang dari taraf signifikan (α). Perhitungan statistik uji pada uji signifikansi parameter model VAR disajikan pada persamaan sebagai berikut,

$$|t| = \frac{\hat{\Phi}_{p(ij)}}{SE(\hat{\Phi}_{p(ij)})} \quad (21)$$

dengan

$$SE(\beta) = |\Sigma_a|^{-\frac{1}{2}} (\mathbf{X}'\mathbf{X})^{-1} \quad (22)$$

dan $\beta = (\phi, \theta)$. Model VMA dengan hipotesis:

$H_0: \Theta_q = 0$ (Parameter tidak signifikan), dan

$H_1: \Theta_q \neq 0$ (Parameter signifikan) dengan $p = 1, 2, \dots, i$, ditetapkan taraf signifikan (α), maka diperoleh daerah penolakan tolak H_0 jika nilai $|t_{hitung}|$ lebih besar dari $t_{\alpha/2, n-p-1}$ atau p -value kurang dari taraf signifikan (α). Perhitungan statistik uji pada uji signifikansi parameter model VAR disajikan pada persamaan sebagai berikut,

$$|t| = \frac{\hat{\Theta}_{q(ij)}}{SE(\hat{\Theta}_{q(ij)})} \quad (23)$$

dengan nilai SE dijelaskan pada Persamaan (22).

3) *Diagnosis Residual*

Residual pada model yang terbentuk harus bersifat *white noise* dan berdistribusi normal multivariat.

a) *White Noise*

Pemeriksaan diagnosis *residual white noise* dilakukan berdasarkan pada analisis sisaan, yaitu

$$\hat{\mathbf{a}}_t = \mathbf{Z}_t - \hat{\Phi}_1 \mathbf{Z}_{t-1} - \dots - \hat{\Phi}_p \mathbf{Z}_{t-p} + \hat{\Theta}_1 \hat{\mathbf{a}}_{t-1} + \dots + \hat{\Theta}_q \hat{\mathbf{a}}_{t-q} \quad (24)$$

dengan \mathbf{Z}_t digunakan untuk menotasikan \mathbf{Z}_t jika $\mu = 0$ dan menotasikan $(\mathbf{Z}_t - \hat{\mu})$ dengan μ merupakan vektor rata-rata, $\hat{\Phi}_p$ adalah nilai taksiran dari parameter Φ_p dan $\hat{\Theta}_q$ adalah nilai taksiran dari parameter Θ_q . Suatu model *time series* dikatakan baik jika *residual* model bersifat *white noise*. Maka dari itu matriks korelasi dari \mathbf{a}_t harus tidak signifikan dan tidak memiliki pola [3]. Pemeriksaan *residual white noise* dapat dilakukan menggunakan uji ARCH-LM dan uji *Durbin Watson*.

Uji *Autoregressive Conditional Heteroscedastic Lagrange Multiplier* (ARCH-LM) yang digagas oleh Engle pada tahun 1982 digunakan untuk menguji adanya heteroskedastisitas pada *residual* model. Hipotesis yang digunakan adalah sebagai berikut [7]: H_0 adalah *residual* identik (varians *residual* bersifat homogen) dan H_1 adalah *residual* tidak identik (varians *residual* bersifat heterogen) ditetapkan taraf signifikan (α), maka diperoleh daerah penolakan tolak H_0 jika nilai F_{hitung} lebih besar dari $F_{\alpha; k; n-k+1}$ atau p -value kurang dari taraf signifikan (α). Perhitungan statistik uji pada uji ARCH-LM disajikan pada persamaan sebagai berikut [8].

$$F = \frac{R^2}{1-R^2} \frac{n-(k+1)}{k} \quad (25)$$

dengan R^2 merupakan koefisien determinasi regresi antara a_t^2 dengan $a_t^2(k)$, a_t adalah data *residual*, k adalah *lag*, dan n merupakan jumlah data.

Uji *Durbin Watson* digunakan untuk mendeteksi adanya autokorelasi (korelasi antar *residual*) atau mengetahui apakah *residual* telah independen dengan hipotesis [9]: H_0 merupakan *residual* independen (tidak terdapat autokorelasi), dan H_1 adalah *residual* tidak independen (terdapat autokorelasi) ditetapkan taraf signifikan (α), maka diperoleh daerah penolakan yang ditunjukkan pada Tabel 1. Di mana nilai d_L dan d_U diperoleh pada tabel *Durbin Watson*. Perhitungan statistik pada uji *Durbin Watson* disajikan pada persamaan sebagai berikut.

