

## PERAMALAN JUMLAH PENUMPANG KERETA API MENGGUNAKAN METODE ARIMA, INTERVENSI DAN ARFIMA (Studi Kasus : Penumpang Kereta Api Kelas Lokal Ekonomi DAOP IV Semarang)

Helmi Panjaitan<sup>1</sup>, Alan Prahutama<sup>2</sup>, Sudarno<sup>3</sup>

<sup>1,2,3</sup> Departemen Statistika, Fakultas Sains dan Matematika, Universitas Diponegoro  
*e-mail* : alan.prahutama@gmail.com

### ABSTRACT

Autoregressive Integrated Moving Average (ARIMA) is stationary time series model after differentiation. Differentiation value of ARIMA method is an integer so it is only able to model in the short term. The best model using ARIMA method is ARIMA([13]; 1; 0) with an MSE value of 1,870844. The Intervention method is a model for time series data which in practice has extreme fluctuations both up and down. In the data plot the number of train passengers was found to be extreme fluctuation. The data used was from January 2009 to June 2017 where fluctuation up significantly in January 2016 (T=85 to T=102) so the intervention model that was suspected was a step function. The best model uses the Intervention step function is ARIMA ([13]; 1; 1) (b=0; s=18; r=0) with MSE of 1124. Autoregressive Fractionally Integrated Moving Average (ARFIMA) method is a development of the ARIMA method. The advantage of the ARFIMA method is the non-integer differentiation value so that it can overcome long memory effect that can not be solve with the ARIMA method. ARFIMA model is capable of modeling high changes in the long term (long term persistence) and explain long-term and short-term correlation structures at the same time. The number of local economy class train passengers in DAOP IV Semarang contains long memory effects, so the ARFIMA method is used to obtain the best model. The best model obtained is the ARMA(0; [1,13]) model with the differential value is 0,367546, then the model can be written into ARFIMA (0; d; [1,13]) with an MSE value of 0,00964. Based on the analysis of the three methods, the best method of analyzing the number of local economy class train passengers in DAOP IV Semarang is the ARFIMA method with the model is ARFIMA (0; 0,367546; [1,13]).

**Keywords:** Train Passengers, ARIMA, Intervention, ARFIMA, Forecasting

### 1. PENDAHULUAN

Transportasi merupakan bidang kegiatan yang sangat penting dalam kehidupan masyarakat Indonesia. Pentingnya transportasi bagi masyarakat Indonesia disebabkan oleh beberapa faktor antara lain, keadaan geografis Indonesia yang terdiri dari ribuan pulau kecil dan besar, perairan yang terdiri dari sebagian besar laut, sungai dan danau yang memungkinkan pengangkutan dilakukan melalui darat, perairan, dan udara guna menjangkau seluruh wilayah Indonesia (Abdulkadir, 1998).

Berdasarkan KAI (2014), Badan Usaha Milik Negara (BUMN) yang menyelenggarakan jasa angkutan perkeretaapian adalah PT Kereta Api Indonesia (PT KAI). PT KAI berkewajiban menyelenggarakan *Public Service Obligation (PSO)* yaitu salah satunya dengan memberikan subsidi kepada beberapa kereta api. PT KAI memprioritaskan KRL dan KA ekonomi jarak dekat untuk pemberian subsidi. Hal tersebut membuat perlunya untuk memprediksikan jumlah penumpang kereta api kelas lokal ekonomi guna meningkatkan pelayanan maupun kenyamanan konsumen. Data jumlah penumpang sendiri merupakan salah satu data runtun waktu yang dapat diprediksi nilainya untuk beberapa tahap ke depan.

Dalam praktik, seringkali ditemui data runtun waktu yang berfluktuasi ekstrem. Fluktuasi ekstrem tersebut dapat disebabkan oleh berbagai faktor, baik eksternal maupun internal yang mempengaruhi pola data. Salah satu metode dalam runtun waktu yang dapat digunakan untuk mengatasi masalah tersebut adalah analisis intervensi. Metode ini dapat digunakan untuk memodelkan dan meramalkan data yang mengandung intervensi baik dari faktor internal maupun eksternal. Di dalam model intervensi terdapat dua fungsi yaitu

fungsi *step* dan *pulse*. Fungsi *step* merupakan suatu bentuk intervensi yang terjadi dalam kurun waktu yang panjang, sedangkan fungsi *pulse* adalah suatu bentuk intervensi yang terjadi hanya dalam waktu sesaat (Wei, 2006). Menurut Cahyandari dan Erviana (2015), adakalanya suatu data runtun waktu menunjukkan pola memori jangka panjang (*long memory*), ini terlihat dari nilai-nilai autokorelasi pada plot ACF yang turun secara lambat untuk jarak waktu (*lag*) yang semakin meningkat dan hasil perhitungan dari statistik *Hurst* (*H*) yang terletak dalam interval  $0,5 < H < 1$ . Identifikasi ini mengindikasikan bahwa nilai dari koefisien pembeda bernilai pecahan, sehingga model yang paling cocok adalah model *Autoregressive Fractionally Integrated Moving Average (ARFIMA)*.

