PERAMALAN JUMLAH PENUMPANG DI STASIUN KEDIRI MENGGUNAKAN MODEL ARIMA

Nabila Lintang Ardani [1*], Kartika Maulida Hindrayani[2], Shindi Shella May Wara[3]
[1,2,3]Sains Data, Universitas Pembangunan Nasional "Veteran" Jawa Timur
*Email Korespondensi : 22083010105@student.upnjatim.ac.id

Abstrak

Peramalan jumlah penumpang kereta api memiliki peranan penting dalam mendukung perencanaan operasional dan peningkatan kualitas layanan transportasi publik. Penelitian ini bertujuan untuk memprediksi jumlah penumpang kereta api di Stasiun Kediri menggunakan model Autoregressive Integrated Moving Average (ARIMA). Data yang digunakan merupakan data historis jumlah penumpang dengan periode tertentu yang kemudian dianalisis melalui tahapan uji stasioneritas, identifikasi pola menggunakan fungsi autokorelasi (ACF) dan autokorelasi parsial (PACF), estimasi parameter, serta pemilihan model optimal berdasarkan kriteria performa model. Hasil analisis menunjukkan bahwa model terbaik yang terpilih adalah ARIMA (5,1,5) dengan nilai Mean Absolute Percentage Error (MSE) sebesar 0,05% dan Root Mean Squared Error (RMSE) sebesar 0,20, yang mengindikasikan tingkat akurasi peramalan berada dalam kategori reasonable forecast. Model tersebut kemudian digunakan untuk melakukan peramalan jumlah penumpang kereta api pada periode Januari 2024 hingga September 2025. Hasil peramalan menunjukkan adanya kecenderungan peningkatan jumlah penumpang pada periode mendatang. Dengan demikian, model ARIMA (5,1,5) dapat digunakan sebagai alat bantu dalam pengambilan keputusan terkait penyediaan fasilitas, penjadwalan perjalanan, serta strategi peningkatan pelayanan transportasi di Stasiun Kediri.

Kata kunci: ARIMA, Jumlah Penumpang, Kereta Api, Peramalan, Stasiun Kediri

Abstract

Forecasting the number of railway passengers plays an essential role in supporting operational planning and improving public transportation service quality. This study aims to predict the number of passengers at Kediri Station using the Autoregressive Integrated Moving Average (ARIMA) model. Historical passenger data were analyzed through several stages, including stationarity testing, pattern identification using Autocorrelation Function (ACF) and Partial Autocorrelation Function (PACF), parameter estimation, and model selection based on performance criteria. The analysis results indicate that the best-performing model is ARIMA (5,1,5) with a Mean Absolute Percentage Error (MAPE) of 0,05% and Root Mean Squared Error (RMSE) sebesar 0,20, which falls under the category of a reasonable forecasting accuracy. The selected model was then utilized to forecast the number of passengers for the period from January 2024 to September 2025. The forecast results demonstrate a potential upward trend in passenger volume in the upcoming periods. Therefore, the ARIMA (5,1,5) model can be used as a decision support tool to enhance operational planning, facility provision, and service improvement in railway transportation at Kediri Station.

Keywords: Additive Outlier, ARIMA, Number of Passengers, Train, Forecasting, Kediri Station