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (26)$$

b) *Berdistribusi Normal*

Pemeriksaan *residual* berdistribusi normal dapat dilakukan dengan menggunakan *Jarque-Bera normality test* sebagai berikut [7]. Dengan hipotesis: H_0 adalah *Residual* berdistribusi normal, dan H_1 adalah *Residual* tidak berdistribusi normal ditetapkan taraf signifikan (α), maka diperoleh daerah penolakan tolak H_0 jika nilai JB lebih besar dari $\chi^2_{\alpha, 2}$ atau p -value kurang dari taraf signifikan (α). Perhitungan statistik uji pada *Jarque-Bera normality test* adalah sebagai berikut [10].

$$b_1 = \frac{\sum_{t=1}^n \frac{r_t^3}{n}}{\left(\sum_{t=1}^n \frac{r_t^2}{n}\right)^{\frac{3}{2}}} \quad (27)$$

$$b_2 - 3 = \frac{\sum_{t=1}^n \frac{r_t^4}{n}}{\left(\sum_{t=1}^n \frac{r_t^2}{n}\right)^2} - 3 \quad (28)$$

$$JB = n \left(\frac{b_1^2}{6} + \frac{(b_2 - 3)^2}{24} \right) \quad (29)$$

dengan b_1 merupakan *skewness*, $b_2 - 3$ adalah kurtosis, n adalah jumlah observasi, dan r_t adalah *residual*.

Tabel 1.
Daerah Pengujian Durbin Watson

Kesimpulan	Daerah Pengujian
Terdapat autokorelasi positif	$d < d_L$
Tidak dapat disimpulkan	$d_L < d < d_U$ atau $4 - d_U < d < 4 - d_L$
Tidak terdapat autokorelasi	$d_U < d < 4 - d_U$
Terdapat autokorelasi negatif	$4 - d_L < d$

Tabel 2.
Struktur Data

Periode (t)	Jenis Pencemar Udara		
	$Z_{1,t}$	$Z_{2,t}$	$Z_{3,t}$
1	$Z_{1,1}$	$Z_{2,1}$	$Z_{3,1}$
2	$Z_{1,2}$	$Z_{2,2}$	$Z_{3,2}$
3	$Z_{1,3}$	$Z_{2,3}$	$Z_{3,3}$
⋮	⋮	⋮	⋮
184	$Z_{1,184}$	$Z_{2,184}$	$Z_{3,184}$

Tabel 3.
Statistika Deskriptif Kualitas Udara di Surabaya ($\mu\text{g}/\text{m}^3$)

Variabel	Rata-rata	Standar deviasi	Minimum	Maksimum
PM ₁₀	48,77	36,23	1	182
NO ₂	62,03	34,76	17	243
O ₃	21,054	11,278	3	68

4) Pemilihan Model Terbaik

RMSE merupakan metode yang digunakan untuk mengevaluasi ketepatan model *time series* dengan mempertimbangkan sisa perhitungan ramalan pada data *out-sample* [3]. RMSE total digunakan untuk mengevaluasi ketepatan model *time series* dengan mempertimbangkan sisa perhitungan ramalan pada data *out-sample* dari beberapa variabel secara bersamaan yang dapat dicari menggunakan rumus sebagai berikut,

$$RMSE_{total} = \sqrt{\frac{\sum_{t=1}^p \sum_{t=1}^n (z_t - \hat{z}_t)^2}{p}} \quad (30)$$

dengan n merupakan banyaknya data *out-sample*, p adalah jumlah kategori, Z_t merupakan data aktual periode ke- t , dan \hat{Z}_t adalah data peramalan periode ke- t .

B. Indeks Standar Pencemar Udara

Indeks Standar Pencemar Udara (ISPU) adalah angka yang tidak mempunyai satuan yang menggambarkan kondisi mutu udara ambien di lokasi tertentu, yang didasarkan kepada dampak terhadap kesehatan manusia, nilai estetika dan makhluk hidup lainnya. ISPU meliputi parameter partikulat (PM₁₀), partikulat (PM_{2,5}), karbon monoksida (CO), nitrogen dioksida (NO₂), sulfur dioksida (SO₂), ozon (O₃) dan hidrokarbon (HC). Tabel konversi nilai konsentrasi parameter ISPU disajikan Gambar 1 [11]. $\mu\text{g}/\text{m}^3$ adalah mikrogram per meter kubik, di mana 1 $\mu\text{g}/\text{m}^3$ sama dengan 0,001 mg/m^3 atau miligram per meter kubik. Kategori dan penjelasan dari ISPU dijelaskan seperti pada Gambar 2.

C. Partikulat

Partikulat atau *particulate matter* (PM) merupakan zat pencemar padat maupun cair yang terdispersi di udara. Partikulat itu dapat berupa debu, abu, jelaga, asap, uap, kabut, atau aerosol. Jenis-jenis partikulat dibedakan berdasarkan ukurannya. PM 10 menyatakan ukuran diameternya antara 2,5 mm hingga 10 mm.