## 2. TINJAUAN PUSTAKA

### 2.1 Kereta Api

Berdasarkan Republik Indonesia (1998), Peraturan Pemerintah Nomor 69 Tahun 1998 tentang prasarana dan sarana kereta api, menyebutkan bahwa moda transportasi kereta api memiliki karakteristik dan keunggulan khusus. Beberapa keunggulan dari kereta api adalah kemampuannya dalam mengangkut baik penumpang maupun barang secara massal, hemat energi, hemat dalam penggunaan ruang, memiliki faktor keamanan yang tinggi, tingkat pencemaran yang rendah, serta lebih efisien untuk angkutan jarak jauh. Stasiun besar yang berada di bawah kendali DAOP IV Semarang adalah Stasiun Semarang Tawang, Stasiun Semarang Poncol, Stasiun Tegal, Stasiun Pekalongan, Stasiun Cepu, Stasiun Ngrombo, dan Stasiun Ambarawa (stasiun kereta wisata). Sedangkan stasiun berkelas menengah diantaranya Stasiun Brumbung, Stasiun Kedungjati, Stasiun Gambringan, Stasiun Weleri, Stasiun Comal, Stasiun Batang Baru dan Stasiun Pemalang. Gudang kereta api berada di kompleks Stasiun Semarang Poncol, sedangkan depo lokomotif berada di sebelah timur Stasiun Semarang Poncol. Berdasarkan KAI (2018), kereta api kelas lokal ekonomi dibawah kendali DAOP IV Semarang terdiri dari KA Kalijaga (Stasiun Semarang Poncol – Stasiun Solo Balapan), KA Kaligung (Stasiun Semarang Poncol – Stasiun Brebes), KA Blorajaya (Stasiun Semarang Poncol – Stasiun Cepu) dan KA Kedung Sepur (Stasiun Semarang Poncol – Stasiun Ngrombo).

### 2.2 Analisis Time Series

Menurut Makridakis dkk. (1995), analisis *time series* adalah pengamatan sekarang ( $Z_t$ ) tergantung pada satu atau beberapa pengamatan sebelumnya ( $Z_{t-k}$ ). Dengan kata lain, model *time series* dibuat karena secara statistik ada korelasi (dependensi) antar deret pengamatan. Tujuan analisis *time series* antara lain memahami dan menjelaskan mekanisme tertentu, meramalkan suatu nilai di masa depan dan mengoptimalkan sistem kendali.

### 2.3 Stasioneritas

Menurut Makridakis dkk. (1995), dalam pembentukan model *time series*, data di asumsikan stasioner terhadap mean dan varian, yang konstan untuk semua waktu. Metode Box-Cox digunakan sebagai salah satu alternatif untuk menstasionerkan data dalam varian, sedangkan untuk menstasionerkan data dalam *mean* salah satu caranya adalah dengan proses *differencing*.

### 2.4 Fungsi Autokorelasi (ACF) dan Fungsi Autokorelasi Parsial (PACF)

Matriks autokorelasi suatu runtun waktu stasioner yang panjangnya  $k$  dalam Soejoeti (1987) adalah sebagai berikut:

$$\rho_{kk} = \begin{bmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-2} \\ \rho_2 & \rho_1 & 1 & \dots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & 1 \end{bmatrix}$$

Secara umum nilai fungsi autokorelasi parsial (PACF) pada lag ke-k adalah:

$$\phi_{kk} = \frac{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_1 \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & \rho_k \end{vmatrix}}{\begin{vmatrix} 1 & \rho_1 & \rho_2 & \dots & \rho_{k-2} & \rho_{k-1} \\ \rho_1 & 1 & \rho_1 & \dots & \rho_{k-3} & \rho_{k-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \dots & \rho_1 & 1 \end{vmatrix}}$$

## 2.5. Model Time Series ARIMA

Model time series ARIMA dapat dibedakan menjadi:

- Model AR dapat dituliskan sebagai berikut:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

- Model MA dapat dituliskan sebagai berikut:

$$Z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

- Model ARMA dapat dituliskan sebagai berikut:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

## 2.6 Model Time Series Tidak Stasioner

### 2.6.1. Model Autoregressive Integrated Moving Average atau ARIMA

Bentuk umum model ini adalah:  $\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t$

#### 2.6.1.1 Estimasi dan Uji Signifikansi Parameter Model ARIMA

Pengujian signifikansi parameter dapat dilakukan dengan langkah berikut:

- AR (*Autoregressive*)

Hipotesis:

$H_0: \phi_i = 0$  (parameter AR tidak berpengaruh terhadap model)

$H_1: \phi_i \neq 0$  (parameter AR berpengaruh terhadap model)

- MA (*Moving Average*)

Hipotesis:

$H_0: \theta_i = 0$  (parameter MA tidak berpengaruh terhadap model)

$H_1: \theta_i \neq 0$  (parameter MA berpengaruh terhadap model)

Statistik uji:

$$t_{hitung} = \frac{(\text{parameter})_i}{SE(\text{parameter})} \text{ dengan } SE(\text{parameter}) = \sqrt{\frac{1-(\text{parameter})^2}{n}}$$

Kriteria uji baik untuk AR atau MA adalah  $H_0$  ditolak apabila  $|t_{hitung}| > t_{\frac{\alpha}{2}(n-p)}$  atau  $p\text{-value} < \alpha$

### 2.6.1.2 Uji Asumsi Residual Model ARIMA

#### a. Uji *White Noise* Residual Model ARIMA

Asumsi dasar bahwa residual bersifat *white noise* artinya tidak terdapat korelasi antar residual dengan *mean* sama dengan nol dan varian konstan. Uji independensi residual (*white noise*) dapat dilakukan dengan menggunakan statistik uji Ljung-Box.

Hipotesis:

$$H_0 : \rho_1 = \rho_2 = \dots = \rho_k \text{ (residual } white \text{ noise)}$$

$$H_1 : \text{Minimal ada satu nilai } \rho_k \neq 0; k=1, 2, \dots, k \text{ (residual tidak } white \text{ noise)}$$

Statistik Uji:

$$Q = n(n+2) \sum_{i=1}^k (n-i)^{-1} \hat{\rho}_i^2$$

Kriteria Uji:

$$H_0 \text{ ditolak apabila } Q > \chi^2_{(\alpha; k-p)} \text{ atau } p\text{-value} < \alpha$$

#### b. Uji Normalitas Residual Model ARIMA

Uji normalitas menggunakan uji *Kolmogorov Smirnov*. Hipotesis yang diuji adalah residual berdistribusi normal.