PENDAHULUAN

Transportasi memiliki peran yang sangat penting dalam mendukung pertumbuhan ekonomi, mobilitas penduduk, serta pemerataan akses antar wilayah (Nugraha & Santoso, 2025). Salah satu moda transportasi publik yang banyak digunakan di Indonesia adalah kereta api (Rahmatunnisa et al., 2021). Selain memiliki kapasitas angkut yang besar, kereta api juga menawarkan tingkat keamanan dan ketepatan waktu yang relatif tinggi dibandingkan moda transportasi darat lainnya (Kusumamingrum & Tohir, 2025). Seiring dengan peningkatan infrastruktur jaringan perkeretaapian yang dilakukan pemerintah, permintaan terhadap moda transportasi ini terus mengalami peningkatan setiap tahunnya (Suryadi, 2014). Stasiun Kediri merupakan salah satu stasiun yang berada di wilayah Daerah Operasi (Daop) 7 Madiun dan memiliki peran strategis dalam melayani arus transportasi masyarakat di Kota Kediri dan daerah sekitarnya. Aktivitas perjalanan kereta api di stasiun ini mencerminkan mobilitas masyarakat pada berbagai sektor seperti pekerjaan, pendidikan, perdagangan, hingga pariwisata. Namun demikian, jumlah penumpang yang dilayani Stasiun Kediri tidak selalu stabil, melainkan berfluktuasi seiring perubahan kondisi sosial ekonomi, hari besar keagamaan, libur nasional, hingga kebijakan pemerintah seperti penyesuaian tarif dan peningkatan jadwal perjalanan. Fluktuasi jumlah penumpang tersebut menimbulkan tantangan dalam pengambilan keputusan operasional (Kanaya et al., 2025). Peramalan atau forecasting jumlah penumpang diperlukan untuk mendukung pengelolaan sarana dan prasarana perkeretaapian, seperti penentuan jadwal keberangkatan, jumlah rangkaian kereta yang dioperasikan, alokasi sumber daya manusia, serta peningkatan fasilitas pelayanan di stasiun. Informasi yang akurat mengenai kondisi di masa mendatang menjadi dasar penting agar pelayanan tetap optimal dan efisien.

Analisis deret waktu (*time series analysis*) merupakan pendekatan yang banyak digunakan dalam memodelkan data berdasarkan pola historis. Salah satu metode yang populer dan efektif dalam peramalan jangka pendek adalah Autoregressive Integrated Moving Average (ARIMA). Model ARIMA mampu menangani data non-stasioner melalui proses diferensiasi dan memanfaatkan komponen autoregresif serta moving average untuk memberikan prediksi yang lebih akurat (Kruba et al., 2025a). Selain itu, ARIMA memiliki fleksibilitas dalam mengakomodasi pola musiman maupun non-musiman dalam data deret waktu (Kruba et al., 2025b). Pada penelitian ini, model ARIMA diterapkan untuk memprediksi jumlah penumpang kereta api di Stasiun Kediri. Model terbaik dipilih berdasarkan evaluasi performa melalui perhitungan kesalahan prediksi, salah satunya *Mean Square Error* (MSE). Hasil peramalan diharapkan dapat memberikan gambaran tren jumlah penumpang di masa mendatang dan menjadi bahan pertimbangan dalam perencanaan strategis perkeretaapian. Dengan demikian, penelitian ini memiliki urgensi dalam upaya meningkatkan kualitas layanan transportasi kereta api, terutama dalam mendukung pemenuhan kebutuhan mobilitas masyarakat yang semakin kompleks. Selain itu, temuan penelitian ini juga diharapkan dapat menjadi acuan bagi pengembangan penelitian lebih lanjut dengan penerapan metode peramalan yang lebih kompleks dan variabel pendukung yang lebih beragam.

METODE PENELITIAN

Penelitian ini dilaksanakan dengan menggunakan metode kuantitatif yang berfokus pada analisis data numerik untuk menguji hipotesis dan merumuskan prediksi berbasis statistik. Pendekatan kuantitatif dipilih karena tujuan utama penelitian adalah peramalan atau *forecasting*, yang membutuhkan ketepatan matematis dan evaluasi kinerja model menggunakan metrik statistik objektif seperti RMSE dan MAE.

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini diperoleh dari PT Kereta Api Indonesia (KAI) melalui pihak pengelola Stasiun Kediri. Data yang dikumpulkan merupakan data harian jumlah penumpang yang tercatat pada periode Januari 2024 hingga September 2025, dengan total sebanyak 639 data observasi. Data tersebut dipilih karena memiliki kontinuitas yang baik serta mencerminkan kondisi operasional harian secara aktual. Informasi ini menjadi dasar dalam analisis deret waktu untuk membangun model peramalan jumlah penumpang secara akurat.