ISPU	24 Jam PM10 ($\mu\text{g}/\text{m}^3$)	24 Jam PM2.5 ($\mu\text{g}/\text{m}^3$)	24 Jam SO ₂ ($\mu\text{g}/\text{m}^3$)	24 Jam CO ($\mu\text{g}/\text{m}^3$)	24 Jam O ₃ ($\mu\text{g}/\text{m}^3$)	24 Jam NO ₂ ($\mu\text{g}/\text{m}^3$)	24 Jam HC ($\mu\text{g}/\text{m}^3$)
0 - 50	50	15,5	52	4000	120	80	45
51 - 100	150	55,4	180	8000	235	200	100
101 - 200	350	150,4	400	15000	400	1130	215
201 - 300	420	250,4	800	30000	800	2260	432
>300	500	500	1200	45000	1000	3000	648

Keterangan:

- Data pengukuran selama 24 jam secara terus-menerus.
- Hasil perhitungan ISPU parameter partikulat (PM_{2.5}) disampaikan tiap jam selama 24 jam.
- Hasil perhitungan ISPU parameter partikulat (PM₁₀), sulfur dioksida (SO₂), karbon monoksida (CO), ozon (O₃), nitrogen dioksida (NO₂) dan hidrokarbon (HC), diambil nilai ISPU parameter tertinggi dan paling sedikit disampaikan setiap jam 09.00 dan jam 15.00.

Gambar 1. Tabel konversi nilai konsentrasi parameter ISPU.

Rentang	Kategori	Penjelasan
1-50	Baik	Tingkat mutu udara yang sangat baik, tidak memberikan efek negatif terhadap manusia, hewan dan tumbuhan.
51-100	Sedang	Tingkat mutu udara masih dapat diterima pada kesehatan manusia, hewan dan tumbuhan.
101-200	Tidak Sehat	Tingkat mutu udara yang bersifat merugikan pada manusia, hewan dan tumbuhan.
201-300	Sangat Tidak Sehat	Tingkat mutu udara yang dapat meningkatkan resiko kesehatan pada sejumlah segmen populasi yang terpapar.
301+	Berbahaya	Tingkat mutu udara yang dapat merugikan kesehatan serius pada populasi dan perlu penanganan cepat.

Gambar 2. Kategori rentang dan penjelasan nilai ISPU.

D. Nitrogen Dioksida

Kegiatan manusia yang menghasilkan *nitric oxide* (NO) dan *nitrogen dioxide* (NO₂) antara lain adalah penggunaan bahan bakar fosil. Gas NO di atmosfer bereaksi dengan oksigen dan terbentuklah gas NO₂ yang digunakan sebagai indeks pencemaran udara [12].

E. Ozon

Ozon (O₃) adalah gas yang reaktif terdiri dari tiga molekul oksigen. Ozon dapat merusak kesehatan makhluk hidup [13]. Ozon dapat mengurangi fungsi paru dan menyebabkan iritasi pernapasan [12].

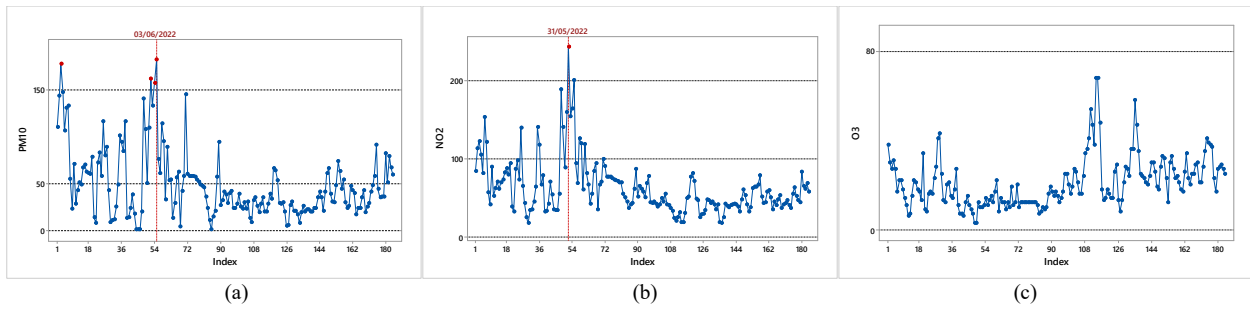
III. METODOLOGI PENELITIAN

A. Metode Pengumpulan Data

Data yang digunakan pada penelitian merupakan data sekunder yang diperoleh dari situs web <https://plumelabs.com/> berupa kualitas udara di Surabaya yang diidentifikasi menggunakan tiga parameter ISPU, yaitu PM₁₀, NO₂, dan O₃. Data yang digunakan merupakan data harian mulai tanggal 10 April hingga 10 Oktober 2022 yang terbagi menjadi 80 persen data *in-sample* dan 20 persen data *out-sample*.

B. Variabel Penelitian

Variabel yang digunakan pada penelitian adalah variabel $Z_{1,t}$ yang menunjukkan konsentrasi PM₁₀, variabel $Z_{2,t}$ yang menunjukkan konsentrasi NO₂, dan variabel $Z_{3,t}$ yang menunjukkan konsentrasi O₃ dengan skala rasio dan satuan $\mu\text{g}/\text{m}^3$. Struktur data yang digunakan pada penelitian disajikan pada Tabel 2.