Hipotesis:

$$H_0 : F(a_t) = F_0(a_t) \text{ untuk semua nilai } a_t \text{ (residual data berdistribusi normal)}$$

$$H_1 : F(a_t) \neq F_0(a_t) \text{ untuk sekurang-kurangnya sebuah nilai } a_t \text{ (residual data tidak berdistribusi normal)}$$

Statistik Uji:

$$D = \sup |S(a_t) - F_0(a_t)|$$

Kriteria uji:

$$H_0 \text{ ditolak apabila } D > D_{1-\alpha/2} \text{ atau } H_0 \text{ ditolak jika } p\text{-value} < \alpha$$

#### c. Uji Homoskedastisitas Residual Model ARIMA

Uji homoskedastisitas digunakan untuk menguji kehomogenan ragam dari residual. Uji *Lagrange Multiplier* (LM) dapat digunakan untuk mendeteksi adanya proses ARCH/GARCH dengan cara meregresikan kuadrat dari residual model.

$$H_0 : \sigma_1^2 = \sigma_2^2 = \dots = \sigma_n^2, \text{ (varian residual sama)}$$

$$H_1 : \text{minimal ada satu nilai } \sigma_t^2 \neq 0, t=1, 2, \dots, n \text{ (varian residual tidak sama)}$$

Statistik uji:

$$LM = \frac{(SSR_0 - SSR_1)/k}{SSR_1/(n-2k-1)}$$

Kriteria Uji:

$$H_0 \text{ ditolak apabila } LM > \chi^2_{(\alpha; k)} \text{ atau } p\text{-value} < \alpha$$

### 2.6.1.3 Evaluasi Model Terbaik

Dalam penelitian ini, evaluasi model dilakukan dengan melihat nilai *Mean Squared Error* (MSE). Penentuan model terbaik untuk peramalan data runtun waktu dapat menggunakan *Mean Squared Error*. Semakin kecil nilai MSE yang diperoleh berarti semakin baik model yang digunakan. Rumus MSE didefinisikan sebagai berikut:

$$MSE = \frac{\sum_{t=1}^n (Z_t - \hat{Z}_t)^2}{n-p}$$

### 2.6.1.4 Peramalan

Peramalan berdasarkan suatu model dengan menurunkan distribusi bersyarat observasi yang akan datang jika diketahui observasi yang lalu. Data yang sudah

ditransformasi, harus dikembalikan ke bentuk data awal untuk mendapatkan nilai sesuai data awal.

### 2.6.2. Model *Autoregressive Integrated Fractionally Moving Average* atau ARFIMA

Model ARFIMA (p,d,q) dapat dituliskan sebagai:

$$\phi_p(B)(1-B)^d Z_t = \theta_q(B)a_t, \text{ dengan } d \text{ merupakan bilangan pecahan}$$

$$\nabla^d = (1-B)^d = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j (B)^j$$

dengan  $\binom{d}{j} = \frac{d!}{j!(d-j)!} = \frac{\Gamma(d+1)}{\Gamma(j+1)\Gamma(d-j+1)}$  dimana  $\Gamma(x)$  merupakan fungsi gamma.

Sehingga diperoleh:

$$(1-B)^d = \binom{d}{0} (-1)^0 (B)^0 + \binom{d}{1} (-1)^1 (B)^1 + \binom{d}{2} (-1)^2 (B)^2 + \dots$$

$$(1-B)^d = 1 - dB - \frac{1}{2}(1-d)dB^2 - \frac{1}{6}(1-d)(2-d)dB^3 + \dots$$

#### 2.6.2.1 Identifikasi pola *long memory*

Bila fungsi autokorelasi antara  $Z_t$  dengan  $Z_{t+k}$  turun lambat secara hiperbolik maka dapat diidentifikasi adanya ketergantungan jangka panjang dalam data yang merupakan ciri dari data memori jangka panjang (*long memory*). Atau dapat dibuktikan dengan nilai statistik *Hurst Exponen (H)* yang berada pada interval  $0,5 < H < 1$ , dapat dihitung dengan mengikuti rumus berikut:

1. Menghitung nilai rata-rata data (*mean*)

$$\mu = \frac{1}{n} \sum_{t=1}^n Z_t$$

2. Menghitung simpangandari masing-masing data  
 $Z(adj)_t = Z_t - \mu$  dengan  $t=1,2,\dots,n$
3. Menghitung simpangan kumulatif data

$$Z_t^* = \sum_{i=1}^t Z(adj)_i$$

4. Menghitung rentangan data  
 $R_t = \max(Z_1^*, Z_2^*, \dots, Z_t^*) - \min(Z_1^*, Z_2^*, \dots, Z_t^*)$
5. Menghitung standar deviasi dari masing-masing data

$$S_t = \sqrt{\frac{1}{t} \sum_{i=1}^t (Z_i - \mu)^2}$$

6. Menghitung *rescaled range statistics (R/S)*  
 $(R/S)_t = R_t/S_t$
7. Menentukan nilai *H* melalui statistik R/S dari data deret waktu

$$(R/S)_t = c \cdot t^H$$

$$\log(R/S)_t = c + H \log t$$

dengan:

$c$  = suatu konstanta

$H$  = *Eksponensial Hurst*

dan menaksir nilai *H* melalui metode *Ordinary Least Square (OLS)*, ditunjukkan pada persamaan berikut:

$$H = \frac{\sum_{t=1}^n (X_t - \mu_{x_t})(Y_t - \mu_{y_t})}{\sum_{t=1}^n (X_t - \mu_{x_t})^2}$$

