

# Peramalan Konsentrasi *Particulate Matter* 2.5 (PM<sub>2.5</sub>) menggunakan Model *Vector Autoregressive* dengan Metode *Maximum Likelihood Estimation*

Atika Tresna Arianto<sup>1, a)</sup>, Kankan Parmikanti<sup>2, b)</sup>, Bambang Suhandi<sup>3, c)</sup>, dan Budi Nurani Ruchjana<sup>4, d)</sup>

<sup>1,2,4</sup>*Departemen Matematika, Fakultas MIPA, Universitas Padjadjaran*

<sup>3</sup>*Balai Pengelola Observatorium Nasional Kupang*

<sup>a)</sup>*email: atikatresna241@gmail.com*

<sup>b)</sup>*email: parmikanti@unpad.ac.id*

<sup>c)</sup>*email: bambang.suhandi@lapan.go.id*

<sup>d)</sup>*email: budi.nurani@unpad.ac.id*

## Abstrak

*Particulate Matter* 2.5 yang selanjutnya disingkat PM<sub>2.5</sub> merupakan partikel udara yang memiliki ukuran  $\leq 2.5\mu m$ . Paparan PM<sub>2.5</sub> dapat mempengaruhi fungsi paru-paru dan memperburuk kondisi kesehatan seperti asma, bronkitis hingga kanker paru-paru. PM<sub>2.5</sub> berasal dari berbagai sumber seperti hasil pembakaran bahan bakar kendaraan bermotor, hingga kebakaran hutan. Untuk meramalkan Konsentrasi PM<sub>2.5</sub> dapat digunakan model *time series* univariat, salah satunya model *Autoregressive* yang selanjutnya disingkat AR. Data Konsentrasi PM<sub>2.5</sub> memiliki pola stasioner, sehingga dapat dibangun model AR. Konsentrasi PM<sub>2.5</sub> di suatu wilayah dipengaruhi oleh Konsentrasi PM<sub>2.5</sub> wilayah di sekitarnya pada waktu-waktu sebelumnya, sehingga dapat dibentuk model *Vector Autoregressive* yang selanjutnya disingkat VAR. Metode yang digunakan untuk menaksir model VAR dalam penelitian ini adalah metode *Maximum Likelihood Estimation* yang selanjutnya disingkat MLE dengan bantuan *software R* dan *Microsoft Excel*. Model VAR diterapkan untuk peramalan jangka pendek Konsentrasi PM<sub>2.5</sub> di Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka. Hasil penelitian menunjukkan bahwa konsentrasi PM<sub>2.5</sub> di ketiga kabupaten memiliki korelasi yang cukup tinggi dan satu sama lain saling mempengaruhi, sehingga penggunaan model VAR dapat memberikan rekomendasi untuk peramalan konsentrasi PM<sub>2.5</sub> pada waktu mendatang dengan memperhitungkan pengaruh konsentrasi PM<sub>2.5</sub> dari lokasi-lokasi terdekat di sekitar lokasi tertentu.

*Kata kunci: MLE, Konsentrasi PM<sub>2.5</sub>, Software R, Model VAR*

## Abstract

Particulate Matter 2.5, hereinafter referred to as PM<sub>2.5</sub>, is an air particle which has a size of  $\leq 2.5\mu m$ . Exposure to PM<sub>2.5</sub> can affect lung function and worsen health conditions such as asthma, bronchitis and lung cancer. PM<sub>2.5</sub> comes from various sources such as the combustion of motor vehicle fuel, to forest fires. To predict PM<sub>2.5</sub> Concentrations, a univariate time series model can be used, one of which is the Autoregressive model, hereinafter abbreviated as AR. PM<sub>2.5</sub> Concentration data has a stationary pattern, so that the AR model can be formed. PM<sub>2.5</sub> Concentration in an area is influenced by PM<sub>2.5</sub>

Concentration in the surrounding area at previous times, so that it can be formed as a Vector Autoregressive model which is then abbreviated as VAR. The method used to estimate the VAR model in this study is the Maximum Likelihood Estimation method, which is hereinafter abbreviated as MLE with the help of R and Microsoft Excel software. The VAR model is applied for the short-term forecasting of PM<sub>2.5</sub> Concentrations in Kab. Cirebon, Kab. Kuningan, and Kab. Majalengka. The results showed that PM<sub>2.5</sub> concentrations in the three districts had a high correlation and influenced each other, so the use of the VAR model could provide recommendations for forecasting PM<sub>2.5</sub> concentrations in the future by calculating the effect of PM<sub>2.5</sub> concentrations from locations closest around a certain location.

**Keywords:** MLE, Concentration PM<sub>2.5</sub>, R Software, VAR Model

## Pendahuluan

Udara merupakan komponen penting bagi kehidupan sehingga menjaga kualitas udara untuk tetap baik dan layak adalah hal yang sangat penting. Pada tahun 2016, 90% populasi penduduk di dunia tidak menghirup udara bersih, dan lebih dari setengah populasi di perkotaan terpapar polusi udara luar ruangan dengan level 2.5 kali lebih tinggi dari standar aman yang telah ditetapkan oleh WHO. Diperkirakan pada tahun 2016 polusi udara luar ruangan di daerah perkotaan dan pedesaan menyebabkan 4.2 juta kematian di seluruh dunia[1].