B. Uji Stasioneritas

Dalam analisis data runtun waktu, stasioneritas merupakan syarat penting agar proses permodelan dapat menghasilkan estimasi parameter yang stabil dan akurat. Suatu data dikatakan stasioner jika nilai rata-rata, varians, dan kovarians antar waktu tidak berubah sepanjang periode pengamatan (Aktivani, 2020). Ketidakstasioneran, khususnya akibat adanya tren atau akar unit (unit root), dapat menyebabkan hasil peramalan menjadi bias dan tidak reliabel (Wijoyo, 2016). Oleh karena itu, diperlukan uji formal untuk memastikan apakah suatu data telah memenuhi sifat stasioner.

Salah satu metode yang paling umum digunakan untuk menguji keberadaan akar unit adalah Augmented *Dickey–Fuller* (ADF) Test (Rizal & Akbar, 2015). Uji ADF merupakan pengembangan dari *Dickey–Fuller* Test yang memasukkan komponen lag dari selisih variabel sehingga dapat mengatasi autokorelasi dalam residual. Secara umum, model ADF dapat dituliskan sebagai berikut:

$$\Delta Y_t = \alpha + \beta t + \gamma Y_{t-1} + \sum_{i=1}^{p} \delta_i \Delta Y_{t-i} + \varepsilon_t \quad (1)$$

Pada persamaan tersebut, ΔY_t merupakan differencing dari variabel, α adalah intercept, βt mewakili komponen tren, sementara γ menjadi parameter utama yang menentukan ada tidaknya akar unit pada data. Istilah δ_i adalah koefisien lag yang digunakan untuk menghilangkan autokorelasi, dan ε_t merupakan *error* yang diasumsikan white noise. Keputusan ditentukan dengan membandingkan nilai statistik ADF dengan nilai kritis *Dickey–Fuller*. Apabila nilai statistik lebih kecil daripada nilai kritis, maka H0 ditolak dan data dianggap stasioner. Sebaliknya, jika H0 tidak dapat ditolak, data dinyatakan tidak stasioner dan perlu dilakukan transformasi seperti differencing atau detrending untuk memenuhi syarat kestasioneran.

C. Identifikasi Model

a. Autocorrelation Function (ACF)

Autocorrelation Function (ACF) digunakan untuk melihat pola hubungan antara suatu observasi dengan observasi lain pada lag tertentu (Kuswanto et al., 2025). ACF membantu mengidentifikasi struktur ketergantungan dalam data runtun waktu, khususnya untuk menentukan ordo komponen *Moving Average* (MA) (Pradana, 2025). Pola ACF dapat menunjukkan apakah data memiliki karakteristik cut off atau tailing off yang menjadi dasar pemilihan model dalam ARIMA.

$$r_k = \frac{\sum_{t=k+1}^{n} (Y_t - \mu)(Y_{t-k} - \mu)}{\sum_{t=1}^{n} (Y_t - \mu)^2} \quad (2)$$

Plot ACF dianalisis untuk melihat pola cut off pada lag tertentu. Jika ACF terputus secara jelas pada lag ke-q, maka data cenderung mengikuti model MA(q). Sebaliknya, jika ACF meluruh perlahan tanpa pola cut off, kemungkinan struktur data lebih sesuai dengan model AR atau ARMA. Dengan demikian, ACF menjadi indikator awal dalam proses identifikasi model pada pendekatan *Box–Jenkins*.

b. Partial Autocorrelation Function (PACF)

Partial Autocorrelation Function (PACF) digunakan untuk mengukur hubungan langsung antara suatu observasi dan observasi pada lag tertentu, setelah efek lag-lag perantara dihilangkan. PACF sangat penting dalam menentukan ordo komponen Autoregressive (AR), karena nilai PACF menunjukkan korelasi "murni" yang tidak dipengaruhi oleh lag lain. PACF pada lag ke-k dapat dihitung menggunakan persamaan Yule–Walker atau regresi berganda. Secara umum, nilai PACF lag ke-k diperoleh dari model regresi:

$$\emptyset_{kk} = \frac{p_k - \sum_{j=1}^{k-1} \emptyset_{k-1,j} p_{k-j}}{1 - \sum_{j=1}^{k-1} \emptyset_{k-1,j} p_j} \quad (3)$$

PACF digunakan untuk melihat apakah terdapat cut off pada lag tertentu. Jika PACF terputus pada lag ke-p, maka data cenderung mengikuti model AR(p). Jika PACF meluruh perlahan tanpa pola cut off, maka model yang sesuai bisa jadi berstruktur MA atau ARMA. Dengan demikian, PACF melengkapi ACF dalam proses identifikasi sehingga penentuan ordo model ARIMA dapat dilakukan secara lebih akurat.