### 2.6.2.2 Estimasi Parameter Pembeda

Menurut Idris dkk. (2014) metode yang digunakan dalam estimasi parameter pembeda (d) adalah metode *Geweke Porter-Hudak (GPH)*. Langkah-langkahnya adalah:

1. Menentukan nilai m dengan persamaan rumus  
 $m = n^{0.5}$  dengan n adalah banyaknya data pengamatan
2. Menentukan nilai  $\omega$  dengan persamaan rumus  
 $\omega_i = \frac{2\pi i}{n}$  dengan  $\pi = 3,14$  dan  $i = 1, 2, \dots, m$
3. Menghitung nilai  $\gamma_0$  dengan rumus  
 $\gamma_0 = \frac{\sum_{t=1}^{n-1} (Z_t - \bar{Z})^2}{n-1}$  dengan  $t = 1, 2, \dots, n$
4. Menentukan nilai dari  $I_z(\omega_i)$

$$I_z(\omega_i) = \frac{1}{2\pi} \left\{ \gamma_0 + 2 \sum_{t=1}^{n-1} \gamma_t \cos(t \cdot \omega_i) \right\}$$

5. Menentukan  $X_i$  sebagai variabel bebas

$$X_i = \ln \left( \frac{1}{4 \sin^2(\omega_i/2)} \right)$$

6. Menentukan  $Y_i$  sebagai variabel tak bebas

$$Y_i = \ln[I_z(\omega_i)]$$

Kemudian dibuat persamaan regresi antara  $X_i$  dan  $Y_i$  yaitu  $Y_i = a + bX_i$ . Nilai estimasi d adalah nilai koefisien pada parameter  $X_i$  atau nilai dari b.

## 2.7 Analisis Intervensi

Menurut Wei (2006), bentuk umum dari model intervensi adalah:

$$Z_t = \sum_{j=1}^k \frac{\omega_{sj}(B) B^{bj}}{\delta_{rj}(B)} I_{jt} + \frac{\theta_q(B)}{\phi_p(B)} a_t$$

Secara umum ada dua macam model intervensi, yaitu model fungsi *step* dan model fungsi *pulse*. Fungsi *step* adalah suatu bentuk intervensi yang terjadinya dalam kurun waktu yang panjang. Secara matematik, bentuk intervensi *step* ini biasanya dinotasikan sebagai berikut:

$$I_t = S_t = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases}$$

Sedangkan fungsi *pulse* adalah suatu bentuk intervensi yang terjadinya hanya dalam suatu waktu tertentu. Secara matematis, bentuk intervensi fungsi *pulse* ini biasanya dinotasikan sebagai berikut:

$$I_t = P_t = \begin{cases} 0, & t \neq T \\ 1, & t = T \end{cases}$$

### 2.7.1 Orde Intervensi

Menurut Nuvitasari dkk. (2009), plot residual ARIMA sebelum intervensi digunakan untuk menentukan orde model intervensi ( $b$ ,  $s$ , dan  $r$ ). Batas yang digunakan untuk menentukan garis signifikansi adalah  $\pm 2\sigma$ . Orde  $b$  merupakan waktu tunda hingga dampak intervensi mulai terjadi. Orde  $s$  merupakan lamanya suatu intervensi berpengaruh pada data setelah  $b$  periode, dan orde  $r$  menunjukkan lag setelah  $b$  dan  $s$  periode pada saat data sudah membentuk pola yang jelas. Dengan adanya kesulitan praktis dalam mengartikan prinsip-prinsip orde  $s$  dan  $r$ , maka dapat ditentukan bahwa  $r + s$  adalah sama

dengan banyaknya lag yang autokorelasinya signifikan sehingga dapat dilakukan proses coba-coba untuk memilih orde  $b$ ,  $s$  dan  $r$  yang menghasilkan model terbaik untuk peramalan.

### 2.7.2 Estimasi Parameter dan Uji Signifikansi Model Intervensi

Nuvitasari dkk. (2009) mengemukakan bahwa estimasi parameter untuk model intervensi dihitung berdasarkan bentuk umum dari model fungsi transfer sebagai berikut:

$$(1 - B)^d Z_t = \frac{\omega_s(B)^b}{\delta_r(B)} I_t + \frac{\theta_q(B)}{\phi_p(B)(1-B)^d} a_t$$

ataudapat dituliskan sebagai:

$$\delta_r(B)\phi_p(B)(1 - B)^d Z_t = \phi_p(B)(1 - B)^d \omega_s(B)^b I_t + \delta_r(B)\theta_q(B)a_t$$

maka diperoleh residualnya adalah:

$$a_t = \frac{\delta_r(B)\phi_p(B)Z_t - \phi_p(B)\omega_s(B)I_{t-b}}{\delta_r(B)\theta_q(B)}$$

dengan asumsi  $a_t$  adalah  $NID(0, \sigma_a^2)$ , maka:

$$S(\delta, \omega, \theta, \phi) = \sum_{t=1}^n \left[ \frac{\delta_r(B)\phi_p(B)Z_t - \phi_p(B)\omega_s(B)I_{t-b}}{\delta_r(B)\theta_q(B)} \right]^2$$

Untuk mendapatkan dugaan parameter digunakan metode *ordinary least square* yaitu:

$$S(\delta, \omega, \theta, \phi) = \sum_{t=1}^n a_t^2$$

## 3. METODOLOGI PENELITIAN

### 3.1. Jenis dan Sumber Data

Jenis data yang digunakan sebagai studi kasus merupakan data sekunder yaitu data jumlah penumpang kereta api kelas lokal ekonomi yang diperoleh dari PT Kereta Api Indonesia (KAI) Daerah Operasi IV Semarang. Data yang diambil ialah data bulanan periode Januari 2009 sampai Juni 2017.

