



Penerapan Model Seasonal Autoregressive Integrated Moving Average (SARIMA) pada Jumlah Penumpang Kereta Api di Sumatera Barat

Serly Cania¹, Darvi Mailisa Putri^{✉2}, Ilham Dangu Rianjaya³

Fakultas Sains dan Teknologi UIN Imam Bonjol Padang^{1,2,3}

email: serlycania378@gmail.com¹, darvimailisa@uinib.ac.id², ilham.rianjaya@uinib.ac.id³

Received 15 Agustus 2023,

Accepted 10 Oktober 2023,

Published 10 Oktober 2023

Abstrak

Penelitian ini bertujuan untuk memperoleh model dan menentukan model terbaik dari data jumlah penumpang kereta api di Sumatera Barat dengan menggunakan *Seasonal Autoregressive Integrated Moving Average* (SARIMA). Data penelitian bersumber dari data sekunder yang diperoleh dari PT.KAI (Persero) Divisi Regional II Sumatera Barat untuk melihat jumlah penumpang kereta api dengan rentang waktu Januari 2017 sampai dengan April 2020. Hasil penelitian menunjukkan model terbaik yang diperoleh adalah SARIMA (0,1,1)(0,1,0)⁶. Pemilihan model terbaik berdasarkan nilai AIC terkecil dari beberapa model yang telah diperoleh melalui plot ACF dan PACF. Berdasarkan model terbaik diperoleh hasil peramalan mendekati data sebenarnya, sehingga model SARIMA(0,1,1)(0,1,0)⁶ layak digunakan untuk peramalan.

Kata Kunci: Penumpang Kereta Api; SARIMA; Peramalan; Deret Waktu.

Abstract

This study aims to obtain a model and determine the best model of the results of the number of train passengers in West Sumatra using Seasonal Autoregressive Integrate Moving Average (SARIMA). The research data comes from secondary data obtained from PT.KAI (Persero) Regional Division II West Sumatra to see the number of train passengers with a time span of January 2017 to April 2020. The results showed that the best model obtained was SARIMA (0,1,1)(0,1,0)⁶. Selection of the best model based on the smallest AIC value of several models that have been obtained through ACF and PACF plots. Based on the best model, the forecasting results are close to the actual data, so the SARIMA(0,1,1)(0,1,0)⁶ model is suitable for forecasting.

Keywords: Number of Train Passengers; SARIMA; Forecasting; Time Series.

✉ Corresponding author

PENDAHULUAN

Saat ini terdapat tiga jenis kereta api yang beroperasi di Sumatera Barat yaitu kereta api Sibinuang, kereta api Lembah Anai dan kereta api Minangkabau Express. Berdasarkan data beberapa tahun terakhir jumlah penumpang kereta api mengalami peningkatan. Terutama di hari-hari libur misalnya, hari sabtu dan minggu, libur hari raya, libur panjang. Peristiwa tersebut mengakibatkan kereta api mengalami lonjakan penumpang sehingga tidak mendapatkan tiket di hari tersebut. Kenaikan jumlah penumpang kereta api sejak 4 tahun terakhir paling tinggi terjadi pada bulan Desember tahun 2019 yaitu 175.128 orang (BPS, 2023).

Berdasarkan data tersebut dapat diketahui kereta api tidak pernah kehilangan minat para penumpang, dikarenakan adanya beberapa faktor yang mempengaruhi masyarakat pecinta kereta api untuk menggunakan kereta api sebagai alat angkut transportasi dibandingkan dengan alat angkut lainnya seperti bus, travel dan lain-lain. Adapun faktor-faktor yang mempengaruhi yaitu, kenyamanan pada sarana dan ruang tunggu, fasilitas dan keamanan dari tindak kriminal, kesesuaian jadwal keberangkatan, terhindar dari pedagang asongan dan harga tiket yang murah sehingga kereta api diminati oleh masyarakat.

Deret waktu merupakan barisan suatu nilai pengamatan yang diukur dalam rentang waktu tertentu dalam interval waktu yang sama [1]. Analisis deret waktu didasarkan pada urutan poin data yang ditempatkan secara merata yang digunakan untuk menemukan bentuk pola dari data masa lalu sehingga dapat melakukan peramalan untuk masa yang akan datang. Peramalan merupakan suatu seni dan ilmu pengetahuan dalam memprediksi peristiwa pada masa mendatang dimana data di masa lampau dikumpul, dipelajari dan dianalisis [2]. Hasil peramalan tidaklah selalu akurat 100% karena peramalan sendiri bersifat tidak pasti.