Gambar 3. Time series plot: (a) PM₁₀, (b) NO₂, dan (c) O₃.

Variable/Lag	1	2	3	4	5	6	7	8	9	10	11	12
PM10	++	++	+++	+++
NO2	++	++	+++	+++	++	++
O3
Variable/Lag	13	14	15	16	17	18	19	20	21	22	23	24
PM10
NO2
O3
Variable/Lag	25	26	27	28	29	30	31	32	33	34	35	36
PM10
NO2
O3

Gambar 4. Plot MCCC.

Variable/Lag	1	2	3	4	5	6	7	8	9	10	11	12
PM10
NO2
O3
Variable/Lag	13	14	15	16	17	18	19	20	21	22	23	24
PM10
NO2
O3
Variable/Lag	25	26	27	28	29	30	31	32	33	34	35	36
PM10
NO2
O3

Gambar 5. Plot MCCC setelah differencing.

Variable/Lag	1	2	3	4	5	6	7	8	9	10	11	12
PM10
NO2
O3
Variable/Lag	13	14	15	16	17	18	19	20	21	22	23	24
PM10
NO2
O3
Variable/Lag	25	26	27	28	29	30	31	32	33	34	35	36
PM10
NO2
O3

Gambar 6. Plot MPCCF.

C. Metode Analisis Data

Metode yang digunakan pada penelitian adalah model VARIMA untuk meramalkan kualitas udara di Surabaya menggunakan konsentrasi PM₁₀, NO₂, dan O₃ dengan tujuan memodelkan kualitas udara di Surabaya dengan model terbaik sehingga dapat diketahui ramalan kualitas udara di Surabaya pada kemudian hari dengan spesifikasi model pada Tabel 2.

$$\begin{aligned}
 \begin{bmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \end{bmatrix} &= \begin{bmatrix} \phi_{111} & \phi_{112} & \phi_{113} \\ \phi_{121} & \phi_{122} & \phi_{123} \\ \phi_{131} & \phi_{132} & \phi_{133} \end{bmatrix} \begin{bmatrix} Z_{1t-1} \\ Z_{2t-1} \\ Z_{3t-1} \end{bmatrix} + \dots + \begin{bmatrix} \phi_{p11} & \phi_{p12} & \phi_{p13} \\ \phi_{p21} & \phi_{p22} & \phi_{p23} \\ \phi_{p31} & \phi_{p32} & \phi_{p33} \end{bmatrix} \\
 &= \begin{bmatrix} Z_{1t-1} \\ Z_{2t-1} \\ Z_{3t-1} \end{bmatrix} - \begin{bmatrix} \phi_{111} & \phi_{112} & \phi_{113} \\ \phi_{121} & \phi_{122} & \phi_{123} \\ \phi_{131} & \phi_{132} & \phi_{133} \end{bmatrix} \begin{bmatrix} a_{1t-1} \\ a_{2t-1} \\ a_{3t-1} \end{bmatrix} - \dots \\
 &= - \begin{bmatrix} \theta_{q11} & \theta_{q12} & \theta_{q13} \\ \theta_{q21} & \theta_{q22} & \theta_{q23} \\ \theta_{q31} & \theta_{q32} & \theta_{q33} \end{bmatrix} \begin{bmatrix} a_{1t-q} \\ a_{2t-q} \\ a_{3t-q} \end{bmatrix} \quad (31)
 \end{aligned}$$

Tabel 4. Nilai AIC

Minimum Information Criterion Based on AICC						
Lag	MA 0	MA 1	MA 2	MA 3	MA 4	MA 5
AR 0	16,2	16,1	15,9	16,0	16,1	16,2
AR 1	16,3	16,2	16,0	16,1	16,2	16,3
AR 2	16,0	16,0	16,0	16,1	16,2	16,3
AR 3	16,0	16,0	16,1	16,2	16,3	16,5
AR 4	16,2	16,1	16,2	16,3	16,5	16,6

Tabel 5. Uji ARCH-LM

Model	Variabel	F _{hitung}	F _{0,05;1;141}	p-value
VARIMA (3,1,0)	PM ₁₀	0,18	3,91	0,670
	NO ₂	0,09		0,760
	O ₃	3,3		0,071
VARIMA (0,1,[1,2,3,5])	PM ₁₀	0,22	0,641	0,641
	NO ₂	0,19		0,663
	O ₃	1,05		0,307
VARIMA (3,1,[1,2,3,5])	PM ₁₀	0,27	0,606	0,606
	NO ₂	0,74		0,393
	O ₃ *	6,34		0,013