D. Metode ARIMA

Metode Autoregressive Integrated Moving Average (ARIMA) merupakan pendekatan statistik yang digunakan untuk memodelkan dan meramalkan data runtun waktu dengan mempertimbangkan pola ketergantungan antar observasi (Firmansyah, 2024). Model ini memadukan tiga komponen utama: autoregressive (AR), integrated (I), dan moving average (MA). Secara umum, ARIMA dinyatakan sebagai ARIMA(p, d, q), di mana p menunjukkan ordo *autoregressive*, d menunjukkan tingkat *differencing* untuk mencapai stasioneritas, dan q menunjukkan ordo *moving average*. Komponen AR merepresentasikan pengaruh nilai masa lalu terhadap nilai saat ini dan dirumuskan sebagai:

$$y_t = \emptyset_1 y_{t-1} + \cdots + \emptyset_p y_{t-p} + \varepsilon_t \quad (4)$$

Komponen differencing digunakan untuk menghilangkan ketidakstasioneran tren, dengan bentuk umum:

$$\Delta^d y_t = y_t - y_{t-1} \quad (5)$$

Sementara itu, komponen MA menangkap pengaruh *error* masa lalu terhadap nilai saat ini, dinyatakan sebagai:

$$y_t = \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (6)$$

Gabungan ketiga komponen tersebut menghasilkan model umum ARIMA:

$$\nabla^d Y_t = \emptyset_1 \nabla^d Y_{t-1} + \cdots + \emptyset_p \nabla^d Y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \tag{7}$$

Model ARIMA efektif digunakan pada data runtun waktu non-musiman yang menunjukkan pola linear, serta mampu memberikan performa yang baik ketika data telah memenuhi asumsi stasioneritas. Namun, model ini sensitif terhadap keberadaan outlier sehingga diperlukan penanganan khusus agar hasil peramalan tidak bias.

E. Evaluasi Model

Evaluasi model dilakukan untuk mengukur sejauh mana model peramalan mampu merepresentasikan pola data secara akurat (Yunianto & Muklason, 2025). Proses ini bertujuan membandingkan nilai hasil prediksi dengan data aktual sehingga dapat diketahui tingkat kesalahan atau *error* yang dihasilkan model. Dua ukuran error yang umum digunakan dalam analisis deret waktu adalah *Root Mean Square Error* (RMSE) dan *Mean Square Error* (MSE). Kedua metrik ini memberikan gambaran kuantitatif mengenai performa model, terutama terkait penyimpangan prediksi terhadap nilai observasi.

a. Root Mean Square Eror (RMSE)

Root Mean Square Error (RMSE) merupakan ukuran error yang menghitung akar dari rata-rata kuadrat selisih antara nilai aktual dan nilai hasil prediksi (Kasandra et al., 2024). RMSE memberikan penalti lebih besar pada kesalahan yang berukuran besar karena proses pengkuadratan error di dalam perhitungannya. Oleh sebab itu, RMSE efektif digunakan untuk mengevaluasi model yang sensitif terhadap outlier atau kesalahan ekstrem.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} \left(Y_t - \hat{Y}_t\right)^2} \tag{8}$$

Semakin kecil nilai RMSE, semakin tinggi akurasi model dalam melakukan prediksi. RMSE juga berada pada satuan yang sama dengan data asli, sehingga lebih mudah untuk diinterpretasikan dalam konteks permasalahan nyata (Farisi et al., 2024).

b. Mean Square Error (MSE)

Mean Square Error (MSE) merupakan ukuran error yang menggambarkan rata-rata dari kuadrat selisih antara nilai aktual dan nilai prediksi (Amansyah et al., 2024a). MSE digunakan secara luas karena perhitungannya sederhana dan mampu menunjukkan seberapa jauh model menyimpang dari data observasi. Seperti RMSE, MSE juga memberikan penalti yang lebih besar terhadap error besar akibat proses pengkuadratan.

$$MSE = \frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \tag{9}$$

Nilai MSE yang lebih kecil menunjukkan bahwa model memiliki kesalahan prediksi yang lebih rendah (Amansyah et al., 2024b). Meskipun tidak berada dalam satuan asli data seperti RMSE, MSE tetap merupakan indikator penting dalam membandingkan performa beberapa model atau konfigurasi parameter yang berbeda.