Oleh Karena itu, perlu dilakukan prakiraan model terbaik yang dapat digunakan untuk mengetahui tingkat kenaikan penumpang, sehingga pihak perusahaan dapat melakukan upaya-upaya dalam mengatasi kenaikan jumlah penumpang kereta api. Jumlah penumpang merupakan suatu data yang bersifat musiman dan *trend*. Hal ini dapat dilihat dari jumlah penumpang pada setiap tahunnya di bulan-bulan tertentu mengalami peningkatan.

Adapun metode-metode yang digunakan dalam melakukan peramalan yaitu metode *Moving Average*, ARIMA, SARIMA, *Exponential Smoothing* dan lain-lain. ARIMA adalah salah satu metode yang sering dipakai untuk peramalan. Salah satunya penelitian yang dilakukan oleh Putri, dkk mengkaji pergerakan dan peramalan harga saham pada suatu Perusahaan Terbatas (PT) [3]. Metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA) merupakan pengembangan dari model ARIMA yang diperkenalkan oleh Georgi Box dan Gwilym Jenkins. Kelebihan metode SARIMA ialah relatif mudah untuk digunakan dalam memprediksi data deret waktu berpola musiman [4]. Oleh karena itu, metode SARIMA adalah salah

satu metode yang dapat digunakan untuk meramalkan jumlah penumpang di masa yang akan datang.

Penelitian mengenai *Seasonal Autoregressive Integrated Moving Average* (SARIMA) ini pernah dilakukan oleh [4] yang berjudul Peramalan Curah Hujan dengan Pendekatan SARIMA studi kasus curah hujan bulanan di kota Ambon, Provinsi Maluku. Hasil yang diperoleh adalah model $(1 - \beta_{12}B^{12})(1 - B)X_t = (1 + b_2B^2 + b_7B^7)\varepsilon_t$ dengan hasil peramalan menggunakan model tersebut untuk 12 bulan ke depan tidak begitu sesuai dibandingkan dengan data aktual tahun 2014. Kemudian penelitian yang dilakukan oleh [5] yang berjudul Penerapan Metode SARIMA untuk Peramalan Jumlah Pengunjung Wisata Taman Nasional Bantimurung Bulusaraung Maros. Hasil yang diperoleh jumlah pengunjung mengalami kenaikan dan penurunan dalam periode satu tahun. Jumlah pengunjung tertinggi terjadi pada bulan Desember tahun 2020 sebanyak 19061 orang dan jumlah terendah terjadi pada bulan Januari tahun 2020 sebanyak 15067 orang. Adapun model terbaik yaitu SARIMA(1,1,3)(2,1,1)¹².

Penelitian mengenai SARIMA juga pernah dilakukan oleh [6] dengan judul Prediksi Kualitas Air Menggunakan metode SARIMA. Hasil yang diperoleh bahwa model SARIMA yang diperoleh dapat digunakan dan memperoleh keakuratan peramalan pada masing-masing data parameter yang diuji. Diperoleh hasil untuk pengujian dataset kedalaman 30 cm dan kedalaman 60 cm untuk parameter temperature, yaitu MSE < 0,1 dan RMSE < 0,02. Untuk parameter PH, MSE < 0,1 dan RMSE < 0,1, serta parameter turbidity MSE < 0,02 dan RMSE < 0,13. Dari hasil tersebut menunjukkan bahwa system ini dapat memprediksi kualitas air dengan data yang telah lalu.

Selanjutnya penelitian yang dilakukan oleh [7] yang berjudul Pemodelan SARIMA Untuk Memprediksi Jumlah Kasus Covid-19 di Padang. Hasil yang diperoleh adalah model SARIMA(0,1,1)(0,1,1)⁷ merupakan model terbaik estimasi parameter yang signifikan dan memenuhi white noise dengan nilai MSE sebesar 330.9333. Hasil peramalan untuk sebulan kedepan menunjukkan kasus pasien Covid-19 masih fluktuatif mulai dari 20 orang sampai 66 orang. Kemudian penelitian yang dilakukan oleh [8] dengan judul Penerapan Metode SARIMA untuk Memprediksi Tingkat Inflasi di Indonesia. Diperoleh model terbaik dengan nilai AIC terendah yaitu SARIMA(1,0,1)(1,1,1)⁷ dengan nilai MAPE sebesar 5.19%. Hasil prediksi menunjukkan bahwa dalam beberapa bulan kedepan akan terjadi peningkatan tingkat inflasi.

Kemudian penelitian yang dilakukan oleh [9] dengan judul Peramalan Jumlah Penumpang Kereta Api di Indonesia Menggunakan Metode SARIMA. Hasil yang diperoleh yaitu model terbaik SARIMA(1,1,2)(0,1,1)¹². Diperoleh prediksi jumlah keseluruhan penumpang kereta api di Indonesia pada tahun 2020 adalah 492.230.700 orang penumpang dengan nilai MSE sebesar 0,046875 dan nilai MAPE sebesar 6,26%.