D. Langkah Analisis

Langkah analisis yang digunakan dalam penelitian sebagai berikut: (1) Mendeskripsikan karakteristik data menggunakan statistika deskriptif dan time series plot. (2) Menganalisis model terbaik menggunakan model VARIMA dengan langkah-langkah sebagai berikut. (a) Memeriksa stasioneritas data in-sample dalam mean menggunakan time series plot dan plot MCCC. (b) Menduga orde model yang telah stasioner terhadap mean menggunakan plot MPCCF, plot MCCC, dan nilai AIC terkecil. (c) Mengestimasi parameter dan uji signifikansi parameter pada model VARIMA. (d) Memeriksa residual white noise dan residual berdistribusi normal. (e) Memilih model terbaik berdasarkan nilai RMSE total yang terkecil. (2) Meramalkan kualitas udara di Surabaya menggunakan dengan langkah-langkah sebagai berikut. (a) Menentukan periode ramalan. (b) Melakukan peramalan dengan model terbaik yang didapatkan. (3) Menginterpretasikan hasil analisis. Terakhir, (4) menarik kesimpulan dan saran.

IV. ANALISIS DAN PEMBAHASAN

A. Karakteristik Konsentrasi PM₁₀, NO₂, dan O₃

Karakteristik konsentrasi PM₁₀, NO₂, dan O₃ menggunakan statistika deskriptif disajikan pada Tabel 3. Tabel 3 menunjukkan bahwa konsentrasi PM₁₀ masuk ke dalam rentang ISPU 0–50 yang artinya rata-rata konsentrasi PM₁₀ masuk pada kategori baik dan pernah memasuki kategori tidak sehat. Konsentrasi NO₂ masuk ke dalam rentang ISPU 0–50 yang artinya rata-rata konsentrasi NO₂ masuk pada kategori baik dan pernah memasuki kategori tidak sehat. Konsentrasi O₃ masuk ke dalam rentang ISPU 0–50 yang artinya rata-rata konsentrasi O₃ masuk pada kategori baik dan berada pada

Tabel 6.
Uji Durbin Watson

Model	Variabel	d_{hitung}	Tabel Durbin Watson
VARIMA (3,1,0)	PM ₁₀	1,977	$d_U = 1,77$ $4 - d_U = 2,23$
	NO ₂	1,933	
	O ₃	1,982	
VARIMA (0,1,[1,2,3,5])	PM ₁₀	1,900	
	NO ₂	1,851	
	O ₃	2,025	
VARIMA (3,1,[1,2,3,5])	PM ₁₀	2,113	
	NO ₂	2,011	
	O ₃ *	1,750*	

Tabel 7.
Uji Durbin Watson

Model	Variabel	JB	$\chi^2_{0,05,2}$	p-value
VARIMA (3,1,0)	PM ₁₀	94,45	5,991	<,0001
	NO ₂	215,64		<,0001
	O ₃	93,67		<,0001
VARIMA (0,1,[1,2,3,5])	PM ₁₀	51,09		<,0001
	NO ₂	149,79		<,0001
	O ₃	120		<,0001
VARIMA (3,1,[1,2,3,5])	PM ₁₀	36,75		<,0001
	NO ₂	41,37		<,0001
	O ₃	50,13		<,0001

kategori sehat. Karakteristik konsentrasi PM₁₀, NO₂, dan O₃ menggunakan *time series plot* disajikan pada Gambar 3.

Gambar 3 menunjukkan bahwa konsentrasi PM₁₀, NO₂, dan O₃ bergerak secara fluktuatif, di mana konsentrasi PM₁₀ berada di kategori baik hingga sedang, namun beberapa kali masuk pada kategori tidak sehat pada tanggal 12 April, 31 Mei, 2 Juni dan tertinggi berada pada tanggal 3 Juni 2022. Konsentrasi NO₂ berada pada kategori baik hingga sedang, namun pada tanggal 31 Mei 2022 berada pada kategori tidak sehat dan konsentrasi O₃ berada pada kategori baik.

B. Pemodelan Data Menggunakan VARIMA

1) Pemeriksaan Stasioneritas Data

Hasil *time series plot* pada Gambar 3 menunjukkan bahwa variabel PM₁₀, NO₂, dan O₃ belum stasioner terhadap *mean* karena *plot* pada *time series plot* cenderung meningkat dan menurun secara acak. Hasil *plot* MCCF disajikan pada Gambar 4. Gambar 4 menunjukkan bahwa pada hasil *plot* MCCF, hampir semua *lag* pada variabel PM₁₀, NO₂, dan O₃ memiliki simbol positif atau negatif, artinya variabel PM₁₀, NO₂, dan O₃ belum stasioner terhadap *mean* sehingga perlu dilakukan *differencing* data dan dilakukan pemeriksaan ulang. Hasil pemeriksaan stasioneritas dalam *mean* variabel PM₁₀, NO₂, dan O₃ setelah dilakukan *differencing* 1 disajikan pada Gambar 5.

Gambar 5 menunjukkan bahwa pada hasil *plot* MCCF, setelah dilakukan *differencing* 1, simbol positif atau negatif hanya muncul pada beberapa *lag* tertentu, artinya variabel PM₁₀, NO₂, dan O₃ yang telah di *differencing* 1 telah stasioner terhadap *mean*.