HASIL DAN PEMBAHASAN

A. Pengumpulan Data

Data yang digunakan dalam penelitian ini merupakan data harian jumlah penumpang kereta api yang diperoleh dari PT Kereta Api Indonesia (KAI) melalui pihak operasional Stasiun Kediri. Data mencakup periode 1 Januari 2024 hingga 30 September 2025 dan terdiri dari 639 observasi harian. Variabel yang dikumpulkan meliputi tanggal dan jumlah penumpang per hari.

Tanggal	Jumlah Penumpang
2024-01-01	1327
2024-01-02	1050
2024-01-03	1002
2024-01-04	803
2024-01-05	976
...	...
2025-09-26	1942
2025-09-27	1823
2025-09-28	2516
2025-09-29	1839
2025-09-30	1320

639 rows × 1 columns

Gambar 1 Dataset

Secara umum, data menunjukkan pola fluktuasi harian yang cukup tinggi. Pada awal periode pengamatan, jumlah penumpang berada pada kisaran 803 hingga 1.327 penumpang per hari. Sementara itu, pada akhir periode, jumlah penumpang dapat mencapai lebih dari 2.500 penumpang per hari, seperti yang terlihat pada tanggal 28 September 2025 dengan nilai 2.516 penumpang. Variasi ini menggambarkan adanya dinamika mobilitas penumpang yang dipengaruhi oleh hari kerja, akhir pekan, dan kemungkinan aktivitas khusus di wilayah Kediri yang memicu lonjakan ataupun penurunan jumlah penumpang.

B. Uji Stasioneritas

Uji stasioneritas menggunakan metode *Augmented Dickey-Fuller* (ADF) dilakukan kembali pada data yang telah melalui proses *differencing* pertama. Hasil pengujian menunjukkan bahwa nilai ADF Statistic sebesar −7.5378, dengan p-value sebesar 3.44×10^{-11}. Nilai p-value yang sangat kecil dan jauh di bawah batas signifikansi 1%, 5%, maupun 10% mengindikasikan penolakan kuat terhadap hipotesis nol yang menyatakan bahwa data mengandung akar unit.

Selain itu, nilai ADF Statistic (−7.5378) lebih kecil secara signifikan dibandingkan nilai critical values pada taraf signifikansi 1% (−3.4409), 5% (−2.8662), dan 10% (−2.5693). Kondisi ini menegaskan bahwa data hasil *differencing* pertama telah memenuhi asumsi stasioneritas. Dengan demikian, dapat disimpulkan bahwa data menjadi stasioner setelah dilakukan *differencing* pertama, sehingga nilai d yang digunakan dalam pemodelan ARIMA adalah d = 1. Proses identifikasi model ARIMA selanjutnya dilakukan berdasarkan pola ACF dan PACF dari data yang sudah stasioner tersebut.

C. Identifikasi Model

Proses identifikasi model ARIMA dilakukan dengan metode stepwise selection untuk memperoleh kombinasi parameter yang menghasilkan nilai *Akaike Information Criterion* (AIC) terendah. Hasil pencarian menunjukkan bahwa model dengan kinerja terbaik adalah ARIMA(5,1,5) dengan nilai AIC = 6670.552 (model tanpa intercept). Nilai AIC ini merupakan yang paling kecil dibandingkan model-model kandidat lainnya, sehingga ARIMA(5,1,5) dipilih sebagai model optimal.