METODOLOGI

Penelitian ini menggunakan data jumlah penumpang kereta api di PT. KAI (Persero) Divisi Regional II Sumatera Barat dengan rentang waktu januari 2017 sampai dengan April 2020. Langkah-langkah analisis data dalam penelitian ini adalah:

1. Membentuk plot deret waktu. Digunakan untuk melihat sebaran data apakah telah sesuai dengan kriteria yang diinginkan.
2. Mengidentifikasi kestasioneran data.

Stasioner deret waktu merupakan asumsi yang mendasari bahwa proses suatu deret pengamatan tidak berubah seiring dengan adanya perubahan waktu [10]. Jika suatu deret waktu Z_t dikatakan stasioner, maka *mean* dan variansi deret waktu tidak dipengaruhi oleh berubahnya waktu pengamatan, sehingga proses berada dalam keseimbangan statistik.

Uji *Augmented Dickey Fuller* atau *unit root test* (uji akar *unit root*) merupakan uji formalitas untuk kestasioneritas suatu data deret waktu. Uji ini dikenalkan oleh David Dickey dan Wayne Fuller. Hipotesis yang digunakan yaitu

$$H_0 : \hat{\delta} = 0 \text{ (data tidak stasioner)}$$

$$H_1 : \hat{\delta} \neq 0 \text{ (data stasioner)}$$

Statistik uji ADF adalah sebagai berikut

$$ADF = \frac{\hat{\delta}}{se(\hat{\delta})} \quad (1)$$

Dimana $se(\hat{\delta})$ adalah *standard error* untuk $\hat{\delta}$. Dengan kriteria jika nilai mutlak statistik uji ADF lebih besar dibandingkan nilai mutlak kritis tabel *Dickey-Fuller* maka tolak H_0 . Sehingga dapat disimpulkan data stasioner. Sedangkan jika nilai mutlak statistik ADF lebih kecil dibandingkan nilai mutlak kritis tabel *Dickey Fuller* maka terima H_0 , dengan kata lain data tidak stasioner [11].

Proses *differencing* dilakukan jika data tidak stasioner dalam *mean*. Proses *differencing* orde pertama merupakan selisih antara data ke t dengan data ke $t - 1$, yaitu

$$\nabla Z_t = Z_t - Z_{t-1} \quad (2)$$

Adapun bentuk *differencing* orde ke- d , yaitu

$$\nabla^d Z_t = (1 - B)^d Z_t \quad (3)$$

Sedangkan transformasi data dapat dilakukan jika data tidak stasioner terhadap varians.

3. Identifikasi model. Hal ini dapat melihat hasil plot ACF dan PACF serta pengujian ADF.

Autocorrelation Function (ACF) adalah suatu hubungan linear antar pengamatan pada waktu ke- t $\{Z_t\}$ dan $\{Z_{t+k}\}$ dari proses yang sama yang hanya terpisah k lag waktu. Dengan rumus

$$\rho_k = \frac{\sum_{t=k+1}^{n-k} (Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^n (Z_t - \bar{Z})^2} = \frac{\gamma_k}{\gamma_0} \quad (4)$$

Partial Autocorrelation Function (PACF) adalah fungsi yang menunjukkan besarnya korelasi parsial antara pengamatan pada waktu ke- t , dengan pengamatan waktu sebelumnya. Dengan rumus

$$\phi_{kk} = \frac{\rho_k - \sum_{j=1}^{k-1} \phi_{k-1, j} \rho_j}{1 - \sum_{j=1}^{k-1} \phi_{k-1, j} \rho_j} \quad (5)$$

4. Uji signifikan parameter model.

Uji signifikansi parameter digunakan untuk menguji apakah suatu parameter model layak digunakan atau tidak. Adapun hipotesis pengujiannya sebagai berikut

Hipotesis:

H_0 : Parameter model tidak signifikan

H_1 : Parameter model signifikan

Kriteria Pengujian:

Tolak H_0 apabila $P - value < \alpha$ dengan α taraf signifikansi yaitu = 0,05.

5. Melakukan uji normalitas residual dan *white noise*.

Uji normalitas residual dilakukan untuk mengetahui apakah residu berdistribusi normal atau tidak. Pengujian ini dapat dilakukan dengan analisis grafik normal *probability plot*. Adapun hipotesis pengujiannya sebagai berikut

Hipotesis:

H_0 : Data residual berdistribusi normal

H_1 : Data residual tidak berdistribusi normal

Kriteria pengujian:

Tolak H_0 jika $probability < \alpha$ dengan α adalah taraf signifikansi 5%. Suatu data dikatakan berdistribusi normal, jika nilai residunya mendekati nol dimana nilai residu merupakan selisih antara nilai sesungguhnya dengan nilai dugaannya.