Material penyebab polusi udara salah satunya adalah partikulat. Partikulat adalah semua material, kecuali air murni, yang berada dalam fasa cair atau padat dan terdapat di atmosfer pada kondisi normal dengan ukuran mikroskopis atau semi mikroskopis, namun lebih besar dibandingkan dimensi molekuler ( $\pm 2 \text{ \AA}$ )[2]. Analisis beban global dari penyakit akibat polusi udara luar ruangan mengenai PM<sub>2.5</sub> mengestimasi bahwa PM<sub>2.5</sub> telah menyebabkan 3% kematian akibat penyakit kardiopulmoner, 5% kematian dari kanker trakea, bronkus, dan paru-paru, serta 1% kematian pada anak-anak akibat infeksi pernapasan di daerah perkotaan di seluruh dunia[3]. Fenomena terkait polusi udara terjadi secara global, tidak hanya di luar negeri, di Indonesia kualitas udara yang buruk juga masih menjadi masalah utama. Untuk menyelesaikan tujuan SDGs nomor 3, yaitu memastikan kehidupan yang sehat dan mendukung kesejahteraan bagi semua manusia, yang termasuk didalamnya mengurangi angka kematian dan penyakit yang disebabkan oleh bahan kimia berbahaya dan juga polusi serta kontaminasi udara, air, dan tanah, maka diperlukan analisis terkait polusi udara yang salah satunya yaitu PM<sub>2.5</sub>.

Model *Vector Autoregressive* (VAR) merupakan model deret waktu yang dapat digunakan untuk menganalisis data deret waktu stasioner. Model VAR merupakan pengembangan dari model *Autoregressive* (AR) dimana pada model VAR variabel endogen yang digunakan lebih dari satu. Beberapa penelitian yang telah dilakukan terkait model VAR diantaranya adalah *Vector Autoregression (VAR) Model for Rainfall Forecast and Isohyet Mapping in Semarang - Central Java - Indonesia*[4]. Pada penelitiannya Nugroho, dkk menggunakan model VAR dengan metode *Ordinary Least Square* dengan 3 variabel yaitu curah hujan, kelembaban, dan temperatur udara. Selain itu, *Vector Autoregressive Model Approach for Forecasting Outflow Cash in Central Java* [2]. Pada penelitiannya, Hoyyi, dkk meramalkan *Outflow Cash* di Jawa Tengah menggunakan model VAR dengan estimasi parameter model menggunakan metode *Ordinary Least Square* dengan 4 variabel endogen.

Terkait permasalahan kualitas udara di Indonesia, pada penelitian ini penulis menganalisis data *time series* Konsentrasi PM<sub>2.5</sub> di Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka. Pada penelitian ini data Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka dimodelkan menggunakan

Model VAR dengan metode *Maximum Likelihood Estimation* (MLE). Model yang didapatkan selanjutnya digunakan untuk meramalkan Konsentrasi PM<sub>2.5</sub> di Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka yang pada akhirnya dapat digunakan sebagai pertimbangan dalam menentukan langkah mitigasi untuk mencegah dampak negatif yang akan ditimbulkan dari Konsentrasi PM<sub>2.5</sub>.

## Metode

### 1. Time Series

*Time series* merupakan serangkaian pengamatan terhadap suatu variabel yang diamati dari waktu ke waktu secara berurutan berdasarkan waktu kejadiannya dengan interval waktu yang tetap. *Time series* merupakan sebuah realisasi atau fungsi sampel dari proses stokastik tertentu[5]. Untuk melakukan peramalan, diperlukan analisis terhadap data time series agar di dapat pola data pada masa lampau. Untuk melakukan analisis terhadap data time series, maka perlu diperhatikan bentuk pola data [6].

### 2. Kestasioneran

Sekumpulan data dikatakan stasioner apabila fluktuasi dari data berada di sekitar rata-rata. Suatu proses stokastik  $\{Z_t, t \in T\}$  dikatakan stasioner jika rata-rata  $E(Z_t) = \mu$  konstan, kovarians  $\gamma_k = cov(Z_t, Z_{t+k}) = E(Z_t - \mu_t)(Z_{t+k} - \mu_{t+k})$  hanya bergantung pada nilai  $k$ , dan varians  $\sigma_t^2 = E(Z_t - \mu)^2 = Var(Z_t)$  artinya nilai variansi dalam periode waktu adalah konstan[7].

### 3. Model Autoregressive (AR)

*Autoregressive* (AR) merupakan model *time series* univariat yang menyatakan bahwa nilai pengamatan suatu variabel dependen pada waktu sekarang dipengaruhi oleh nilai pengamatan variabel itu sendiri pada waktu-waktu sebelumnya. Jika pengamatan pada waktu ke- $t$  dari set data deret waktu tergantung pada *lag* waktu pengamatan ordo ke- $p$ , maka disebut dengan AR( $p$ ), dengan kata lain pengamatan sekarang dipengaruhi oleh " $p$ " waktu pengamatan sebelumnya[8]. Model AR( $p$ ) dinyatakan sebagai persamaan berikut[7]:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + \varepsilon_t \quad (1)$$

dengan,

$Z_t$  : variabel acak pada waktu  $t$

$\phi_p$  : parameter model AR pada lag ke-  $p$

$\varepsilon_t$  : galat pada waktu  $t$ ,  $\varepsilon_t \stackrel{iid}{\sim} N(0, \sigma^2)$

### 4. Model Vector Autoregressive (VAR)

Model *Vector Autoregressive* (VAR) merupakan model *time series* pengembangan dari model *Autoregressive* (AR). Model VAR menjelaskan keterkaitan antar pengamatan pada peubah itu sendiri pada waktu sebelumnya dan juga keterkaitannya dengan pengamatan pada peubah lain pada waktu sebelumnya. Jika nilai pengamatan di suatu lokasi/ wilayah pada waktu ke- $t$  dipengaruhi oleh nilai pengamatan di-" $p$ " lag sebelumnya beserta pengamatan di lokasi lain, maka model ini berbentuk VAR( $p$ )[9]. Secara umum model VAR( $p$ ) didefinisikan sebagai berikut[8]:

$$\vec{Z}_t = \Phi_1 \vec{Z}_{t-1} + \dots + \Phi_p \vec{Z}_{t-p} + \vec{\varepsilon}_t \quad (2)$$

dengan,

$\vec{Z}_t$  : vektor ( $N \times 1$ ) dari  $N$  lokasi pada period ke  $t$

$\vec{Z}_{t-p}$  : vektor ( $N \times 1$ ) dari  $N$  lokasi pada period ke  $t - p$