2) Identifikasi Model VARIMA

Hasil *plot* MPCCF disajikan pada Gambar 6, pada Gambar 6 menunjukkan pada *plot* MPCCF, *lag* yang keluar dari batas spesifikasi *standard error* adalah *lag* 1, 2, dan 3. *Plot* MCCF yang disajikan pada Gambar 5 diatas menunjukkan bahwa *lag* yang keluar dari batas spesifikasi *standard error* adalah *lag* 1, 2, 3, dan 5. Selain dari *plot* MPCCF dan MCCF diatas, orde model juga dapat diketahui berdasarkan nilai AIC terkecil yang disajikan pada Tabel 4.

Tabel 4 menunjukkan pada nilai AIC terkecil berada pada *lag* AR 0 dan MA 2. Hasil *plot* MPCCF, *plot* MCCF, dan nilai AIC terkecil menunjukkan bahwa model dugaan pada data kualitas udara adalah VARIMA (3,1,0), VARIMA (0,1,[1,2,3,5]), dan VARIMA (3,1,[1,2,3,5]).

3) Estimasi Parameter dan Pengujian Signifikansi Model

Estimasi model dilakukan dengan mengeluarkan parameter yang tidak signifikan menggunakan proses *restrict*. *Restrict* dilakukan hingga semua parameter pada model signifikan. Hasil analisis menunjukkan bahwa ketiga parameter dari ketiga model dugaan VARIMA kualitas udara di Surabaya telah signifikan sehingga dapat dilanjutkan dengan pembentukan matriks koefisien model VARIMA sebagai berikut. Model 1, VARIMA (3,1,0)

$$\Phi_1 = \begin{bmatrix} -0,26 & 0 & 0 \\ 0 & -0,28 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\Phi_2 = \begin{bmatrix} -0,28 & 0 & 0 \\ 0 & -0,39 & 0 \\ 0 & 0 & -0,25 \end{bmatrix}$$

$$\Phi_3 = \begin{bmatrix} -0,27 & 0,27 & 0,43 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Model 2, VARIMA (0,1,[1,2,3,5])

$$\Theta_1 = \begin{bmatrix} 0,42 & 0 & 0 \\ 0 & 0,41 & 0,34 \\ 0 & 0 & 0 \end{bmatrix} \quad \Theta_2 = \begin{bmatrix} 0,64 & -0,44 & 0 \\ 0,27 & 0 & 0 \\ 0 & 0 & 0,31 \end{bmatrix}$$

$$\Theta_3 = \begin{bmatrix} 0,26 & -0,19 & -0,32 \\ 0 & 0 & 0 \\ 0 & 0 & 0,19 \end{bmatrix} \quad \Theta_4 = \begin{bmatrix} 0 & 0,23 & 0 \\ 0 & 0,25 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

Model 3, VARIMA (3,1,[1,2,3,5])

$$\Phi_1 = \begin{bmatrix} 0 & -1,45 & 0 \\ 0 & -0,59 & -0,22 \\ 0 & 0 & 0 \end{bmatrix} \quad \Phi_2 = \begin{bmatrix} -0,49 & 0 & 0 \\ 0,20 & 0 & 3,51 \\ 0 & 0 & -0,19 \end{bmatrix}$$

$$\Phi_3 = \begin{bmatrix} 1,77 & -0,62 & 7,52 \\ 0,58 & 0 & 0 \\ 0 & 0 & -0,13 \end{bmatrix} \quad \Theta_1 = \begin{bmatrix} 0 & -1,05 & 0 \\ -0,33 & 0 & 0 \\ 0 & 0 & 0,14 \end{bmatrix}$$

$$\Theta_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0,80 & 0 & 3,49 \\ 0 & 0 & 0 \end{bmatrix} \quad \Theta_3 = \begin{bmatrix} 3,47 & -1,27 & 7,08 \\ 0,88 & 0 & 0 \\ -0,15 & 0,16 & 0 \end{bmatrix}$$

$$\Theta_5 = \begin{bmatrix} 0 & 0,73 & -0,36 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

4) Pemeriksaan Residual White Noise dan Berdistribusi Normal

a) Pemeriksaan Residual White Noise

Pemeriksaan homogenitas *residual* menggunakan uji ARCH-LM dengan Hipotesis: H₀ adalah *Residual* identik (varians *residual* bersifat homogen) dan H₁ adalah *Residual* tidak identik (varians *residual* bersifat heterogen) ditetapkan taraf signifikan (α) sebesar 0,05, maka diperoleh daerah penolakan tolak H₀ jika nilai F_{hitung} lebih besar dari F_{0,05;1;141} atau p-value kurang dari 0,05. Hasil perhitungan statistik uji disajikan sebagai pada Tabel 5.