Menurut [12] Proses $\{Z_t\}$ dikatakan *white noise* dengan rata-rata 0 dan ragam σ^2 , dapat ditulis

$$\{Z_t\} \sim WN(0, \sigma^2) \quad (6)$$

Hipotesis:

H_0 : $\rho_1 = \rho_2 = \dots = \rho_k$ (residual memenuhi *white noise*)

H_1 : Minimal ada satu nilai $\rho_k \neq 0; k = 1, 2, 3, \dots$ (residual tidak memenuhi *white noise*).

Statistik uji Ljung-Box:

$$Q = n(n+2) \sum_{K=1}^K \frac{\rho_K^2}{n-K} \quad (7)$$

Kriteria Pengujian:

Tolak H_0 apabila $P - value < \alpha$.

6. Dilakukan pemilihan model terbaik dengan menggunakan AIC terkecil dari masing-masing model.

Sebelum menentukan model terbaik berdasarkan beberapa model yang signifikan, hanya satu model yang akan dipilih paling baik untuk digunakan dalam peramalan. salah satu cara pemilihan kriteria model terbaik yaitu dengan menggunakan *Akaike's Information Criteria* (AIC). Dengan rumus sebagai berikut

$$AIC = -2\log\hat{\sigma}^2 + 2k \quad (8)$$

Keterangan:

$\log\hat{\sigma}^2$ = Ukuran *likelihood*

k = Banyak parameter

n = Banyak pengamatan

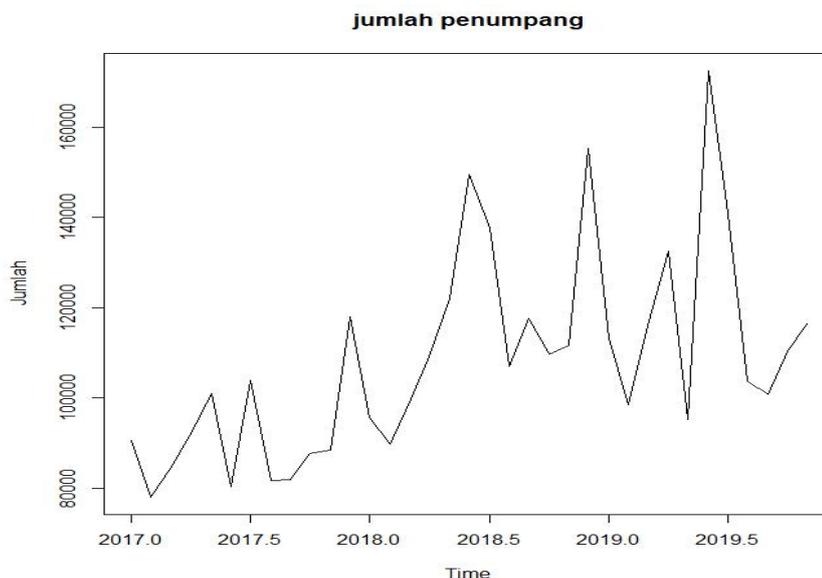
7. Merumuskan kesimpulan.

HASIL DAN PEMBAHASAN

Data yang digunakan pada penelitian ini adalah data jumlah penumpang kereta api di PT.KAI regional II Sumatera Barat. Data yang digunakan adalah data per bulan dari Januari 2017 sampai dengan April 2020 sebanyak 40 data, dimana data dibagi menjadi dua yaitu data *insample* dan *outsample*. Data *insample* yaitu Januari 2017 sampai dengan November 2019 data ini yang akan digunakan untuk menentukan model. Sedangkan data *outsample* yaitu dari Desember 2019 sampai dengan April 2020. Pada data akan diaplikasikan metode SARIMA menggunakan *software* statistik RStudio untuk mencari model terbaik dari jumlah penumpang kereta api di Sumatera Barat. Adapun langkah-langkah yang dilakukan adalah sebagai berikut.