$\Phi_p$  : matriks ( $N \times N$ ) yang berisikan parameter model pada lag ke- $p$

$\varepsilon_t$  : vektor ( $N \times 1$ ) yang berisikan nilai galat pada periode  $t$ ,  $\varepsilon_t \sim N(0, \sigma^2 I)$

### 5. Maximum Likelihood Estimation (MLE) pada Model VAR

Galat pada model VAR(1) diasumsikan berdistribusi normal, sehingga penaksiran parameter model VAR(1) dapat diestimasi menggunakan metode MLE dengan cara memaksimalkan fungsi *ln likelihood*[7]. Misal, untuk variabel pengamatan  $N = 3$  dan waktu  $T = 4$ , maka persamaannya menjadi:

$$\begin{bmatrix} \begin{pmatrix} Z_{1,2} \\ Z_{2,2} \\ Z_{3,2} \end{pmatrix} \\ \begin{pmatrix} Z_{1,3} \\ Z_{2,3} \\ Z_{3,3} \end{pmatrix} \\ \begin{pmatrix} Z_{1,4} \\ Z_{2,4} \\ Z_{3,4} \end{pmatrix} \end{bmatrix} = \begin{bmatrix} \begin{pmatrix} Z_{1,1} & Z_{2,1} & Z_{3,1} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Z_{1,1} & Z_{2,1} & Z_{3,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & Z_{1,1} & Z_{2,1} & Z_{3,1} \end{pmatrix} \\ \begin{pmatrix} Z_{1,2} & Z_{2,2} & Z_{3,2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Z_{1,2} & Z_{2,2} & Z_{3,2} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & Z_{1,2} & Z_{2,2} & Z_{3,2} \end{pmatrix} \\ \begin{pmatrix} Z_{1,3} & Z_{2,3} & Z_{3,3} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & Z_{1,3} & Z_{2,3} & Z_{3,3} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & Z_{1,3} & Z_{2,3} & Z_{3,3} \end{pmatrix} \end{bmatrix} \begin{bmatrix} \phi_{11} \\ \phi_{12} \\ \phi_{13} \\ \phi_{21} \\ \phi_{22} \\ \phi_{23} \\ \phi_{31} \\ \phi_{32} \\ \phi_{33} \end{bmatrix} + \begin{bmatrix} \begin{pmatrix} \varepsilon_{1,2} \\ \varepsilon_{2,2} \\ \varepsilon_{3,2} \end{pmatrix} \\ \begin{pmatrix} \varepsilon_{1,3} \\ \varepsilon_{2,3} \\ \varepsilon_{3,3} \end{pmatrix} \\ \begin{pmatrix} \varepsilon_{1,4} \\ \varepsilon_{2,4} \\ \varepsilon_{3,4} \end{pmatrix} \end{bmatrix} \quad (3)$$

$y \qquad X \qquad \beta \qquad \varepsilon$

persamaan (3) disederhanakan menjadi model linear, yaitu:

$$y = X\beta + \varepsilon \quad (4)$$

Persamaan (4) ditaksir menggunakan MLE dengan fungsi kepadatan peluang yang didefinisikan sebagai

$$f(\varepsilon) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \exp \left( -\frac{\varepsilon^2}{2\sigma^2} \right) \quad (5)$$

sehingga didapat fungsi *likelihood*nya yaitu

$$L(\beta, \sigma^2) = \frac{1}{(\sqrt{2\pi\sigma^2})^n} \exp \exp \left( -\frac{1}{2\sigma^2} (y - X\beta)^T (y - X\beta) \right) \quad (6)$$

Selanjutnya, fungsi *likelihood* di *ln* kan untuk mendapatkan fungsi *ln likelihood* sebagai berikut:

$$\ln \ln (L(\beta, \sigma^2)) = (2\pi) - \frac{n}{2} \ln \ln (\sigma^2) - \frac{1}{2\sigma^2} (y - X\beta)^T (y - X\beta) \quad (7)$$

untuk mendapatkan penaksiran *maximum likelihood*, fungsi *ln likelihood* diturunkan terhadap  $\beta$  sebagai berikut:

$$\frac{\partial \ln \ln (L(\beta, \sigma^2))}{\partial \beta} = \frac{1}{\sigma^2} (X^T y - X^T X \beta) \quad (8)$$

Setelah didapatkan turunan pertama fungsi *ln likelihood*, maka selanjutnya turunan pertama fungsi *ln likelihood* disamadengankan 0, untuk memperoleh penduga parameter secara *likelihood*, yaitu:

$$\begin{aligned} \frac{1}{\sigma^2} (X^T y - X^T X \beta) &= 0 \\ X^T X \beta &= X^T y \end{aligned}$$

$$\begin{aligned}(X^T X)^{-1} X^T X \beta &= (X^T X)^{-1} X^T y \\ \hat{\beta} &= (X^T X)^{-1} X^T y\end{aligned}\quad (9)$$

untuk menjamin fungsi *In likelihood* maksimum, maka turunan keduanya harus bernilai  $< 0$ , persamaan (8) diturunkan terhadap  $\beta$  sehingga diperoleh:

$$\frac{\partial^2 \ln L(\beta, \sigma^2)}{\partial^2 \beta} = -\frac{1}{\sigma^2} X^T X \quad (10)$$

karena turunan kedua fungsi *In likelihood* terhadap  $\beta < 0$ , maka fungsi *In likelihood* maksimum.