### 3.2. Variabel Penelitian

Variabel penelitian yang digunakan dalam metode ARIMA dan ARFIMA adalah jumlah penumpang kereta api ( $Z_t$ ). Variabel penelitian yang digunakan pada metode intervensi *step* terdiri dari satu variabel yaitu jumlah penumpang kereta api sebelum dan sesudah terjadinya kenaikan drastis. Secara matematis variabel intervensi fungsi *step* dapat dinotasikan sebagai berikut:

$$I_t = S_t = \begin{cases} 0, & t < T \\ 1, & t \geq T \end{cases}$$

T adalah waktu terjadinya intervensi yaitu kenaikan ekstrem jumlah penumpang kereta api pada saat  $T=85$  (Januari 2016) yang diduga terjadi karena peningkatan *Public Service Obligation (PSO)* melalui subsidi yang meningkat sebesar 21,24%.

### 3.3. Teknik Pengolahan Data

#### 3.3.1 Metode ARIMA

1. Membuat *plot time series* data jumlah penumpang kereta api
2. Melakukan transformasi apabila data tidak stasioner dalam varian
3. Melakukan pembedaan (*differencing*) terhadap data apabila tidak stasioner dalam *mean*
4. Membuat plot ACF dan PACF data yang telah dilakukan transformasi dan pembedaan (*differencing*)

5. Melakukan pemodelan dengan metode ARIMA.
6. Melakukan uji signifikansi dan uji asumsi parameter model ARIMA.
7. Melakukan evaluasi model ARIMA terbaik dengan melihat MSE yang terkecil.
8. Melakukan peramalan berdasarkan model ARIMA terbaik.

### 3.3.2 Metode Intervensi

Langkah-langkah pemodelan intervensi hampir sama dengan ARIMA. Namun, dalam intervensi terlebih dahulu dianalisis model sebelum intervensi, lalu menentukan orde intervensi dan melakukan estimasi parameter model intervensi.

### 3.3.3 Metode ARFIMA

Langkah-langkah pemodelan ARFIMA hampir sama dengan ARIMA. Namun, dalam ARFIMA perlu dibuktikan apakah terdapat efek *long memory* atau tidak, apabila terdapat efek *long memory* maka diolah menggunakan ARFIMA. Kemudian dilakukan estimasi pembeda (d) pada data.

## 4. HASIL DAN PEMBAHASAN

### 4.1 Metode ARIMA

#### 4.3.1 Stasioneritas

Data jumlah penumpang tidak stasioner dalam varian, diperoleh nilai  $\lambda=0,5$  maka dilakukan transformasi. Setelah dilakukan transformasi, nilai diperoleh  $\lambda=0,5$ . Maka dilakukan transformasi yang kedua diperoleh  $\lambda=1$ . Maka data sudah stasioner dalam varian. Selain itu dilakukan pengecekan stasioneritas dalam *mean*. Diperoleh bahwa data tidak stasioner dalam *mean*, maka perlu dilakukan *differencing*. Setelah dilakukan *differencing*, data sudah stasioner dalam *mean*.

#### 4.3.2 Estimasi Parameter dan Uji Signifikansi model ARIMA

Model ARIMA yang mungkin terbentuk adalah:

- |                         |                               |
|-------------------------|-------------------------------|
| 1. ARIMA (0, 1, 1)      | 9. ARIMA ([13], 1, 1)         |
| 2. ARIMA (0, 1, [13])   | 10. ARIMA ([13], 1, [13])     |
| 3. ARIMA (0, 1, [1,13]) | 11. ARIMA ([13], 1, [1,13])   |
| 4. ARIMA (1, 1, 0)      | 12. ARIMA ([1,13], 1, 0)      |
| 5. ARIMA (1, 1, 1)      | 13. ARIMA ([1,13], 1, 1)      |
| 6. ARIMA (1, 1, [13])   | 14. ARIMA ([1,13], 1, [13])   |
| 7. ARIMA (1, 1, [1,13]) | 15. ARIMA ([1,13], 1, [1,13]) |
| 8. ARIMA ([13], 1, 0)   |                               |

Dari model tersebut dilakukan uji signifikansi parameter, sehingga pada taraf signifikansi  $\alpha = 5\%$ , terdapat parameter model yang signifikan dan tidak signifikan. Model yang signifikan ialah ARIMA (0,1,[13]) dan ARIMA ([13],1,0).

#### 4.3.3 Evaluasi asumsi residual Model ARIMA

- a. Uji White Noise Residual Model ARIMA  
Pada taraf signifikansi  $\alpha = 5\%$ , kedua model yang signifikan yaitu model ARIMA (0, 1, [13]) dan ARIMA ([13], 1, 0) memenuhi asumsi *white noise* residual.
- b. Uji Normalitas Residual Model ARIMA  
Model ARIMA (0,1,[13]) dan ARIMA ([13],1,0) tidak memenuhi asumsi normalitas residual.
- c. Uji Homoskedastisitas Residual Model ARIMA



Model ARIMA (0,1,[13]) dan ARIMA ([13],1,0) memenuhi asumsi homoskedastisitas residual yang artinya memiliki varian residual yang sama.