Plot Data

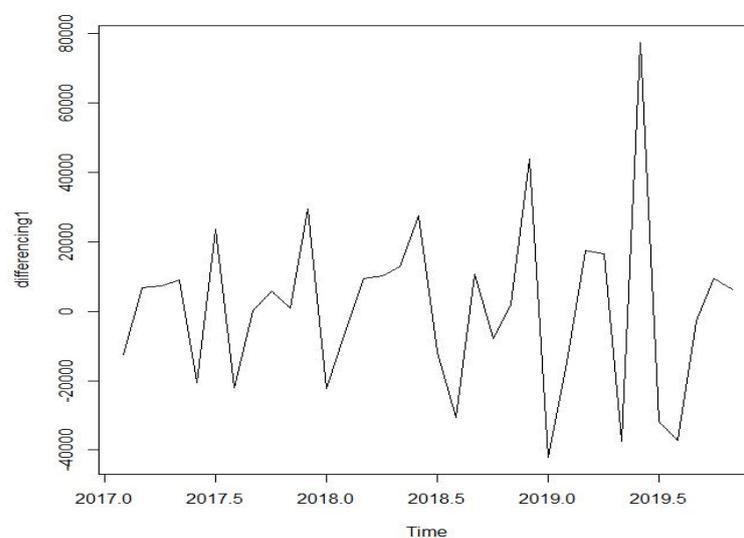
Plot data deret waktu digunakan untuk melihat secara kasat mata bentuk data penelitian apakah data sudah stasioner terhadap *mean* dan variansi serta melihat apakah data mengandung unsur musiman atau tidak. Saat data tidak stasioner terhadap *mean* maka dilakukan *differencing* sedangkan jika data tidak stasioner terhadap variansi maka dilakukan transformasi. Berikut hasil plot data awal menggunakan *software* statistik RStudio.



Gambar 1. Plot Deret waktu Data Aktual Jumlah Penumpang Kereta Api Bulan Januari 2017-November 2019

Pada Gambar 1, menunjukkan data bulanan dari bulan Januari 2017 sampai dengan November 2019 karena data tersebut yang digunakan untuk mendapatkan model. Berdasarkan Gambar 2 data menunjukkan adanya *trend* dan musiman karena data mengalami kenaikan dan terjadi pola yang berulang. Pola musiman dapat dilihat dari pola yang diulang dimana data akan menunjukkan naik dan turun dalam jangka waktu yang tetap.

Selanjutnya data jumlah penumpang kereta api dilakukan *differencing* satu kali. Hal ini dilakukan agar data stasioner terhadap nilai tengah (mean). Berikut hasil plot data setelah dilakukan *differencing* satu kali.



Gambar 2. Plot Data Differencing Satu Kali Non Musiman

Berdasarkan Gambar 2, setelah dilakukan *differencing* satu kali terlihat bahwa data tidak ada lagi yang mengandung unsur *trend* dan sudah berfluktuasi di sekitar nilai tengah. Hal ini menunjukkan data telah stasioner terhadap nilai tengah.

Uji Stasioneritas Data

Kestasioneran suatu data juga dapat diperiksa melalui uji ADF. Berikut hasil uji ADF pada data jumlah penumpang kereta api di Sumatera Barat.

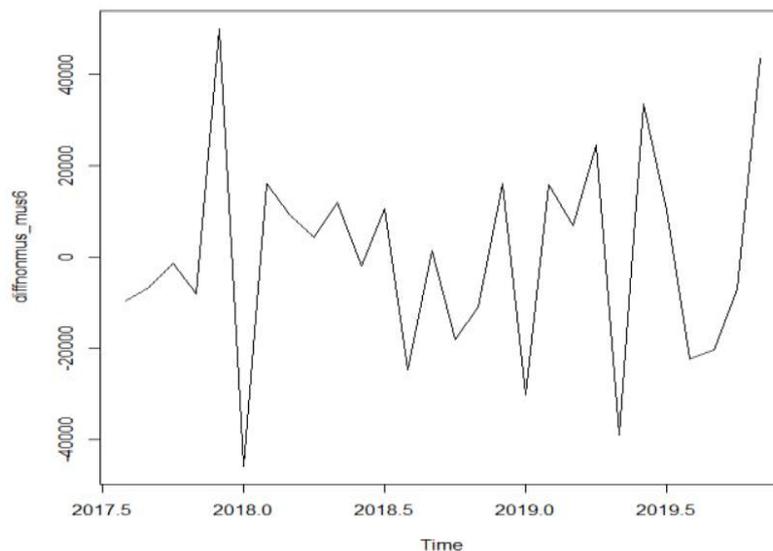
Tabel 1. Hasil Uji ADF Jumlah Penumpang Kereta Api di Sumatera Barat Bulan Januari 2017-November 2019 Setelah Differencing Satu Kali

	t-Statistic	probabilitas
Augmented Dickey-Fuller test statistic	-6,273776	0,0000
Test critical values	1% level	-3,679322
	5% level	-2,967767
	10% level	-2,622989

Berdasarkan uji ADF pada Tabel 1 diperoleh nilai statistik sebesar -6,273776. Nilai kritis atau t_{tabel} 5% sebesar -2,967767 dengan nilai $\alpha = 0.05$. kestasioneran data

disimpulkan dengan membandingkan nilai t_{stat} dengan t_{tabel} . Sehingga dapat disimpulkan bahwa nilai statistik $|ADF| > |t_{tabel}|$ sebesar $6,273776 > 2,967776$ dengan $P - value$ sebesar 0 terlihat bahwa ADF lebih besar dari nilai kritis tabel dan $P - value < 5\%$. Maka data telah stasioner terhadap nilai tengah atau tolak H_0 dan terima H_1 . Karena data telah stasioner pada *differencing* pertama, maka diperoleh orde non musiman $d = 1$.