Taksiran parameter model VAR(1) untuk variabel pengamatan  $N = 3$  dan waktu  $T = 4$ , maka persamaannya menjadi:

$$\hat{\Phi} = \left( Z_{t-1}^T \left( (3 \times 3) \times (3 \times 3) \right) Z_{t-1} \left( (3 \times 3) \times (3 \times 3) \right) \right)^{-1} Z_{t-1}^T \left( (3 \times 3) \times (3 \times 3) \right) \vec{Z}_t \left( (3 \times 3) \times 1 \right) \quad (11)$$

Taksiran MLE untuk model VAR(1) untuk  $N$  variabel pengamatan dan waktu  $T$  secara umum, yaitu:

$$\hat{\Phi}_{(N \times N)} = \left( Z_{t-1}^T \left( (N \times N) \times (N \times (T-1)) \right) Z_{t-1} \left( (N \times (T-1)) \times (N \times N) \right) \right)^{-1} Z_{t-1}^T \left( (N \times N) \times (N \times (T-1)) \right) \vec{Z}_t \left( (N \times (T-1)) \times 1 \right) \quad (12)$$

## 6. Q-Q Plot

*Q-Q (Quantile-Quantile)* Plot merupakan uji asumsi normalitas dengan pendekatan analisis grafik. Pada pendekatan *Q-Q Plot*, jika titik-titik (*dots*) menyebar jauh dari garis diagonal, maka indikasi normalitas tidak terpenuhi. Sebaliknya, jika titik-titik menyebar sangat dekat pada garis diagonal, maka asumsi normalitas terpenuhi[9].

## 7. Mean Absolute Percentage Error (MAPE)

*Mean Absolute Percentage Error (MAPE)* merupakan persentase dari rata-rata mutlak galat pada tiap periode dibagi dengan nilai aktual pada periode tersebut. MAPE mengindikasikan seberapa besar persentase *error* untuk mengetahui ketepatan dugaan model. MAPE dihitung dengan rumus sebagai berikut[7]:

$$MAPE = \left( \sum_{i=1}^N \sum_{t=2}^T \left| \frac{Z_{i,t} - \hat{Z}_{i,t}}{Z_{i,t}} \right| \times \frac{1}{N(T-1)} \right) \times 100\% \quad (13)$$

## 8. Particulate Matter 2.5 (PM<sub>2.5</sub>)

*Particulate Matter 2.5 (PM<sub>2.5</sub>)* merupakan partikel udara yang memiliki ukuran  $\leq 2.5 \mu\text{m}$ . PM<sub>2.5</sub> dapat terhirup oleh saluran pernapasan hingga mencapai paru-paru. Paparan PM<sub>2.5</sub> dapat mempengaruhi fungsi paru-paru dan memperburuk kondisi kesehatan seperti asma, bronkitis hingga kanker paru-paru. PM<sub>2.5</sub> berasal dari berbagai sumber seperti hasil pembakaran bahan bakar kendaraan bermotor, sehingga kebakaran hutan[10].

## 9. Software R

*R* merupakan suatu sistem analisis data statistik yang lengkap sebagai hasil dari kolaborasi penelitian berbagai ahli statistik di seluruh dunia. *Software R* dapat diperoleh secara gratis di *CRAN-archive* yaitu *The Comprehensive R Archive Network* di alamat <http://cran.r-project.org>. Fungsi dan kemampuan dari *software R* sebagian besar dapat diperoleh melalui *Add-on packages/library*. Suatu *library* merupakan kumpulan perintah atau fungsi yang dapat digunakan untuk melakukan analisis tertentu[11].

## Hasil dan Diskusi

### Data Penelitian

Data yang digunakan pada penelitian ini adalah data sekunder berupa data Konsentrasi  $PM_{2.5}$  yang diperoleh dari situs web Berkeley Earth (Nirlaba independen yang berfokus pada ilmu iklim dan analisa strategis). Data yang digunakan adalah data dari tiga wilayah yaitu Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka berupa data per jam dari tanggal 12 September 2019 hingga 17 September 2019. Banyaknya data yang digunakan dalam penelitian ini adalah 124 data Konsentrasi  $PM_{2.5}$ . Pada penelitian ini, data Konsentrasi  $PM_{2.5}$  dibagi menjadi data *in-sample*, yaitu data  $PM_{2.5}$  selama 120 jam (12 September – 16 September 2019), dan data *out-sample* yaitu data Konsentrasi  $PM_{2.5}$  selama 4 jam (pukul 00.00-00.03 tanggal 17 September 2019). Variabel penelitian yang digunakan pada penelitian ini dijelaskan sebagai berikut:

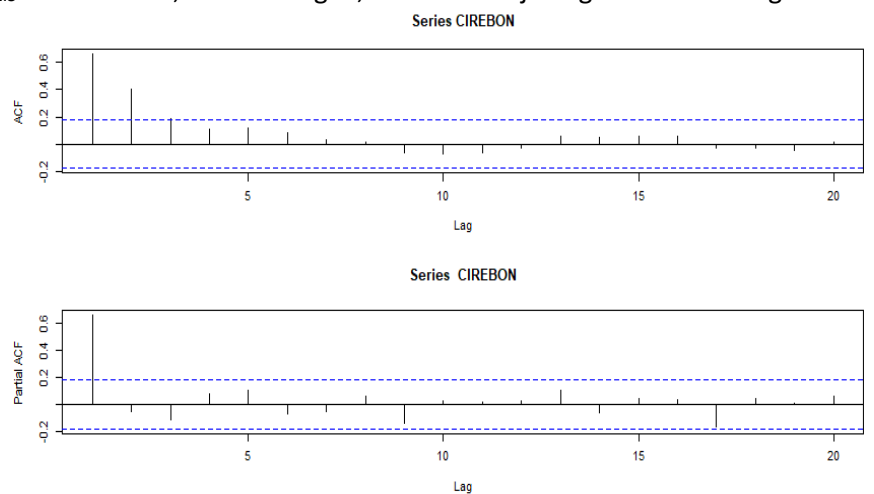
$Z_{1,t}$ : Konsentrasi  $PM_{2.5}$  di Kab. Cirebon