Tabel 5 menunjukkan bahwa *residual* variabel O₃ pada model VARIMA (3,1,[1,2,3,5]) bersifat heterogen atau *residual* tidak identik dan variabel lainnya memiliki *residual* bersifat homogen atau *residual* identik. Selanjutnya, dilakukan pemeriksaan independensi *residual* menggunakan uji Durbin Watson dengan hipotesis: H₀ adalah *residual* independen (tidak terdapat autokorelasi) dan H₁: *Residual* tidak

Tabel 8.
Pemilihan Model Peramalan Terbaik

Model	Variabel	RMSE	Total RMSE
VARIMA (3,1,0)	PM ₁₀	20,516	15,595
	NO ₂	15,324	
	O ₃	8,595	
VARIMA (0,1,[1,2,3,5])	PM ₁₀	20,022	18,566
	NO ₂	15,216	
	O ₃	8,640	
VARIMA (3,1,[1,2,3,5])	PM ₁₀	30,305	21,169
	NO ₂	19,389	
	O ₃	7,078	

Tabel 9.
Pemilihan Model Peramalan Terbaik

Tanggal	PM ₁₀	NO ₂	O ₃
11/10/2022	56,407	57,22	25,5548
12/10/2022	63,858	61,509	26,1097
13/10/2022	61,252	60,257	25,9557
14/10/2022	60,73	58,902	25,8018
15/10/2022	60,897	59,765	25,8445
16/10/2022	61,341	60,066	25,8872
17/10/2022	60,913	59,643	25,8754
18/10/2022	61,09	59,639	25,8635
19/10/2022	61,138	59,807	25,8668
20/10/2022	61,077	59,763	25,8701
21/10/2022	61,026	59,709	25,8692
22/10/2022	61,089	59,741	25,8683
23/10/2022	61,093	59,754	25,8685
24/10/2022	61,073	59,737	25,8688
25/10/2022	61,068	59,737	25,8687

independen (terdapat autokorelasi) ditetapkan taraf signifikan (α) sebesar 0,05, maka diperoleh daerah penolakan tolak H_0 jika nilai d_{hitung} lebih kecil dari d_L , lebih besar dari $4 - d_L$, dan berada pada rentang $d_L < d < d_U$ atau $4 - d_U < d < d_L$. Hasil perhitungan statistik uji disajikan pada Tabel 6.

Tabel 6 menunjukkan bahwa terdapat autokorelasi pada variabel O₃ pada model VARIMA (3,1,[1,2,3,5]) atau *residual* tidak independen dan tidak terdapat autokorelasi pada variabel lainnya atau *residual* independen.

b) Pemeriksaan Residual Berdistribusi Normal

Pemeriksaan *residual* berdistribusi normal dilakukan dengan menggunakan uji *Jarque Bera Normality* dengan hipotesis: H_0 adalah *Residual* berdistribusi normal, dan H_1 adalah *Residual* tidak berdistribusi normal, ditetapkan taraf signifikan (α) sebesar 0,05, maka diperoleh daerah penolakan tolak H_0 jika nilai JB lebih besar dari $\chi^2_{0,05;2}$ atau *p-value* kurang dari 0,05. Hasil perhitungan statistik uji disajikan di Tabel 7.

Tabel 7 menunjukkan bahwa setiap variabel pada masing-masing model dugaan memiliki *residual* yang tidak berdistribusi normal. Diduga *residual* tidak berdistribusi normal karena adanya *residual* yang *outlier*.

5) Pemilihan Model Terbaik

Model peramalan terbaik dipilih berdasarkan nilai RMSE total terkecil dari data *outsample* yang disajikan pada Tabel 8. Tabel 8 menunjukkan bahwa model yang terpilih untuk meramalkan kualitas udara di Surabaya adalah model VARIMA (3,1,0) dengan interpretasi sebagai berikut. Model Konsentrasi PM₁₀

$$\hat{Z}_1(t) = Z_1(t-1) - 0,26Z_1(t-1) + 0,26Z_1(t-2) + 0,28Z_1(t-3) - 0,27Z_1(t-3) + 0,27Z_2(t-3) + 0,43Z_3(t-3) + 0,27Z_1(t-4) - 0,27Z_2(t-4) - 0,43Z_3(t-4) + a_1(t)$$

Konsentrasi PM₁₀ pada waktu ke- t dipengaruhi oleh konsentrasi PM₁₀ pada 1, 2, 3, dan 4 hari sebelumnya, dan konsentrasi NO₂ dan O₃ pada 3 dan 4 hari sebelumnya.

a) Model Konsentrasi NO₂

$$\hat{Z}_2(t) = Z_2(t-1) - 0,28Z_2(t-1) + 0,28Z_2(t-2) - 0,34Z_2(t-2) + 0,34Z_2(t-3) + a_2(t)$$

Konsentrasi NO₂ pada waktu ke- t dipengaruhi oleh konsentrasi NO₂ pada 1, 2, dan 3 hari sebelumnya.

b) Model Konsentrasi O₃

$$\hat{Z}_3(t) = Z_3(t-1) - 0,25Z_3(t-2) + 0,25Z_3(t-3) + a_3(t)$$

Konsentrasi O₃ pada waktu ke- t dipengaruhi oleh konsentrasi O₃ pada 1, 2, dan 3 hari sebelumnya.