#### 4.3.4 Evaluasi Model ARIMA

Model	Signifikansi Parameter	White Noise	Normalitas	Homoskedastisitas	MSE
ARIMA(0,1,[13])	Ya	Ya	Tidak	Ya	1,899554
ARIMA([13],1,0)	Ya	Ya	Tidak	Ya	1,870844

Sehingga model terbaik untuk data jumlah penumpang kereta api DAOP IV Semarang dengan menggunakan metode ARIMA adalah model ARIMA ([13],1,0) dengan bentuk persamaan:

$$Z_t = Z_{t-1} - 0,250632Z_{t-13} + 0,250632Z_{t-14} + a_t$$

#### 4.3.5 Peramalan Model ARIMA

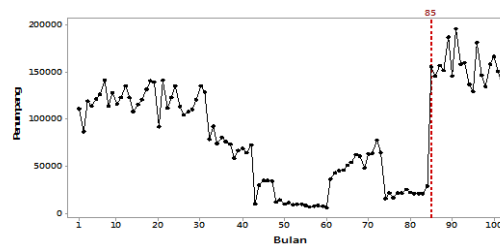
Hasil Peramalan Jumlah Penumpang Model ARIMA ([13],1,0)

Bulan	Hasil Peramalan
Jul-17	145.253
Agu-17	134.440
Sep-17	142.249
Okt-17	141.900
Nov-17	147.596
Des-17	149.607

## 4.2 Metode Intervensi

### 4.2.1. Pemodelan data sebelum intervensi

Berikut merupakan plot time series data penumpang sebelum dan sesudah intervensi:



#### 4.2.1.1 Stasioneritas

Data jumlah penumpang sebelum intervensi tidak stasioner dalam varian, diperoleh nilai  $\lambda=0,5$  maka dilakukan transformasi. Setelah dilakukan transformasi, nilai diperoleh  $\lambda=1$ . Maka data sudah stasioner dalam varian. Selain itu dilakukan pengecekan stasioneritas dalam *mean*. Diperoleh bahwa data tidak stasioner dalam *mean*, maka perlu dilakukan *differencing*. Setelah dilakukan *differencing*, data sudah stasioner dalam *mean*.

#### 4.2.1.2 Estimasi Parameter dan Uji Signifikansi model ARIMA sebelum Intervensi

Model ARIMA yang mungkin terbentuk adalah:

1. ARIMA (0, 1, 1)
2. ARIMA (0, 1, [13])
3. ARIMA (0, 1, [1,13])
4. ARIMA (1, 1, 0)
9. ARIMA ([13], 1, 1)
10. ARIMA ([13], 1, [13])
11. ARIMA ([13], 1, [1,13])
12. ARIMA ([1,13], 1, 0)

5. ARIMA (1, 1, 1)
6. ARIMA (1, 1, [13])
7. ARIMA (1, 1, [1,13])
8. ARIMA ([13], 1, 0)
13. ARIMA (([1,13], 1, 1)
14. ARIMA ([1,13], 1, [13])
15. ARIMA (([1,13], 1, [1,13])

Model yang signifikan ialah ARIMA (0,1,1), ARIMA (0,1,[13]), ARIMA (0,1,[1,13]), ARIMA (1,1,0), ARIMA (1,1,[13]), ARIMA ([13],1,0), ARIMA ([13],1,1) dan ARIMA ([1,13],1,0).

#### 4.2.1.3 Evaluasi asumsi residual Model ARIMA sebelum Intervensi

- a. Uji White Noise Residual Model ARIMA sebelum Intervensi  
Pada taraf signifikansi  $\alpha = 5\%$ , model yang memenuhi asumsi *white noise* residual adalah ARIMA (0,1, 1), ARIMA (0,1,[1,13]), ARIMA (1,1,0), ARIMA (1,1,[13]), ARIMA ([13],1,1) dan ARIMA ([1,13],1,0).
- b. Uji Normalitas Residual Model ARIMA sebelum Intervensi  
Model ARIMA sebelum intervensi tidak memenuhi asumsi normalitas residual.
- c. Uji Homoskedastisitas Residual Model ARIMA sebelum Intervensi  
Semua model memenuhi asumsi homoskedastisitas artinya model tersebut memiliki varian residual yang sama.

#### 4.2.1.4 Evaluasi Model ARIMA sebelum Intervensi

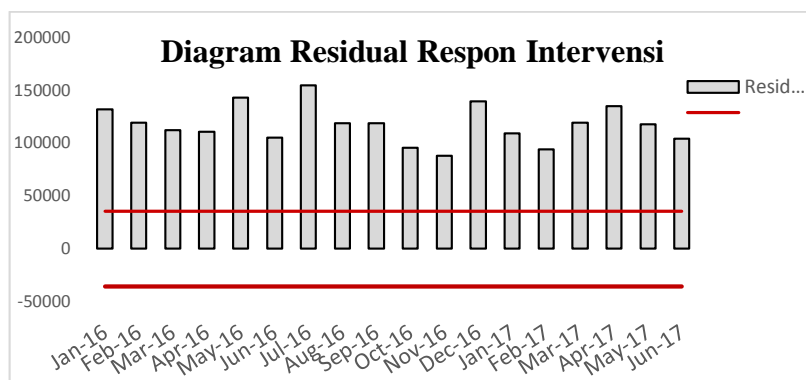
Model	Signifikansi Parameter	White Noise	Normalitas	Homo-skedastisitas	MSE
ARIMA (0, 1, 1)	Ya	Ya	Tidak	Ya	1218,622
ARIMA (0, 1, [13])	Ya	Tidak	Tidak	Ya	1184,900
ARIMA (0, 1, [1,13])	Ya	Ya	Tidak	Ya	1049,082
ARIMA (1, 1, 0)	Ya	Ya	Tidak	Ya	1222,511
ARIMA (1, 1, [13])	Ya	Ya	Tidak	Ya	1087,421
ARIMA ([13], 1, 0)	Ya	Tidak	Tidak	Ya	1128,761
<b>ARIMA ([13], 1, 1)</b>	<b>Ya</b>	<b>Ya</b>	<b>Tidak</b>	<b>Ya</b>	<b>1044,231</b>
ARIMA ([1,13], 1, 0)	Ya	Ya	Tidak	Ya	1069,586