D. Metode ARIMA

Model terbaik yang diperoleh dari proses identifikasi dan pemilihan menggunakan nilai AIC, BIC, serta evaluasi diagnostik adalah ARIMA(5,1,5). Estimasi parameter dilakukan menggunakan metode Maximum Likelihood Estimation (MLE). Ringkasan hasil estimasi ditunjukkan pada Gambar output model.

```
                               SARIMAX Results
==============================================================================
Dep. Variable:                      y   No. Observations:                  510
Model:               SARIMAX(5, 1, 5)   Log Likelihood               -3324.912
Date:                Thu, 04 Dec 2025   AIC                           6673.824
Time:                        00:12:11   BIC                           6724.613
Sample:                    01-02-2024   HQIC                          6693.738
                         - 05-25-2025
Covariance Type:                  opg
==============================================================================
                 coef    std err          z      P>|z|      [0.025      0.975]
------------------------------------------------------------------------------
intercept      4.3168      6.815      0.633      0.526      -9.041      17.675
ar.L1         -0.1929      0.112     -1.723      0.085      -0.412       0.026
ar.L2          0.4481      0.064      7.044      0.000       0.323       0.573
ar.L3         -0.6265      0.031    -19.959      0.000      -0.688      -0.565
ar.L4         -0.7843      0.057    -13.646      0.000      -0.897      -0.672
ar.L5          0.3545      0.106      3.338      0.001       0.146       0.563
ma.L1         -0.1718      0.105     -1.633      0.102      -0.378       0.034
ma.L2         -0.6333      0.052    -12.240      0.000      -0.735      -0.532
ma.L3          0.6230      0.054     11.458      0.000       0.516       0.730
ma.L4          0.5701      0.051     11.283      0.000       0.471       0.669
ma.L5         -0.6503      0.087     -7.476      0.000      -0.821      -0.480
sigma2       3.03e+04   1610.009     18.820      0.000    2.71e+04    3.35e+04
===================================================================================
Ljung-Box (L1) (Q):                   0.08   Jarque-Bera (JB):               174.18
Prob(Q):                              0.78   Prob(JB):                         0.00
Heteroskedasticity (H):               1.14   Skew:                             0.63
Prob(H) (two-sided):                  0.38   Kurtosis:                         5.58
===================================================================================
```

Gambar 2 Model ARIMA

Secara umum, sebagian besar parameter AR dan MA signifikan pada taraf 5%, ditunjukkan oleh nilai p-value yang sangat kecil (< 0.05), khususnya pada koefisien AR2, AR3, AR4, AR5, serta MA3, MA4, MA5. Parameter yang signifikan ini menunjukkan bahwa komponen *autoregressive* dan *moving average* memiliki peran penting dalam menangkap dinamika pola data jumlah penumpang. Nilai AIC = 6673.824, BIC = 6724.613, dan HQIC = 6693.738, yang menunjukkan bahwa model ini lebih optimal dibandingkan model-model lain yang memiliki nilai kriteria informasi lebih tinggi.

Uji *Ljung–Box* menunjukkan nilai Prob(Q) = 0.78, menandakan bahwa residual tidak memiliki autokorelasi dan dapat dianggap sebagai white noise. Hal ini menunjukkan bahwa model telah mampu menangkap pola dependensi dalam data. Uji heteroskedastisitas (H) menghasilkan Prob(H) = 0.38, sehingga tidak ditemukan masalah heteroskedastisitas yang signifikan pada residual. Namun, uji normalitas *Jarque–Bera* menunjukkan Prob(JB) = 0.00, sehingga residual tidak sepenuhnya berdistribusi normal. Meskipun demikian, dalam praktik deret waktu, ketidaknormalan residual masih dapat diterima terutama jika autokorelasi sudah hilang. Dengan terpenuhinya sebagian besar asumsi diagnostik, model ARIMA(5,1,5) dinyatakan layak digunakan sebagai model peramalan untuk jumlah penumpang.