$Z_{2,t}$ : Konsentrasi  $PM_{2.5}$  di Kab. Kuningan

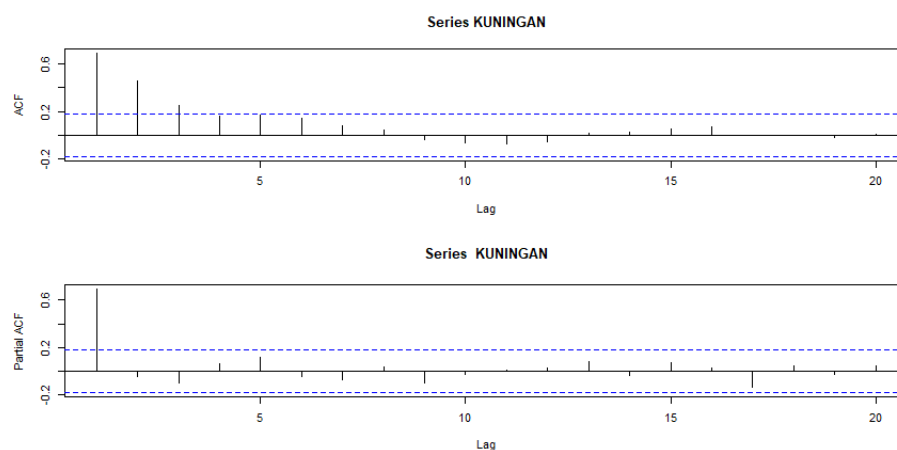
$Z_{3,t}$ : Konsentrasi  $PM_{2.5}$  di Kab. Majalengka

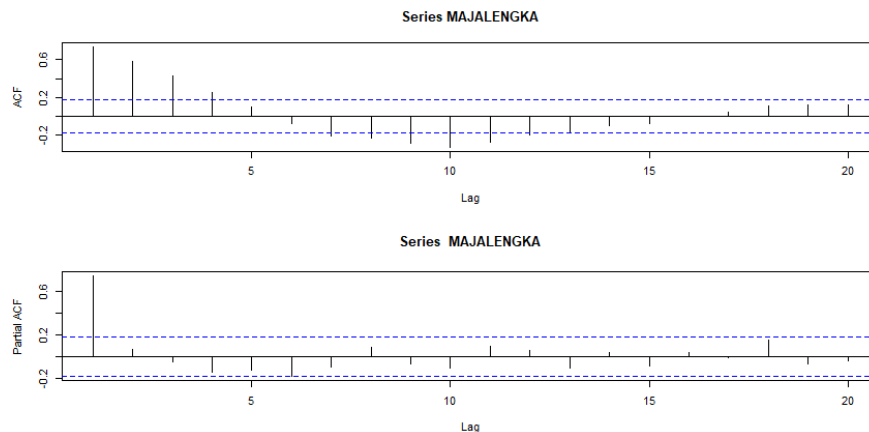
### 1. Identifikasi Model *Autoregressive* (AR)

Model *Autoregressive* (AR) dapat diidentifikasi melalui plot ACF dan PACF. Plot ACF dan PACF data konsentrasi  $PM_{2.5}$  Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka adalah sebagai berikut:



**Gambar 1.** Plot ACF dan PACF Konsentrasi  $PM_{2.5}$  Kab. Cirebon



**Gambar 2.** Plot ACF dan PACF Konsentrasi PM<sub>2.5</sub> Kab. Kuningan**Gambar 3.** Plot ACF dan PACF Konsentrasi PM<sub>2.5</sub> Kab. Majalengka

Dilihat dari plot ACF dan PACF, plot ACF ketiga wilayah menurun secara eksponensial, sedangkan pada plot PACF terlihat *cuts-off* pada *lag* pertama, yang artinya data Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka dapat dimodelkan menggunakan model AR(1).

## 2. Penaksiran Model AR(1)

Pada penelitian ini, model AR(1) ditaksir menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan bantuan *software R*. Tabel 1. menunjukkan hasil taksiran parameter model AR(1) dengan metode MLE.

**Tabel 1.** Nilai Taksiran Parameter Model AR(1) Data Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka

Wilayah	$\hat{\phi}_1$
Kab. Cirebon	0.9931
Kab. Kuningan	0.9914
Kab. Majalengka	0.9866

Berdasarkan Tabel 1 diperoleh hasil penaksiran parameter model AR(1) dengan metode MLE pada data Konsentrasi PM<sub>2.5</sub> di tiga wilayah pengamatan, untuk wilayah pengamatan Kab. Cirebon diperoleh nilai taksiran  $\hat{\phi}_1 = 0.9931$ , untuk wilayah pengamatan Kab. Kuningan diperoleh nilai taksiran  $\hat{\phi}_1 = 0.9914$ , dan untuk wilayah pengamatan Kab. Majalengka diperoleh nilai taksiran  $\hat{\phi}_1 = 0.9866$ .