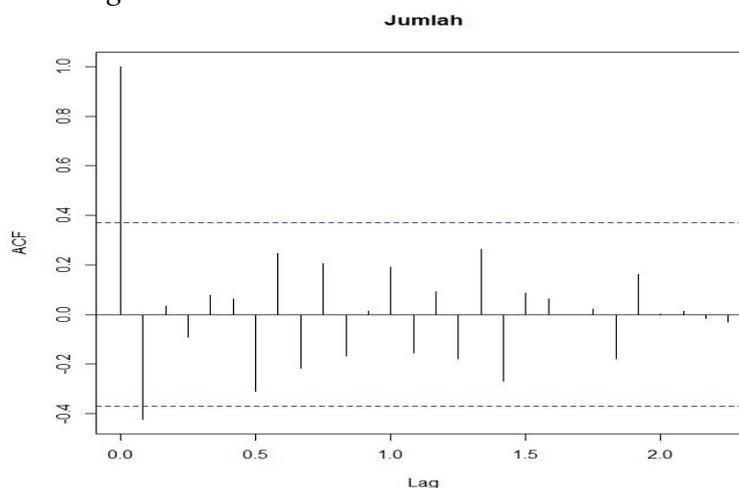
Langkah selanjutnya adalah membuat *differencing* musiman terhadap data jumlah penumpang kereta api untuk menstasionerkan data. Karena data jumlah penumpang kereta api di Sumatera Barat memiliki komponen musiman dan data tersebut merupakan bulanan yang memiliki peningkatan jumlah penumpang setiap 6 bulan terakhir, maka dilakukan *differencing* sebanyak satu kali musiman menggunakan orde musiman $s = 6$. Berikut plot hasil *differencing* 1 kali orde musiman 6.



Gambar 3. Plot Data Hasil *Differencing* Satu Kali Musiman 6

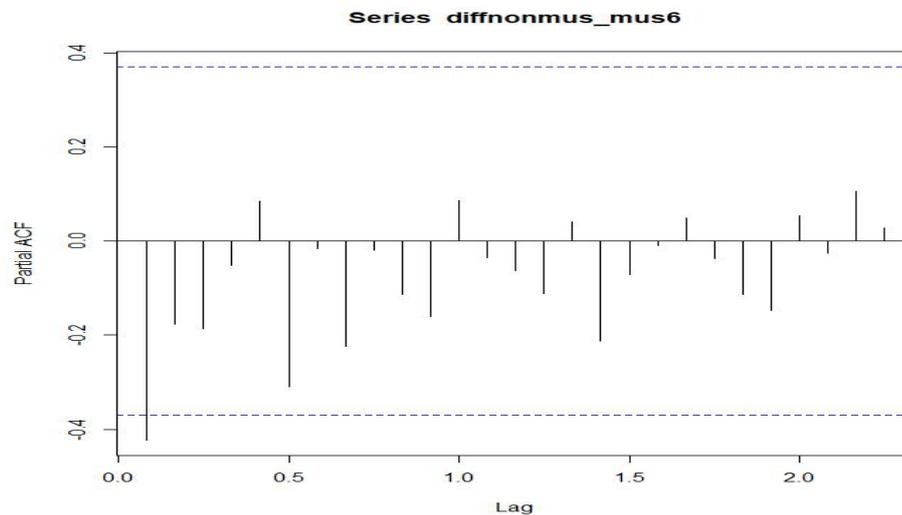
Identifikasi Model dan Dugaan Parameter Model

Identifikasi model SARIMA dilakukan dengan melihat orde non musiman dan musiman $(p,d,q)(P,D,Q)$ dari *correlogram* ACF dan PACF *differencing* musiman. Berikut dipeoleh *correlogram* ACF dan PACF musiman.



Gambar 4. Correlogram ACF Differencing Satu Kali Musiman 6

Gambar 4, menunjukkan bahwa *correlogram* ACF data *cuts off* setelah lag 1 sehingga MA(1) untuk non musiman dan MA(0) untuk musiman karena, lag 6 tidak melewati batas kendali signifikansi.