E. Evaluasi Model

Kinerja model dievaluasi menggunakan metrik *Mean Squared Error* (MSE) dan *Root Mean Squared Error* (RMSE). Hasil evaluasi menunjukkan bahwa model menghasilkan MSE sebesar 0,05 dan RMSE sebesar 0,20. Secara teknis, nilai MSE yang sangat kecil menunjukkan bahwa rata-rata kuadrat selisih antara nilai observasi dan nilai hasil prediksi berada pada tingkat yang rendah. Metrik ini sensitif terhadap error berukuran besar, sehingga nilai MSE yang kecil mengindikasikan bahwa model tidak menghasilkan kesalahan prediksi ekstrem. Nilai RMSE sebesar 0,20 memberikan gambaran kuantitatif tentang besarnya error dalam satuan asli data sebelum transformasi. Dengan RMSE < 1, error prediksi dapat dikategorikan sangat rendah dibandingkan rentang variabilitas data penumpang (sekitar 800 hingga 2500 penumpang). Hal ini menunjukkan bahwa deviasi prediksi terhadap nilai aktual berada dalam skala yang sangat kecil dan tidak berdampak signifikan secara praktis. Kedua metrik tersebut menunjukkan konsistensi bahwa model ARIMA/SARIMA yang digunakan memiliki performa prediksi yang kuat, stabil, dan mampu menangkap pola pada data dengan baik tanpa menghasilkan residual yang besar.

KESIMPULAN

Berdasarkan hasil penelitian yang telah dilakukan, data jumlah penumpang menunjukkan pola fluktuatif dengan variasi yang cukup tinggi sehingga diperlukan pemodelan runtun waktu untuk menghasilkan prediksi yang akurat. Uji stasioneritas menggunakan *Augmented Dickey–Fuller* (ADF) menunjukkan bahwa data asli tidak stasioner, namun setelah dilakukan *differencing* pertama data menjadi stasioner, ditunjukkan oleh nilai ADF statistic sebesar −7,537 dan p-value $3{,}44\times10^{-11}$ yang berada jauh di bawah tingkat signifikansi 0,05. Proses identifikasi model melalui analisis ACF dan PACF serta pencarian otomatis berdasarkan kriteria AIC menghasilkan model terbaik berupa ARIMA(5,1,5). Hasil estimasi parameter menunjukkan bahwa sebagian besar koefisien AR dan MA signifikan pada tingkat kepercayaan 5%, menandakan model mampu menangkap pola dependensi dalam data secara tepat. Evaluasi performa model juga menunjukkan hasil yang sangat baik dengan nilai MSE sebesar 0,05 dan RMSE sebesar 0,20, yang berarti tingkat kesalahan prediksi sangat kecil dibandingkan skala data aktual. Secara keseluruhan, model ARIMA(5,1,5) dapat disimpulkan mampu merepresentasikan karakteristik data dengan baik serta memberikan akurasi prediksi yang tinggi sehingga layak digunakan sebagai dasar perencanaan dan pengambilan keputusan terkait operasional jumlah penumpang di masa mendatang.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Universitas Abdurachman Saleh Situbondo (UNARS) atas terselenggaranya Seminar Nasional & Call for Paper Series 4 yang menjadi wadah publikasi ilmiah dalam penelitian ini. Penghargaan yang sebesar-besarnya juga penulis sampaikan kepada PT Kereta Api Indonesia (Persero) Daerah Operasi Madiun, Stasiun Kediri, yang telah menyediakan data dan informasi yang diperlukan sehingga penelitian dapat dilaksanakan dengan baik. Ucapan terima kasih turut disampaikan kepada pimpinan program studi, dosen pembimbing, serta seluruh pihak yang telah memberikan dukungan, bimbingan, dan fasilitas selama proses penelitian hingga penyusunan artikel ini. Kontribusi dan kerja sama dari berbagai pihak sangat membantu terwujudnya penelitian ini secara optimal.

REFERENSI

Aktivani, S. (2020). Uji Stasioneritas Data Inflasi Kota Padang Periode 2014-2019. *Statistika*, *20*(2), 83–90.

Amansyah, I., Indra, J., Nurlaelasari, E., & Juwita, A. R. (2024a). Prediksi penjualan kendaraan menggunakan regresi linear: Studi kasus pada industri otomotif di indonesia. *Innovative: Journal Of Social Science Research*, *4*(4), 1199–1216.

Amansyah, I., Indra, J., Nurlaelasari, E., & Juwita, A. R. (2024b). Prediksi penjualan kendaraan menggunakan regresi linear: Studi kasus pada industri otomotif di indonesia. *Innovative: Journal Of Social Science Research*, *4*(4), 1199–1216.