Model AR(1) untuk data PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka dapat ditulis ke dalam persamaan sebagai berikut:

$$\hat{Z}_{1,t} = 0.9931Z_{1,t-1} \quad (14)$$

$$\hat{Z}_{2,t} = 0.9914Z_{2,t-1} \quad (15)$$

$$\hat{Z}_{3,t} = 0.9866Z_{3,t-1} \quad (16)$$

Dari persamaan (14), dapat diartikan bahwa Konsentrasi PM<sub>2.5</sub> di Kab. Cirebon setiap jam nya dipengaruhi oleh 99,31% konsentrasi PM<sub>2.5</sub> satu jam sebelumnya. Dari persamaan (15), dapat diartikan bahwa Konsentrasi PM<sub>2.5</sub> di Kab. Kuningan setiap jam nya dipengaruhi oleh 99,14%

konsentrasi PM<sub>2.5</sub> satu jam sebelumnya. Dari persamaan (16), dapat diartikan bahwa Konsentrasi PM<sub>2.5</sub> di Kab. Majalengka setiap jam nya dipengaruhi oleh 98,66% konsentrasi PM<sub>2.5</sub> satu jam sebelumnya.

### 3. Penaksiran Model VAR(1)

Pada penelitian ini, model VAR(1) ditaksir menggunakan metode *Maximum Likelihood Estimation* (MLE) dengan bantuan *software R*. Tabel 2 menunjukkan hasil taksiran parameter model VAR(1) dengan metode MLE.

**Tabel 2.** Nilai Taksiran Parameter Model VAR(1) Data Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka

Taksiran Parameter	Nilai
$\hat{\phi}_{11}$	0.4693
$\hat{\phi}_{12}$	0.1807
$\hat{\phi}_{13}$	0.1112
$\hat{\phi}_{21}$	-0.2183
$\hat{\phi}_{22}$	0.8938
$\hat{\phi}_{23}$	0.1030
$\hat{\phi}_{31}$	0.4515
$\hat{\phi}_{32}$	-0.3836
$\hat{\phi}_{33}$	0.7060

Nilai taksiran parameter model VAR(1) di Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka dapat dinyatakan dalam bentuk matriks sebagai berikut:

$$\begin{bmatrix} Z_{1,t} & Z_{2,t} & Z_{3,t} \end{bmatrix} = \begin{bmatrix} 0.4693 & 0.1807 & 0.1112 & -0.2183 & 0.8938 & 0.1030 & 0.4515 \\ -0.3836 & 0.7060 \end{bmatrix} \begin{bmatrix} Z_{1,t-1} & Z_{2,t-1} & Z_{3,t-1} \end{bmatrix}$$

atau dapat ditulis kedalam persamaan sebagai berikut:

$$\hat{Z}_{1,t} = 0.4693Z_{1,t-1} + 0.1807Z_{2,t-1} + 0.1112Z_{3,t-1} \quad (17)$$

$$\hat{Z}_{2,t} = -0.2183Z_{1,t-1} + 0.8938Z_{2,t-1} + 0.1030Z_{3,t-1} \quad (18)$$

$$\hat{Z}_{3,t} = 0.4515Z_{1,t-1} - 0.3836Z_{2,t-1} + 0.7060Z_{3,t-1} \quad (19)$$

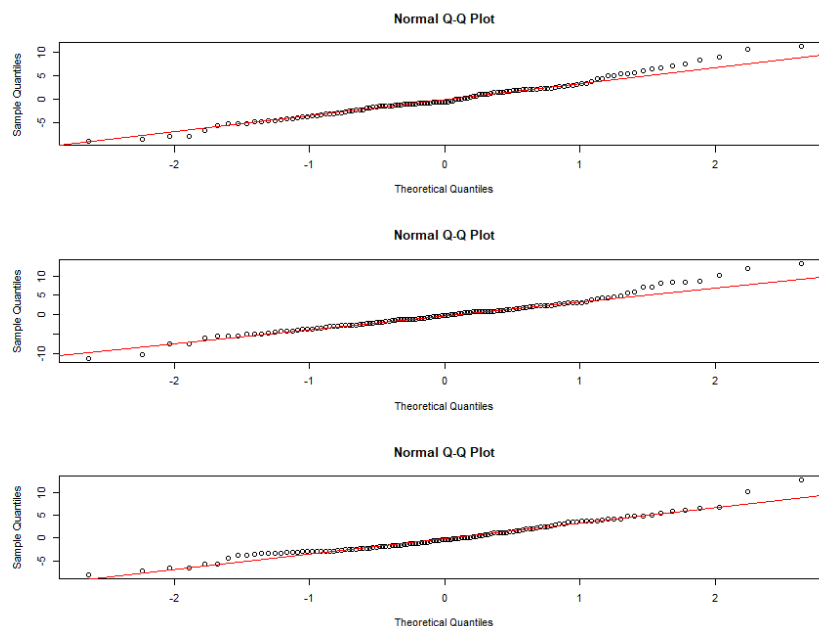
Dari persamaan (17) diketahui bahwa Konsentrasi PM<sub>2.5</sub> di Kab. Cirebon setiap jam nya dipengaruhi oleh penjumlahan dari 46,93% Konsentrasi PM<sub>2.5</sub> satu jam sebelumnya, 18,07% Konsentrasi PM<sub>2.5</sub> Kab. Kuningan satu jam sebelumnya, dan 11,12% Konsentrasi PM<sub>2.5</sub> Kab. Majalengka satu jam sebelumnya. Dengan cara yang sama, dari persamaan (8) dan (9) dapat dijelaskan bahwa variasi Konsentrasi PM<sub>2.5</sub> waktu sekarang dipengaruhi oleh Konsentrasi PM<sub>2.5</sub> satu waktu sebelumnya di lokasi tersebut dan pengaruh Konsentrasi PM<sub>2.5</sub> satu waktu sebelumnya dari dua lokasi lain di sekitar lokasi tersebut.