C. Peramalan pada Periode Selanjutnya

Hasil peramalan menggunakan model VARIMA (3,1,0) pada 15 periode selanjutnya disajikan pada Tabel 9. Tabel 9 menunjukkan bahwa pada ramalan kualitas udara di Surabaya memiliki nilai tertinggi pada tanggal 12 Oktober 2022 dan berdasarkan pada kategori ISPU, kondisi kualitas udara di Surabaya pada tanggal 11 Oktober hingga 25 Oktober 2022 memiliki konsentrasi PM₁₀ pada kategori sedang dan konsentrasi NO₂ dan O₃ pada kategori baik, di mana tingkat kualitas udara masih dapat diterima pada kesehatan manusia, hewan, dan tumbuhan.

V. KESIMPULAN/RINGKASAN

Model peramalan terbaik pada kualitas udara di Surabaya adalah model VARIMA (3,1,0). Hasil peramalan kualitas udara di Surabaya pada tanggal 11 Oktober hingga 25 Oktober 2022 berdasarkan pada kategori ISPU berada pada kategori baik hingga sedang, di mana tingkat kualitas udara masih dapat diterima pada kesehatan manusia, hewan, dan tumbuhan. Hasil peramalan secara lengkap ditunjukkan di Tabel 9.

DAFTAR PUSTAKA

- [1] V. A. Putri, "Kajian Hubungan Dampak Paparan NO₂, O₃, dan PM₁₀ Terhadap Kejadian Penyakit ISPA Pneumonia dan Non-Pneumonia di Kebonsari, Surabaya dan Sekitarnya," Departemen Teknik Lingkungan, Institut Teknologi Sepuluh Nopember, Surabaya, 2021.
- [2] A. P. Prasetyo dan A. D. Syaifei, "Kajian konsentrasi NO₂, O₃, dan PM₁₀ di udara terhadap kejadian penyakit ISPA pneumonia dan non-pneumonia di Wonorejo, Surabaya dan sekitarnya," *J. Purifikasi*, vol. 21, no. 2, hal. 80-92, 2022, doi: <https://doi.org/10.12962/purifikasi.v21i2.461>.
- [3] W. W. S. Wei, *Time Series Analysis: Univariate and Multivariate Methods*, 2nd ed. New York, USA: Pearson Addison Wesley, 2006. ISBN: 0-321-32216-9.
- [4] G. C. Tiao dan G. E. P. Box, "Modeling multiple time series with applications," *J. Am. Stat. Assoc.*, vol. 76, no. 376, hal. 802-816, 1981, doi: 10.1080/01621459.1981.10477728.
- [5] J. F. Heyse dan W. W. S. Wei, "Modelling the advertising-sales relationship through use of multiple time series techniques," *J. Forecast.*, vol. 4, no. 2, hal. 165-181, Jan 1985, doi: <https://doi.org/10.1002/for.3980040206>.
- [6] H. Akaike, *Selected Papers of Hirotugu Akaike*. New York, USA: Springer, 1998. ISBN: 978-1-4612-1694-0.
- [7] R. S. Tsay, *Analysis of Financial Time Series*, 2nd ed. New Jersey: John Wiley & Sons, 2005. ISBN: 978-0-471-69074-0.
- [8] J. Welc dan P. J. R. Esquerdo, *Applied Regression Analysis for Business*. Cham, Switzerland: Springer, 2018. ISBN: 978-3-319-89041-8.
- [9] D. N. Gujarati, *Basic Econometrics*, 5th ed. New York, USA: Mcgraw-Hill, 2003. ISBN: 978-0-07-337577-9.
- [10] J. C. Brocklebank dan D. A. Dickey, *SAS for Forecasting Time Series*, 2nd ed. North California, USA: John Wiley & Sons, Inc, 2003. ISBN: 0-471-39566-8.

- [11] Menteri Lingkungan Hidup dan Kehutanan, "Peraturan Menteri Lingkungan Hidup dan Kehutanan Republik Indonesia Nomor P.14/MENLHK/SETJEN/KUM.1/7/2020 Tentang Indeks Standar Pencemar Udara," Kementerian Lingkungan Hidup dan Kehutanan, Jakarta, 2020.
- [12] H. J. Mukono, *Pencemaran Udara Dalam Ruang Berorientasi Kesehatan Masyarakat*. Surabaya: Airlangga University Press (AUP), 2014. ISBN: 978-602-7924-59-8.
- [13] S. Hermawan, *Ilmu Lingkungan Bermetode Service Learning*. Sleman: Kanisius, 2020. ISBN: 978-979-21-6507-4.