Farisi, I., Shadiq, J., Priyadi, W., Maulana, D., Acep, A., & Gusril, S. F. (2024). Penerapan Model Recurrent Neural Network (RNN) untuk Prediksi Curah Hujan Berbasis Data Historis. *Information System for Educators and Professionals: Journal of Information System*, *9*(2), 217–226.

Firmansyah, A. F. E. (2024). Prediksi Harga Coin Cryptocurrency BNB Menggunakan Metode Auto Regressive Integrated Moving Average. *SUBMIT: Jurnal Ilmiah Teknologi Infomasi Dan Sains*, *4*(2), 17–21.

Kanaya, I. K. K., Sitanggang, B. E. I., & Setiawati, N. L. P. L. S. (2025). PERAMALAN KEDATANGAN KAPAL PESIAR MENGGUNAKAN METODE SINGLE EXPONENTIAL SMOOTHING PADA PT. PELABUHAN INDONESIA. *Jurnal Media Akademik (JMA)*, *3*(11).

Kasandra, P., Audila, V., Zainal, F., & Dermawan, A. A. (2024). Prediksi Nilai Emas Menggunakan Algoritma Regresi Linear. *Industrial Engineering System and Management Journal*, *5*(2), 152–165.

Kruba, R., Sofyan, H., Marshanda, D., & Syazana, N. (2025a). Peramalan Saham Indofood di Indonesia Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA). *Jurnal Manajemen Dan Keuangan*, *14*(1), 102–117.

Kruba, R., Sofyan, H., Marshanda, D., & Syazana, N. (2025b). Peramalan Saham Indofood di Indonesia Menggunakan Metode Seasonal Autoregressive Integrated Moving Average (SARIMA). *Jurnal Manajemen Dan Keuangan*, *14*(1), 102–117.

Kusumamingrum, S. P. A., & Tohir, M. (2025). Peran Kereta Api dalam Sistem Logistik Multimoda untuk Ekspor Produk Tekstil dari Semarang ke Pelabuhan Tanjung Perak. *Greenation Management and Business Review*, *1*(3), 105–112.

Kuswanto, H., Utomo, P. E. P., Khaira, U., & Waladi, A. (2025). Prediksi Nilai Ekspor Migas Indonesia Menggunakan Metode SARIMA dan LSTM. *SATESI: Jurnal Sains Teknologi Dan Sistem Informasi*, *5*(1), 69–79.

Nugraha, B., & Santoso, A. D. (2025). Peran Transportasi Multimoda dalam Meningkatkan Konektivitas dan Efisiensi Logistik di Pulau Jawa. *Innovative: Journal Of Social Science Research*, *5*(2), 1935–1948.

Pradana, B. L. (2025). Time Series Forecasting of LQ45 Stock Index Using ARIMA: Insights and Implications. *Review of Management, Accounting and Tourism Studies*, *1*(1), 27–40.

Rahmatunnisa, S. N., Utami, A., & Nurhidayat, A. Y. (2021). Probabilitas Perpindahan Penumpang Transportasi Massal Berbasis Rel (Studi Kasus Kereta Api Argo Parahyangan Terhadap Kereta Cepat Jakarta–Bandung). *Ge-STRAM: Jurnal Perencanaan Dan Rekayasa Sipil*, *4*(2), 91–96.

Rizal, J., & Akbar, S. (2015). Perbandingan Uji Stasioner Data Timeseries Antara Metode: Control Chart, Correlogram, Akar Unit Dickey Fuller, dan Derajat Integrasi. *Jurnal Gradien*, *11*(1), 1040–1046.

Suryadi, S. (2014). Kinerja Dan Peramalan Pertumbuhan Angkutan Kereta Api Menggunakan Model Sarima. *Warta Penelitian Perhubungan*, *26*(7), 381–394.

Wijoyo, N. A. (2016). Peramalan nilai tukar rupiah terhadap USD dengan menggunakan model GARCH. *Kajian Ekonomi Dan Keuangan*, *20*(2), 169–189.

Yunianto, R. M., & Muklason, A. (2025). Penerapan Metode Hybrid LSTM dan Seasonal Decomposition untuk Peramalan Permintaan Tiket Harian PT KAI. *ILKOMNIKA*, *7*(2), 1–14.