### 4. Cek Diagnostik Model VAR(1)

Dengan bantuan *Software R*, pengecekan distribusi normal galat dari model VAR(1) data Konsentrasi PM<sub>2.5</sub> di ketiga wilayah pengamatan dapat dilakukan melalui QQ-plot dan QQ-line seperti disajikan pada Gambar 4. Tampak bahwa QQ-plot berada di sekitar garis lurus, sehingga dapat



diasumsikan bahwa galat model VAR(1) mendekati distribusi normal. Dengan demikian model VAR(1) yang dibangun dapat digunakan untuk peramalan pada waktu mendatang dengan memperhitungkan pengaruh konsentrasi PM<sub>2.5</sub> di lokasi tersebut dan lokasi-lokasi lain di sekitarnya.



**Gambar 4.** QQ-Plot Galat Model VAR(1) Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka

**5. Penerapan Model VAR(1)**

Model VAR(1) yang telah diperoleh pada persamaan (7), (8), dan (9) diterapkan untuk meramalkan Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka. Peramalan dilakukan pada data *in-sample* Konsentrasi PM<sub>2.5</sub> ketiga wilayah pengamatan. Dibantu dengan *software Microsoft Excel 2013*, perbandingan nilai ramalan dan data aktual Konsentrasi PM<sub>2.5</sub> di ketiga wilayah pengamatan dapat dilihat pada Tabel 3.

**Tabel 3.** Perbandingan Nilai Ramalan dan Data Aktual *Out-sample* Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka

Tanggal	Pukul	Kab. Cirebon		Kab. Kuningan		Kab. Majalengka	
		$Z_{1,t}$	$\hat{Z}_{1,t}$	$Z_{2,t}$	$\hat{Z}_{2,t}$	$Z_{3,t}$	$\hat{Z}_{3,t}$
17-Sep-19	00.00	50.70	33.02	52.40	33.63	50.00	17.72
17-Sep-19	00.01	53.20	38.83	54.60	40.92	52.40	38.09
17-Sep-19	00.02	31.50	40.66	31.70	42.59	28.40	40.07
17-Sep-19	00.03	35.20	23.67	34.50	24.38	31.20	22.11

**6. Perhitungan MAPE Model VAR (1)**

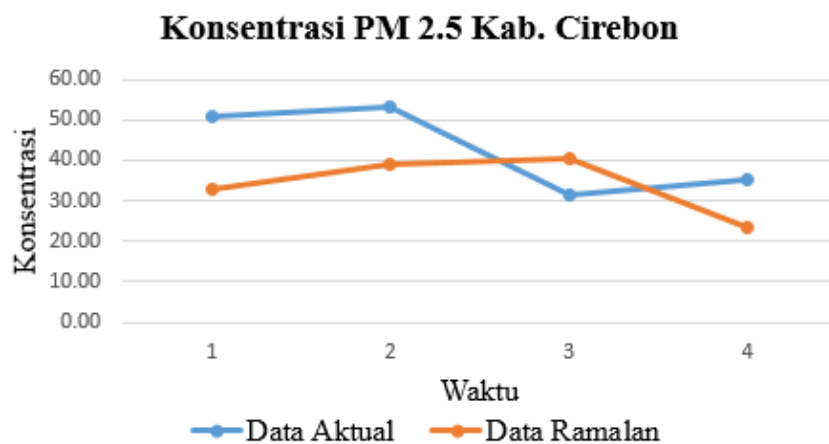
Hasil perhitungan MAPE Konsentrasi PM<sub>2.5</sub> *in-sample* dan *out-sample* ketiga wilayah pengamatan dapat dilihat pada Tabel 4.

**Tabel 4.** Hasil Perhitungan MAPE Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka

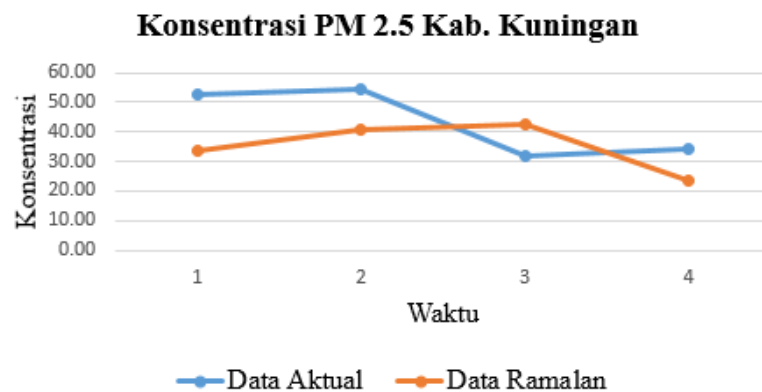
Wilayah	MAPE (%)	
	<i>In-sample</i>	<i>Out-sample</i>

Kab. Cirebon	9.37	13.75
Kab. Kuningan	8.75	13.84
Kab. Majalengka	5.72	18.01

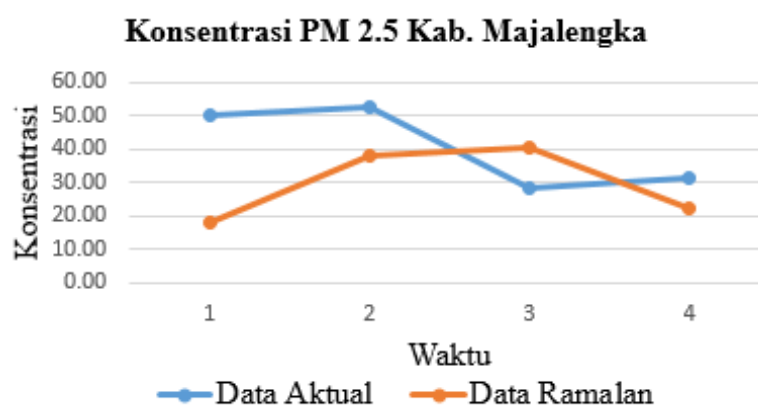
Dapat dilihat pada Tabel 4. nilai MAPE *in-sample* memiliki nilai 10-20%, sehingga menunjukkan model sesuai untuk peramalan waktu mendatang. Demikian halnya dengan MAPE *out-sample* Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka pada jangka pendek termasuk dalam kategori baik untuk peramalan, sehingga model VAR(1) dapat digunakan sebagai alternatif untuk melakukan peramalan Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka. Hal ini didukung dengan kemiripan pola data dengan ramalan pada jangka pendek seperti disajikan pada Gambar 5, 6 dan 7.