Metode ARIMA terbaik sebelum intervensi adalah adalah ARIMA ([13],1,1). Persamaan dari model tersebut adalah:

$$Z_t = Z_{t-1} - 0,384553Z_{t-13} + 0,384553Z_{t-14} + a_t + 0,300783a_{t-1}$$

#### 4.2.2. Penentuan Orde Intervensi

Dapat diketahui bahwa intervensi yang terjadi dimulai pada saat T=85 atau pada bulan Januari 2016 yang menyebabkan dampak yang bersifat jangka panjang terhadap jumlah penumpang kereta api sehingga model intervensi yang diduga ialah intervensi fungsi *step*. Pada T=85 (Januari 2016) yang artinya intervensi mulai terjadi pada saat itu juga sehingga waktu tunda (b) adalah 0. Orde *s* merupakan lamanya satu intervensi berpengaruh pada data setelah b periode. Dapat dilihat bahwa plot-plot residual respon yang keluar dari garis signifikansi merupakan banyaknya intervensi sehingga diperoleh nilai *s* adalah 18. Menurut Makridakis dkk. (1995), orde *r* dapat ditentukan dengan nilai maksimum *r+s* adalah banyaknya lag keluar yang signifikan. Berdasarkan analisis tersebut, maka diperoleh model intervensi dengan dugaan orde b=0, s=18, r=0.



### 4.2.3. Analisis Intervensi

#### 4.2.3.1 Estimasi Parameter dan Uji Signifikansi model Intervensi

Dilakukan dilakukan uji signifikansi parameter, sehingga pada taraf signifikansi  $\alpha = 5\%$ , Model yang signifikan ialah ARIMA([13],1,1) dengan orde  $b=0$ ,  $s=18$ ,  $r=0$ .

#### 4.2.3.2 Evaluasi asumsi residual Model Intervensi

- Uji White Noise Residual Model Intervensi  
Pada taraf signifikansi  $\alpha = 5\%$ , model intervensi ARIMA ([13],1,1) dengan  $b=0$ ,  $s=18$ , dan  $r=0$  memenuhi asumsi *white noise* residual atau tidak terdapat korelasi antar residual.
- Uji Normalitas Residual Model Intervensi  
Model ARIMA ([13],1,1) dengan  $b=0$ ,  $s=18$ , dan  $r=0$  tidak memenuhi asumsi normalitas residual.
- Uji Homoskedastisitas Residual Model Intervensi  
Model ARIMA ([13],1,1),  $b=0$   $s=18$   $r=0$  memenuhi asumsi homoskedastisitas atau varian residual sama.

#### 4.2.3.3 Evaluasi Model Intervensi

Persamaan dari model ARIMA ([13], 1, 1),  $b=0$   $s=18$   $r=0$  ialah:

$$Z_t = Z_{t-1} + 227,55798S_t - 0,1S_{t-18} + \frac{(a_t - 0,34723a_{t-1})}{(1 - B + 0,37157B^{13} - 0,37157B^{14})}$$

dengan:

$$S_t = \begin{cases} 0, & t < (T = 85) \\ 1, & t \geq (T = 85) \end{cases}$$

dimana  $T=85$  merupakan intervensi yang terjadi pada bulan Januari 2016

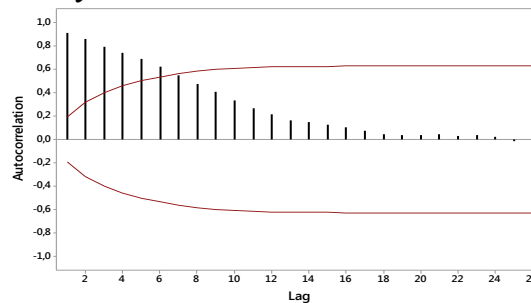
#### 4.2.3.4 Peramalan Model Intervensi

Hasil Peramalan Jumlah Penumpang Model ARIMA ([13], 1, 1),  $b=0$   $s=18$   $r=0$  :

Bulan	Hasil Peramalan
Juli 2017	152.528
Agustus 2017	135.468
September 2017	147.913
Oktober 2017	147.370
November 2017	156.061
Desember 2017	159.045

### 4.3 Metode ARFIMA

#### 4.3.1 Identifikasi *Long Memory*



Pada gambar tersebut, dapat dilihat bahwa lag yang muncul turun lambat secara hiperbolik. Secara visual dapat dikatakan bahwa data mengandung efek memory jangka panjang. Identifikasi model secara formal dapat dilihat melalui perhitungan nilai *Hurst Exponent* ( $H$ ). Nilai  $H$  sebesar 0,69, karena nilai *Hurst* terletak pada interval  $0,5 < H < 1$  maka dapat disimpulkan bahwa terdapat memori jangka panjang pada data.

#### 4.3.2 Estimasi Parameter Pembeda ( $d$ )

Estimasi nilai  $d$  dilakukan menggunakan program R dengan metode *GPH* (*Geweke and Porter-Hudak*). Diperoleh nilai  $d$  sebesar 1,222974 dan dilakukan diferensiasi terhadap data sesuai dengan nilai  $d$  tersebut.

#### 4.3.3 Estimasi Parameter dan Uji Signifikansi model ARFIMA

Dari model tersebut dilakukan uji signifikansi parameter, pada taraf signifikansi  $\alpha = 5\%$ , terdapat parameter model yang signifikan dan tidak signifikan. Model yang signifikan ialah ARFIMA (0,  $d$ , [1,13]).