Gambar 5. Correlogram PACF Differencing Satu Kali Musiman 6

Gambar 5 menunjukkan bahwa *correlogram* PACF data *cuts off* setelah lag 1 sehingga AR(1) untuk non musiman dan AR(0) untuk musiman karena, lag 6 tidak melewati batas kendali signifikansi. Berdasarkan *correlogram* ACF dan PACF data *differencing* musiman sebanyak 1 kali, sehingga diperoleh orde $D = 1$. Dengan demikian, model dugaan awal untuk musiman yang terbentuk adalah SARIMA $(1,1,1)(0,1,0)^6$, $(0,1,1)(0,1,0)^6$, $(1,1,0)(0,1,0)^6$. Langkah selanjutnya adalah mengestimasi masing-masing parameter model.

Tabel 2. Estimasi Parameter Masing-Masing Model

No	Model	Parameter	Dugaan	P-value	Uji signifikansi
1	SARIMA(1,1,1)(0,1,0) ⁶	AR(1)	0,27669	0,1405	Tidak Signifikan
		MA(1)	-1,00000	1,57E-10	Signifikan
2	SARIMA(0,1,1)(0,1,0) ⁶	MA(1)	-0,6938	0,002925	Signifikan
3	SARIMA(1,1,0)(0,1,0) ⁶	AR(1)	-0,47262	0,007154	Signifikan

Uji Asumsi Residual

Selanjutnya pemeriksaan uji asumsi residual yang terdiri atas *white noise* dan distribusi normal terhadap model yang signifikan. Berikut tabel hasil yang memenuhi uji *white noise* dan berdistribusi normal.

Tabel 3. Uji Asumsi White Noise dan Distribusi Normal

No	Model	Parameter	White Noise	Distribusi Normal	Keterangan
1	SARIMA(1,1,1)(0,1,0) ⁶	AR(1)	0,6961	0,8259	Bersifat White Noise dan berdistribusi Normal
		MA(1)			

2	SARIMA(0,1,1)(0,1,0) ⁶	MA(1)	0,8491	0,5233	Bersifat White Noise dan berdistribusi Normal
3	SARIMA(1,1,0)(0,1,0) ⁶	AR(1)	0,4861	0,264	Bersifat White Noise dan berdistribusi Normal

Tabel 3, menunjukkan bahwa hasil uji asumsi *white noise* dan distribusi normal pada model SARIMA (1,1,1)(0,1,0)⁶, SARIMA (0,1,1)(0,1,0)⁶ dan SARIMA (1,1,0)(0,1,0)⁶ diperoleh *P – value* dari masing-masing model yaitu lebih besar dari taraf signifikansi $\alpha = 0,05$. Tolak H_0 apabila *P – value* Ljung-Box $< \alpha$ dengan $\alpha = 0,05$ dengan kata lain, ketiga model tersebut telah memenuhi uji asumsi *white noise*. Pada Tabel 3, juga diperoleh bahwa ketiga model berdistribusi normal karena *P – value* $> \alpha$. Sehingga, dapat disimpulkan bahwa ketiga model tersebut bersifat *white noise* dan berdistribusi normal.

Pemilihan Model Terbaik

Pemilihan model terbaik dilakukan dengan melihat nilai AIC dari masing-masing model, dimana model dengan nilai AIC terkecil adalah model yang paling baik. Berikut tabel nilai AIC beberapa model yang telah dipilih.

Tabel 4. Nilai AIC dari Masing-Masing Model SARIMA

No	Model	Parameter	AIC
1	SARIMA(1,1,1)(0,1,0) ⁶	AR(1) MA(1)	636,55
2	SARIMA(0,1,1)(0,1,0) ⁶	MA(1)	635,89
3	SARIMA(1,1,0)(0,1,0) ⁶	AR(1)	638,49

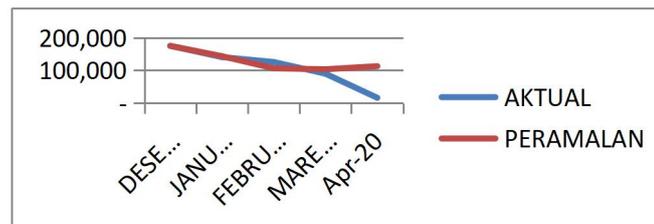
Pada Tabel 4, diperoleh nilai AIC yang paling kecil adalah model SARIMA (0,1,1)(0,1,0)⁶ sehingga, model tersebut dipilih sebagai model terbaik dan layak digunakan untuk peramalan.

Peramalan

Perbandingan hasil peramalan dan data aktual dapat dilihat pada Tabel 5. Secara visual perbandingan data tersebut dapat dilihat pada Gambar 6.