**Gambar 5.** Plot Data Aktual dan Data Ramalan Konsentrasi PM<sub>2.5</sub> Kab. Cirebon



**Gambar 6.** Plot Data Aktual dan Data Ramalan Konsentrasi PM<sub>2.5</sub> Kab. Kuningan



**Gambar 7.** Plot Data Aktual dan Data Ramalan Konsentrasi PM<sub>2.5</sub> Kab. Majalengka

Berdasarkan Gambar 5, 6, dan 7 dapat dilihat bahwa peramalan Konsentrasi PM<sub>2.5</sub> Kab. Cirebon, Kab. Kuningan, dan Kab. Majalengka memiliki pola kemiripan dengan data aktual pada jam ke-1 sampai jam ke-2, namun pada jam ke-3 dan ke-4 memiliki pola yang berbeda dengan data aktual, sehingga model VAR(1) untuk data Konsentrasi PM<sub>2.5</sub> pada data penelitian ini hanya sesuai untuk peramalan jangka waktu pendek. Hal ini disebabkan karena terdapat pencilon pada data awal saat penaksiran, namun model VAR(1) ini diharapkan dapat dijadikan salah satu alternatif untuk meramalkan Konsentrasi PM<sub>2.5</sub> maupun data-data *time series* lainnya.

### Kesimpulan

Berdasarkan hasil dan pembahasan, maka diperoleh kesimpulan bahwa penerapan model VAR(1) dengan metode *Maximum Likelihood Estimation* pada data Konsentrasi PM<sub>2.5</sub> di Kab. Cirebon, Kab. Kuningan, Kab. Majalengka menghasilkan model ramalan Konsentrasi PM<sub>2.5</sub> yang sesuai hanya dalam jangka waktu pendek di satu wilayah pengamatan yang dipengaruhi bukan hanya oleh Konsentrasi PM<sub>2.5</sub> di wilayah itu sendiri, namun juga dipengaruhi oleh Konsentrasi PM<sub>2.5</sub> di dua wilayah pengamatan lainnya. Hasil ramalan untuk jam pertama dan kedua memiliki kemiripan dengan data aktual, namun untuk jam ketiga dan keempat memiliki pola yang berbeda. Walaupun demikian, untuk jangka pendek MAPE dari nilai ramalan ketiga wilayah pengamatan berada pada kategori baik, sehingga model VAR(1) dapat digunakan sebagai alternatif untuk melakukan peramalan Konsentrasi PM<sub>2.5</sub> selanjutnya.

### Ucapan Terima Kasih

Penulis mengucapkan terima kasih kepada Rektor Universitas Padjadjaran yang telah memberikan dukungan dana penelitian melalui *Academic Leadership Grant* tahun 2021 dengan nomor kontrak 1959/UN6.3.1/PT00/ 2021 untuk diseminasi paper oleh dosen dan mahasiswa.

### Referensi

- [1] World Health Organization, *Monitoring Health for The SDGs, Sustainable Development Goals*. Geneva: World Health Organization, 2018.
- [2] Hoyyi, A., Tarno, Maruddani, D. A., and Rahmawati, R., *Vector Autoregressive Model Approach for Forecasting Outflow Cash in Central Java, ISNPINSA-7*, 2018.
- [3] Cohen, A. J., *The Global Burden of Disease due to Outdoor Air Pollution, Journal of Toxicology and Environmental Health*, 2005.

- [4] Nugroho, A., Subanar, Hartati, S., and Mustofa, K., Vector Autoregression (VAR) Model for Rainfall Forecast and Isohyet Mapping in Semarang-Central Java-Indonesia. *International Journal of Advanced Computer Science and Applications*, 2014.
- [5] Seinfeld, J. H., and Pandis, S. N., *Atmospheric Chemistry and Physics, From Air Pollution to Climate Change*. Wiley-Interscience, 1997.
- [6] Hakimi, F. D.D.,dkk. "Peramalan Kecepatan Angin yang Direkamoleh Sistem AWS dengan Analisis Fuzzy Time Series", *Jurnal Publikasi Ilmiah Matematika Vol. 2 No. 2*, pp. 24-32, 2017.
- [7] Wei, W. W., *Time Series Analysis: Univariate and Multivariate Methods. Second Edition*. United States of America: Pearson Education, Inc., 2006.
- [8] Wijaya, F. B., *Pendekatan Space Time Autoregressive (STAR) dengan Generalized Space Time Autoregressive (GSTAR) melalui metode Autoregressive (AR) dan Vector Autoregressive(VAR)*. Tesis tidak dipublikasikan. Bogor: IPB, 2015.
- [9] Gio, P. U., dan Irawan, D. E. , *Belajar Statistika dengan R*. Medan: USU Press, 2016.
- [10] Department of Health, New York State. (n.d.). *Fine Particles (PM 2.5) Questions and Answers*. Retrieved from The Official Web Site of New York State: [https://www.health.ny.gov/environmental/indoors/air/pmq\\_a.htm](https://www.health.ny.gov/environmental/indoors/air/pmq_a.htm)
- [11] Suhartono, *Analisis Data Statistik dengan R*. Surabaya: Lab. Statistik Komputasi, ITS, 2008.