#### 4.3.4 Evaluasi asumsi residual Model ARFIMA

- Uji White Noise Residual Model ARFIMA  
Model ARFIMA (0,  $d$ , [1,13]) memenuhi asumsi *white noise* residual karena memiliki *p-value* yang lebih besar dari  $\alpha = 5\%$
- Uji Normalitas Residual Model ARFIMA  
Model ARFIMA (0,  $d$ , [1,13]) tidak memenuhi asumsi normalitas.
- Uji Homoskedastisitas Residual Model ARFIMA  
Pada saat taraf signifikansi  $\alpha = 5\%$ , dapat disimpulkan bahwa model ARFIMA (0,  $d$ , [1,13]) memenuhi asumsi homoskedastisitas residual.

#### 4.3.5 Evaluasi Model ARFIMA

Nilai MSE dari model ARFIMA (0,  $d$ , [1,13]) adalah 0,00964. Bentuk persamaannya adalah:

$$Z_t = \frac{(1 + 0,666414B + 0,304954B^{13}) a_t}{(1 - B)^{0,367546}}$$

dengan  $(1 - B)^{0,367546}$  dapat didekati dengan ekspansi binomial sebagai berikut:

$$(1 - B)^{0,367546} = 1 - 0,367546B - 0,116228B^2 - 0,063246B^3 + \dots$$

#### 4.3.6 Peramalan dengan Metode ARFIMA

Hasil Peramalan Jumlah Penumpang Model ARFIMA (0, d, [1,13])

Bulan	Hasil Peramalan
Juli 2017	129.543
Agustus 2017	121.919
September 2017	123.230
Oktober 2017	124.348
November 2017	132.299
Desember 2017	143.719

#### 5. KESIMPULAN

Berdasarkan pada hasil analisis penelitian yang telah dilakukan, dapat disimpulkan beberapa hal sebagai berikut:

1. Model ARIMA yang terbentuk adalah ARIMA ([13],1,0) dengan MSE sebesar 1,870844. Model tersebut ditulis dengan persamaan:

$$Z_t = Z_{t-1} - 0,250632Z_{t-13} + 0,250632Z_{t-14}$$

2. Model Intervensi fungsi *step* yang terbentuk adalah ARIMA ([13],1,1) b=0, s=18, r=0, dengan MSE sebesar 1124. Dapat ditulis dengan persamaan:

$$Z_t = Z_{t-1} + 227,55798S_t - 0,1S_{t-18} + \frac{(a_t - 0,34723a_{t-1})}{(1 - B + 0,37157B^{13} - 0,37157B^{14})}$$

3. Model ARFIMA yang terbentuk pada adalah ARFIMA (0,d,[1,13]) dimana d=0,367546. Nilai MSE model tersebut adalah 0,00964, dapat ditulis dengan persamaan:

$$Z_t = \frac{(1 + 0,666414B + 0,304954B^{13}) a_t}{(1 - B)^{0,367546}}$$

$$(1 - B)^{0,367546} = 1 - 0,367546B - 0,116228B^2 - 0,063246B^3 + \dots$$

4. Dari metode ARIMA, Intervensi dan ARFIMA, metode terbaik dalam meramalkan jumlah penumpang adalah metode ARFIMA karena memiliki MSE terkecil yaitu 0,00964 dengan model ARFIMA (0,d,[1,13]) dimana d=0,367546.

#### DAFTAR PUSTAKA

- Abdulkadir, M. 1998. *Hukum Pengangkutan Niaga*. Bandung: Citra Aditya Bakti.
- Cahyandari, R., dan Erviana, R. 2015. *Peramalan Kurs Jual Uang Kertas MataUangSingapore Dollar (SGD) terhadap Rupiah Menggunakan ModelARFIMA(Autoregressive Fractionally Integrated Moving Average)*.Kubik Vol.1, No.1.
- Chatfield, C. 2000. *Time Series Forecasting*. Florida: CRC Press.
- Hosking, J. R. M. 1981. *Fractional Differencing*. Biometrika, Vol.68 Page 165-176.
- KAI. 2014. *Pemerintah Subsidi Penumpang KA Kelas Ekonomi*. <https://www.kereta-api.co.id>. Diakses: 10 Mei 2018.
- KAI. 2018. *Angkutan Penumpang*. <https://www.kereta-api.co.id>. Diakses: 10 Mei 2018.
- Makridakis, S., Wheelright, S.C., dan McGee V. E. 1995. *Metode dan Aplikasi Peramalan*. Jilid satu. Edisi kedua. Terjemahan dari: Ir. Untung Sus Andriyanto, M.Sc. dan Ir Abdul Basith, M.Sc. Jakarta: Erlangga.

- Nuvtasari, E., Suhartono., dan Wibowo, S. H. 2009. *Analisis Intervensi Multi Input Fungsi Step dan Fungsi Pulse untuk Peramalan Kunjungan Wisatawan ke Indonesia*. Thesis. Institut Teknologi Sepuluh November, Surabaya.
- Rosadi, D. 2012. *Ekonometrika & Analisis Runtun Waktu Terapan dengan Eviews*. Yogyakarta: Andi.
- Soejoeti, Z. 1987. *Materi Pokok Analisis Runtun Waktu*. Jakarta: Karunika.
- Suhartono. 2007. *Teori dan Aplikasi Model Intervensi Fungsi Pulse*. Jurnal Ilmiah MatStat, Vol.7, No.2.
- Republik Indonesia. 1998. *Peraturan Pemerintah Nomor 69 Tahun 1998 tentang prasarana dan sarana kereta api*. Jakarta.
- Wei, W. W. S. 2006. *Time Series Analysis, Univariate and Multivariate Methods*. Canada: Addison Wesley Publishing Company.