Tabel 5. Data Jumlah Penumpang Kereta Api Bulan Desember 2019-April 2020

Bulan	Data Aktual	Peramalan
Desember 2019	175.128	175.197
Januari 2020	140.667	143.430
Februari 2020	125.619	106.267
Maret 2020	89.823	103.421
April 2020	15.983	112.846



Gambar 6. Plot Data Aktual dan Peramalan

Pada Tabel 5, diperoleh ramalan jumlah penumpang kereta api di Sumatera Barat. Berdasarkan model terbaik SARIMA $(0,1,1)(0,1,0)^6$, diperkirakan bahwa jumlah penumpang kereta api tertinggi terjadi pada pertengahan dan akhir tahun. Gambar 6, menunjukkan bahwa hasil peramalan telah mendekati nilai sebenarnya.

SIMPULAN

Pengolahan data yang dilakukan pada data jumlah penumpang kereta api di PT. KAI (Persero) Divisi Regional II Sumatera Barat dari bulan Januari 2017 sampai dengan April 2020 dengan menggunakan metode *Seasonal Autoregressive Integrated Moving Average* (SARIMA) diperoleh model terbaik SARIMA $(0,1,1)(0,1,0)^6$. Model ini terpilih karena telah memenuhi uji signifikansi, *white noise*, distribusi normal dan memiliki nilai AIC terkecil. Model SARIMA $(0,1,1)(0,1,0)^6$ memiliki persamaan sebagai berikut

$$Z_t = \mu + (Z_{t-1} - \mu) + (Z_{t-6} - \mu) - (Z_{t-7} - \mu) + e_t - \theta_1 e_{t-1} \quad (16)$$

DAFTAR PUSTAKA

- [1] G. E. P. and Box, *Time Series Analysis Forecasting and Control*.
- [2] F. Dunia, W. Abdullah, and C. Sasongko, *int t S en*. 2019.
- [3] D. M. Putri and Aghsilni, "Estimasi Model Terbaik Untuk Peramalan Harga Saham PT. Polychem Indonesia Tbk Dengan Arima," *MAp J.*, vol. 1, no. 1, pp. 1-12, 2019.
- [4] A. Assidiq, P. Hendikawati, and N. Dwidayati, "Perbandingan Metode Weighted Fuzzy Time Series, Seasonal Arima, dan Holt-Winter's Exponential Smoothing untuk Meramalkan Data Musiman," *Indones. Gedung D7 Lt.1, Kampus Sekaran Gunungpati*, vol. 6, no. 2, pp. 129-142, 2017, [Online]. Available: <http://journal.unnes.ac.id/sju/index.php/ujm>
- [5] M. Munira Anwar, Khalilah Nurfadilah, and Wahidah Alwi, "Penerapan Metode SARIMA untuk Peramalan Jumlah Pengunjung Wisata Taman Nasional Bantimurung Bulusaraung Maros," *J. Math. Theory Appl.*, vol. 3, no. 1, pp. 1-7, 2021, doi: 10.31605/jomta.v3i1.1221.
- [6] A. Agustin, F. F. Rahani, and F. I. Indikawati, "Prediksi Kualitas Air Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA)," *J. Manaj. Inform.*, vol. 12, no. 2, pp. 137-150, 2022, doi: 10.34010/jamika.v12i2.8022.
- [7] W. Rahmalina, "Pemodelan Seasonal Autoregressive Integrated Moving Average Untuk Memprediksi Jumlah Kasus Covid-19 di Padang," *J. Mat. Integr.*, vol. 17, no. 1, p. 23, 2021, doi: 10.24198/jmi.v17.n1.32024.23-31.
- [8] M. I. Rizki and T. A. Taqiyyuddin, "Penerapan Model SARIMA untuk Memprediksi Tingkat Inflasi di Indonesia," *J. Sains Mat. dan Stat.*, vol. 7, no. 2, pp. 62-72, 2021, doi: 10.24014/jsms.v7i2.13168.
- [9] P. Utomo and A. Fanani, "Peramalan Jumlah Penumpang Kereta Api di Indonesia Menggunakan Metode Seasonal Autoregressive Integrated Moving Average

- (SARIMA)," *J. Mhs. Mat. Algebr.*, vol. 1, no. 1, pp. 169–178, 2020, [Online]. Available: <http://jurnalsaintek.uinsby.ac.id/mhs/index.php/algebra/article/view/6/6>
- [10] S. C. Hillmer and W. W. S. Wei, "Time Series Analysis: Univariate and Multivariate Methods.," *J. Am. Stat. Assoc.*, vol. 86, no. 413, p. 245, 1991, doi: 10.2307/2289741.
- [11] D. N. Gujarati, *Single-equation regression models*. 2013.
- [12] W. D. Ray, P. J. Brockwell, and R. A. Davis, *Time Series: Theory and Methods.*, vol. 153, no. 3. 1990. doi: 10.2307/2982983.